Biochimica et Biophysica Acta, 1171 (1993) 299–303

BBAEXP 90434

Short Sequence-Paper

A yeast gene (*BLH1*) encodes a polypeptide with high homology to vertebrate bleomycin hydrolase, a family member of thiol proteinases

Ulla Magdolen [a], Günter Müller [b], Viktor Magdolen [c] and Wolfhard Bandlow [a]

[a] *Institut für Genetik und Mikrobiologie, München (Germany),* [b] *c/o Hoechst Aktiengesellschaft Frankfurt am Main, Frankfurt (Germany) and* [c] *Max-Planck-Institut für Biochemie, Martinsried (Germany)*

(Received 28 September 1992)

Key words: Bleomycin hydrolase gene; Antitumor drug; Cysteine proteinase; (*S. cerevisiae*)

We have purified bleomycin hydrolase from yeast (molecular mass 55 000 Da). Using protein sequence-derived degenerate oligonucleotide primers and amplification by polymerase chain reaction, the yeast gene *BLH1* was isolated and characterized. The deduced amino acid sequence (483 amino acids) exhibits surprisingly high homology to vertebrate bleomycin hydrolase (43% identical residues and 22% conserved exchanges). It contains three blocks of sequences found conserved in other members of the thiol proteinase family and thought to be associated with the catalytic centre. *BLH1* is non-essential under all growth conditions tested. However, in the presence of 3.5 mg bleomycin/ml medium wild-type cells have a slight growth advantage compared to *blh1* mutant cells.

Bleomycin is a mixture of closely related glycopeptide antibiotics secreted from *Streptomyces verticillus* [1]. It causes radiomimetic DNA damage in bacteria and eukaryotes. It also inhibits growth of a number of squamous and small cell type tumors in situ and in vitro, although its exact mechanism of action is unknown. Due to its lack of hepatic, renal and bone marrow toxicities, the drug is extensively used in tumor therapy [1,2]. A more general use of the drug is, however, primarily hampered by two drawbacks, (i) the pulmonary toxicity [3] and (ii) the resistance of certain cell types and tissues to the drug due to high intracellular activities of the thiol proteinase bleomycin hydrolase [4–6]. The ubiquitous and abundant occurrence of this enzyme in all animal tissues and in microorganisms points to an important, although as yet unknown function. In order to study bleomycin resistance, to identify the physiological role of the proteinase and to understand why certain tumors are refractory to therapy by the drug, it may be helpful that we have isolated a yeast gene displaying high homology of the encoded protein to bleomycin hydrolase from rabbit.

We have isolated the bleomycin hydrolase protein from yeast during a study of a cAMP receptor glycolipid-anchored to plasma membranes [7]. We have enriched the latter protein by isolating crude plasma membranes from yeast spheroplasts by Percoll gradient centrifugation and sequential purification of the detergent-solubilized protein by cAMP-agarose affinity chromatograpy, molecular sieving and phenyl Sepharose lipophilic chromatography which uses the hydrophobicity of the glycolipid anchor-containing protein as a parameter for separation. When the eluted proteins were separated on a sodium dodecylsulfate polyacrylamide gel, one of the major bands comigrated with the authentic, 8-N_3-[^{32}P]cAMP photoaffinity-labeled product (apparent molecular mass 54 kDa, as deduced from electrophoretic mobility) [7]. This material was excised from the gel, eluted and cleaved either with trypsin or CNBr. The peptides were separated by high performance liquid chromatography. Two major fragments (one from each cleavage) were sequenced (indicated in Fig. 2). It turned out from the sequence data (see below) that they obviously do not belong to the glycosyl-phosphatidylinositol-anchored cAMP-binding protein from yeast plasma membranes. Whether the enriched protein with homology to bleomycin hydrolase binds to cAMP agarose and phenyl Sepharose directly or interacts with the native cAMP receptor protein

Correspondence to: W. Bandlow, Institut für Genetik und Mikrobiologie, Maria-Ward-Strasse 1a, D-8000 München 19, Germany.
The nucleotide sequence data reported in this paper have been submitted to the EMBL/Genbank Data Libraries under the accession number X69124.

fortuitously or in a functional complex remains to be established in the future.

The corresponding gene was isolated by using the polymerase chain reaction method. Since the intramolecular arrangement of the two eluted and sequenced peptides was unknown, three primers were designed from each of them, two overlapping primers in the one and a third one in the reverse direction. Pairwise amplification by polymerase chain reaction and crosswise hybridization led to the identification of a DNA fragment, which encoded sequences from both protein fragments. The genomic DNA clone, pPMY 11.0, subsequently isolated from a gene bank constructed from *Bam*HI-digested yeast DNA, contained about 11 kbp of genomic DNA and was characterized by restriction and hybridization analyses. The region cross-hybridizing to the amplified DNA fragment could be narrowed down to two *Hin*dIII fragments, which were sequenced subsequently from the central *Hin*dIII site in both directions to yield the complete DNA sequence of the yeast *BLH1* gene (Fig. 1A, B). The nucleotide sequence analysis and the deduced primary structure of the protein are displayed in Fig. 2. The open reading frame comprises 483 triplets with a deduced molecular mass of the encoded protein of 55.5 Da, the calculated isoelectric point is at pH 9.58. Structural analyses by using the algorithms of Kyte and Doolittle, and Hopp and Woods do not reveal any striking structural features and suggest a globular, hydrophilic protein. Southern hybridization with a labeled DNA fragment, derived from the coding region of the *BLH1* gene shows that *BLH1* occurs in single copy in the yeast genome (one example is given in Fig. 1D, lane 1). The codon bias index [8] of 0.35 points to a relatively high expression. The DNA sequence flanking the gene at its 5′ side displays no obvious TATA box. However, many yeast genes are known which lack obvious TATA sequences. In the putative promoter region of *BLH1* a perfect consensus sequence for an abundant DNA binding protein, Grf2p [9], is found, immediately followed by a pyrimidine-rich sequence (Fig. 2). Grf2p may have a role in nucleosome positioning by excluding nucleosomes from the vicinity of its binding site. The same combination of sequence motifs as in the *BLH1*-upstream region also occurs in the promoter region of the yeast profilin gene, *PFY1* [10]. Deletion of these elements in the *PFY1* promoter dramatically reduce transcription of *PFY1* (H. Herrmann and W. Bandlow, unpublished results).

Data base searches using the Genetics Computer Group (GCG) Sequence Analysis Software Package [11] reveal significant sequence similarities to bleomycin hydrolase, a cysteine proteinase isolated from rabbit lung and liver. A cDNA fragment encoding the central portion of this enzyme has been obtained from rabbit liver and sequenced [6]. The alignment of the deduced

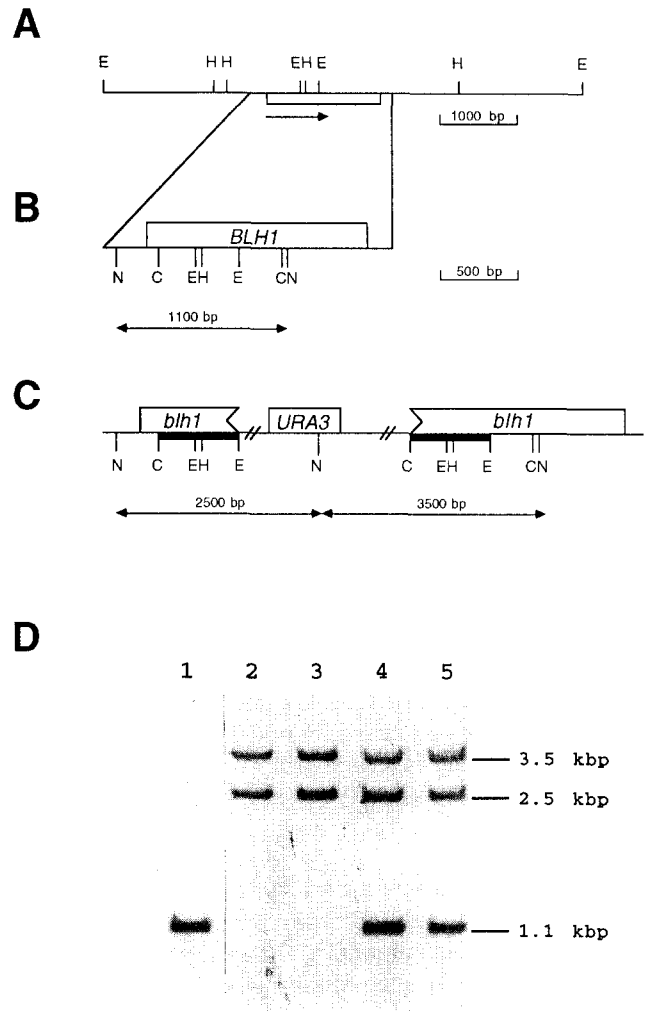

Fig. 1. Inactivation of the *BLH1* gene. (A) Genomic restriction map of the *BLH1* locus. The arrow indicates direction of transcription. (B) Restriction map of the sequenced region. Restriction sites relevant for gene disruption are shown. (C) Genomic organization after integration of the linearized plasmid pPMP9-A at the *BLH1* locus by homologous recombination. Vector sequences between the duplicated *BLH1*-coding sequences (indicated by thick bars) are not drawn to scale. (D) Genomic Southern blot from wild-type and disruption mutants. Genomic DNA was isolated from the haploid wild-type strain W303-1A (lane 1), from haploid disruption strains (lanes 2 and 3), from heterozygous diploid strains (lanes 4 and 5) and restricted with *Nde*I. The blots were probed with a digoxigenin-labeled fragment corresponding to the *BLH1* fragment cloned into pPMP9-A (indicated by thick bars in (C). C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; N, *Nde*I.

rabbit and yeast protein sequences is shown in Fig. 3A. Within the region of the known rabbit sequence, the homology is highly significant (40 S.D. units, using the program ALIGN [12] with the mutation data matrix, 100 random runs and a penalty for a gap of 20). The yeast sequence extends beyond the known rabbit sequence by 21 amino acids to the N-terminus and 158 residues to the C-terminus. The overlap of the two proteins contains 43% identical amino acids and 22% conservative exchanges at comparable positions (Fig.

3A). The search for protein motifs and domains reveals that a stretch of 17 amino acids (pos. 92–108 of Blh1p) is highly homologous to the region containing the cysteine residue conserved in the catalytic centre of eukaryotic thiol proteinases (Fig. 3B). These include calpain, plant and protozoan cystein proteinases, vertebrate lysosomal and secretory cathepsins, and two developmentally co-regulated cysteine proteinases from the slime mold *Dictyostelium discoideum*. In addition, it can be seen from Fig. 3B that Blh1p from yeast also has significant homology to two sequence blocks of other cysteine proteinases (which were previously unknown from the rabbit enzyme) implied to contribute to the active centre, but conserved only in a subset of thiol proteinases.

In order to gain some information about the physiological role of the protein, we replaced the structural *BLH1* gene with two different incomplete *blh1* alleles

```
-260                                                                                    GRF2
CGTATATTCA GTGCTTTTTG AAATTTAGAA AATAGTCGCT TTATTGATTA TGCTACATAT GTTATTTTGC TTTAGTATGT GTCAAATAGC ATGCAATCGC CGGGTAGTGA GTGACACATC CTTTTTTCCC

.....
CTCCCACCAG AAAGTCACCT GCTCTCTTAA TCAATGGGGT TGCATGTGGT AATCAATTGA TAACGCCGGC TGACAAGTCG CCGACGGGCA CCCATAAATA AACGATAAAT AGCCCTTGCT CTTTTGTTAC

+1
    ATG CTT CCT ACT TCT GTT TCC CGG TCC TTG TAT CTC AAG ACG TTT AGA AGC CAC CTT TTA CGA GCA CCA CAA ATT GTT TTA AAA AGA ATG TCC TCT TCC ATC
    Met Leu Pro Thr Ser Val Ser Arg Ser Leu Tyr Leu Lys Thr Phe Arg Ser His Leu Leu Arg Ala Pro Gln Ile Val Leu Lys Arg Met Ser Ser Ser Ile

+103
    GAT ATC AGT AAG ATC AAC TCT TGG AAC AAA GAG TTT CAA TCC GAC TTA ACC CAT CAA TTG GCA ACC ACT GTC CTT AAG AAT TAT AAT GCC GAT GAT GCA CTG
    Asp Ile Ser Lys Ile Asn Ser Trp Asn Lys Glu Phe Gln Ser Asp Leu Thr His Gln Leu Ala Thr Thr Val Leu Lys Asn Tyr Asn Ala Asp Asp Ala Leu

+205
    TTG AAC AAG ACT AGA CTG CAA AAG CAA GAT AAC AGG GTT TTC AAC ACT GTT GTC TCT ACT GAT TCC ACT CCA GTC ACC AAC CAA AAA AGC TCT GGT AGA TGT
    Leu Asn Lys Thr Arg Leu Gln Lys Gln Asp Asn Arg Val Phe Asn Thr Val Val Ser Thr Asp Ser Thr Pro Val Thr Asn Gln Lys Ser Ser Gly Arg Cys

+307
    TGG TTG TTT GCC GCT ACC AAT CAA CTG CGC TTG AAT GTT CTT TCT GAA TTG AAC TTA AAA GAA TTC GAA CTG TCC CAA GCT TAC TTG TTC TTT TAC GAT AAG
    Trp Leu Phe Ala Ala Thr Asn Gln Leu Arg Leu Asn Val Leu Ser Glu Leu Asn Leu Lys Glu Phe Glu Leu Ser Gln Ala Tyr Leu Phe Phe Tyr Asp Lys

+409
    TTG GAA AAG GCC AAC TAT TTC TTG GAC CAA ATC GTC TCC TCG GCT GAT CAA GAC ATC GAC TCA CGT CTC GTG CAA TAC TTG TTA GCA GCT CCA ACA GAA GAC
    Leu Glu Lys Ala Asn Tyr Phe Leu Asp Gln Ile Val Ser Ser Ala Asp Gln Asp Ile Asp Ser Arg Leu Val Gln Tyr Leu Leu Ala Ala Pro Thr Glu Asp

+511
    GGT GGC CAA TAC AGC ATG TTT TTG AAC CTA GTC AAG AAG TAT GGC CTT ATG CCC AAA GAT TTA TAC GGG GAT CTA CCT TAT TCT ACT ACT GCG TCC AGA AAA
    Gly Gly Gln Tyr Ser Met Phe Leu Asn Leu Val Lys Lys Tyr Gly Leu Met Pro Lys Asp Leu Tyr Gly Asp Leu Pro Tyr Ser Thr Thr Ala Ser Arg Lys

+613
    TGG AAT TCT CTG TTG ACT ACT AAA CTG AGA GAA TTT GCC GAG ACT CTA AGA ACA GCT TTG AAA GAG CGT TCT GCC GAT GAT TCC ATA ATT GTC ACT CTG AGA
    Trp Asn Ser Leu Leu Thr Thr Lys Leu Arg Glu Phe Ala Glu Thr Leu Arg Thr Ala Leu Lys Glu Arg Ser Ala Asp Asp Ser Ile Ile Val Thr Leu Arg

          primer 1-A
+715 ------------------------>
    GAG CAA ATG CAA AGA GAA ATC TTC AGG TTG ATG TCG TTG TTC ATG GAC ATA CCT CCA GTG CAA CCA AAC GAG CAA TTC ACT TGG GAA TAC GTT GAC AAA GAC
    Glu Gln Met Gln Arg Glu Ile Phe Arg Leu Met Ser Leu Phe Met Asp Ile Pro Pro Val Gln Pro Asn Glu Gln Phe Thr Trp Glu Tyr Val Asp Lys Asp

+817
    AAG AAA ATC CAC ACT ATC AAA TCG ACT CCG TTA GAA TTT GCC TCC AAA TAC GCA AAA TTG GAC CCT TCC ACG CCA GTC TCA TTG ATC AAT GAT CCA AGA CAC
    Lys Lys Ile His Thr Ile Lys Ser Thr Pro Leu Glu Phe Ala Ser Lys Tyr Ala Lys Leu Asp Pro Ser Thr Pro Val Ser Leu Ile Asn Asp Pro Arg His

+919
    CCA TAT GGT AAA TTA ATT AAG ATC GAT CGT TTA GGA AAC GTC CTT GGC GGA GAT GCC GTG ATT TAC TTA AAT GTT GAC AAT GAA ACA CTA TCT AAA TTG GTT
    Pro Tyr Gly Lys Leu Ile Lys Ile Asp Arg Leu Gly Asn Val Leu Gly Gly Asp Ala Val Ile Tyr Leu Asn Val Asp Asn Glu Thr Leu Ser Lys Leu Val

+1021                                                                                  <------
    GTT AAG AGA TTA CAA AAT AAC AAA GCT GTC TTT TTT GGA TCT CAC ACT CCA AAG TTC ATG GAC AAG AAA ACT GGT GTC ATG GAT ATT GAA TTG TGG AAC TAT
    Val Lys Arg Leu Gln Asn Asn Lys Ala Val Phe Phe Gly Ser His Thr Pro Lys Phe Met Asp Lys Lys Thr Gly Val Met Asp Ile Glu Leu Trp Asn Tyr

    primer 2-A          primer 2-C
    ----------          <----------
    CCT GCC ATC GGC TAT AAT TTA CCT CAG CAA AAG GCA TCG CGT ATT AGA TAC CAT GAA AGT TTG ATG ACT CAT GCT ATG TTG ATC ACT GGC TGC CAC GTC GAT
    Pro Ala Ile Gly Tyr Asn Leu Pro Gln Gln Lys Ala Ser Arg Ile Arg Tyr His Glu Ser Leu Met Thr His Ala Met Leu Ile Thr Gly Cys His Val Asp

+1225
    GAA ACG TCT AAA TTA CCA CTT CGC TAC CGC GTT GAA AAT TCC TGG GGT AAA GAC TCC GGT AAA GAC GGA TTA TAC GTG ATG ACT CAA AAG TAC TTC GAG GAG
    Glu Thr Ser Lys Leu Pro Leu Arg Tyr Arg Val Glu Asn Ser Trp Gly Lys Asp Ser Gly Lys Asp Gly Leu Tyr Val Met Thr Gln Lys Tyr Phe Glu Glu

+1327
    TAC TGC TTT CAA ATT GTG GTC GAT ATC AAT GAA TTG CCA AAA GAG CTG GCT TCA AAA TTC ACC TCA GGT AAG GAA GAG CCG ATT GTC TTG CCC ATC TGG GAC
    Tyr Cys Phe Gln Ile Val Val Asp Ile Asn Glu Leu Pro Lys Glu Leu Ala Ser Lys Phe Thr Ser Gly Lys Glu Glu Pro Ile Val Leu Pro Ile Trp Asp

+1429
    CCA ATG GGT GCT TTG GCC AAA TAA  ATAGTT TCAGCAGCTC TGATGTAGAT ACACGTATCT CGACATGTTT TATTTTTACT ATACATACAT AAAAAGAAAT AAAAAATGAT AACGTGTATA
    Pro Met Gly Ala Leu Ala Lys ***
```

Fig. 2. Nucleotide and deduced amino acid sequences of *BLH1*. The putative Grf2p-binding site is overlined, the adjacent pyrimidine-rich region is indicated by dots. The two peptides sequenced from the purified protein are underlined. The DNA regions corresponding to the primers effective in PCR amplification are overlined by arrows. The sequences of these degenerate oligonucleotides were: 5′ CAA/GATGCAA/GC/AGI/CGAA/GAT (1-A); 5′ CCIATI/CGCI/CGGA/GTAA/GTT (2-A); 5′ TGT/CTGI/CGGI/CAA/GA/GTTA/GTA (2-C). I denotes inosine.

as detailed in Fig. 1C. For this purpose the coding sequence contained between the *Cla*I site at pos. +100 and the *Eco*RI site at pos. +615 was cloned into the yeast integration vector pRS306 harbouring the *URA3* gene as a selective marker [13] to yield clone pPMP9-A. After cleavage of pPMP9-A with *Hin*dIII (pos. +383 of the coding region of *BLH1*) the resulting linear plasmid DNA was transformed into either haploid or diploid yeast wild-type cells. The correct integration of the plasmid by homologous recombination at the *BLH1* locus leads to the duplication of those *BLH1* DNA sequences cloned into pRS306, the resulting two incomplete copies being separated by sequences from the vector (Fig. 1C). The integration event was proven by Southern blotting of genomic DNA restricted with *Nde*I. The blot was probed with the *BLH1*-derived sequences cloned into pPMP9-A. DNA from wild-type cells yields a single signal corresponding to 1.1 kbp (Fig. 1D, lanes 1), whereas the haploid disruption mutants display two signals as expected, corresponding to 2.5 and 3.5 kbp (Fig. 1D, lanes 2 and 3) instead of the 1.1 kbp band (cf. Fig. 1B and C). Heterozygous diploids contain both, the wild-type and the disrupted alleles. Consequently, all three signals can be seen with DNA from heterozygous diploid disruption strains (Fig. 1D, lanes 4 and 5).

blh1 disruption mutants were found to be viable in the haploid state. When the growth behaviour of the haploid mutant and the isogenic wild-type strains was compared, no differences were observed on glucose, galactose and glycerol media. Also long-term experiments, where cells were shifted from one carbon source to another (eight changes of media), revealed no disadvantage for the *blh1* mutant strains compared to wild-type cells. In addition, heterozygous diploids sporulate well, and the spores germinate. It must be concluded from these results that the gene is dispensable for vegetative and meiotic growth, cell viability and differentiation. Finally, the haploid wild-type strain (with a functional *BLH1* allele) was compared to the disruption strain with respect to growth in the presence of bleomycin. For this purpose co-cultivation experiments of the haploid strain W303–1A and of the isogenic *blh1* mutant strain were performed: cells of both strains pre-grown overnight in the presence of bleomycin were mixed and incubated for 2 days in the absence of 3.5 mg bleomycin/ml medium (which leads to about 50% growth inhibition). The ratio of the wild-type and mutant cells was monitored by replica-plating at the beginning of the experiment and after two days of growth by testing uracil auxotrophy (*URA3* has been used as the selective marker to inactivate *BLH1*). The initial

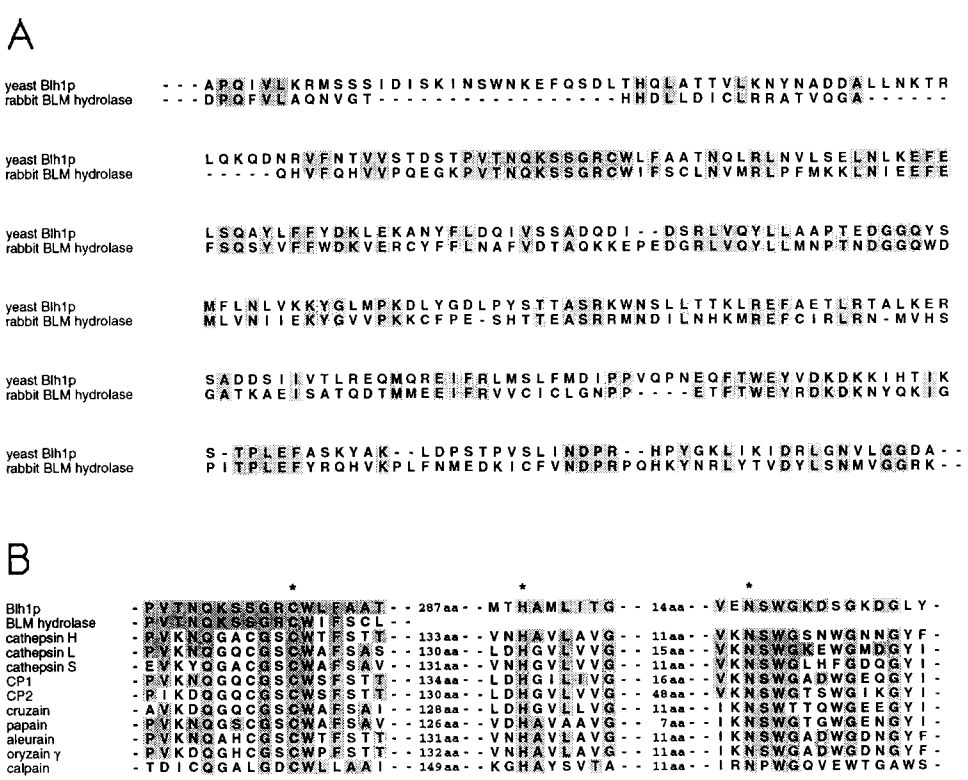

Fig. 3. Comparison of the deduced amino acid sequence of BLH1 to other cysteine proteinases: (A) residues identical to the known sequence of rabbit bleomycin hydrolase [6] at corresponding positions are shaded. (B) Alignment of Blh1p with other cysteine proteinases in the regions containing the putative active-site cysteine, histidine and asparagine (indicated by stars) [14]. Amino acids identical to Blh1p are shaded. Sequences were taken from [6,14–19].

ratio of 55% *blh1* mutant and 45% wild-type cells changed after two days (i.e., 26 generations) to 27% mutant and 73% wild-type cells, indicating a slightly higher sensitivity of the mutant cells to the drug. In contrast, the ratios of both types of cells did not change significantly in the absence of bleomycin (55%/45% versus 56%/44% *blh1/BLH1* cells). These results make it likely that *BLH1*, in fact, encodes a cysteine proteinase capable of hydrolyzing bleomycin. The lack of a detectable phenotype in *blh1* mutant cells in the absence of bleomycin as well as the only slightly increased sensitivity to bleomycin suggest that yeast may contain other proteinases with overlapping substrate specificities in addition to bleomycin hydrolase.

We thank E.-M. Wetekam for excellent technical assistence. The work was supported by a grant of the Deutsche Forschungsgemeinschaft to W.B. (Ba415/19-4).

References

1 Umezawa, H., Maeda, K., Takeuchi, T. and Akami, Y. (1966) J. Antibiot. Ser. A 19, 200–209.
2 Povirk, L.F. (1983) In Molecular Aspects of Anti-Cancer Drug Action (Neidle, S. and Waring H., eds.), pp. 157–181, Macmillan Publisher, London.
3 Lazo, J.S. and Humphreys, C.J. (1983) Proc. Natl. Acad. Sci. USA 80, 3064–3068.
4 Akiyama, S.I., Ikezaki, K., Kuramochi, H., Takahashi, K. and Kuwano, M. (1981) Biochem. Biophys. Res. Commun. 101, 55–60.
5 Lazo, J.S., Sebti, S.M. and Filderman, A.E. (1987) In Metabolism and Mechanism of Action of Anti-Cancer Drugs (Powis, G. and Prough, R.A., eds.), pp. 194–210, Taylor and Francis, London.
6 Sebti, S.M., Mignano, J.E., Jani, J.P., Srimatkandada, S. and Lazo, J.S. (1989) Biochemistry 28, 6544–6548.
7 Müller, G. and Bandlow, W. (1991) Biochemistry 30, 10181–10190.
8 Bennetzen, J.L. and Hall, B.D. (1982) J. Biol. Chem. 257, 3026–3031.
9 Buchman, A.R. and Kornberg, R.D. (1990) Mol. Cell. Biol. 10, 887–897.
10 Magdolen, V., Oechsner, U., Müller, G. and Bandlow, W. (1988) Mol. Cell. Biol. 8, 5108–5115.
11 Devereux, J., Haeberli, P. and Marquess, P. (1984) Nucleic Acids Res. 12, 387–395.
12 George. D.G., Barker, W.C. and Hunt, L.T. (1986) Nucleic Acids Res. 14, 11–17.
13 Sikorski, R.S. and Hieter, P. (1989) Genetics 122, 19–27.
14 Eakin, A.E., Mills, A.A., Harth, G., McKerrow, J.H. and Craik, C.S. (1992) J. Biol. Chem. 267, 7411–7420.
15 Higgins, D.G., McConnell, D.J. and Sharp, P.M. (1989) Nature 340, 604.
16 Watanabe, H., Abe, K., Emori, Y., Hosoyama, H. and Arai, S. (1991) J. Biol. Chem. 266, 16897–16902.
17 Ritonja, A., Colic, A., Dolenc, I., Ogrinc, T., Podobnik, M. and Turk, V. (1991) FEBS Lett. 283, 329–331.
18 Pears, C.J., Mahbubani, H.M. and Williams, J.G. (1985) Nucleic Acids Res. 13, 8853–8866.
19 Ohno, S., Emori, Y., Imajoh, S., Kawasaki, H., Kisaragi, M. and Suzuki, K. (1984) Nature 312, 566–570.