FEBS Letters 346 (1994) 151–155

FEBS 14068

Cloning and expression of a human pro(tea)some β-subunit cDNA: a homologue of the yeast PRE4-subunit essential for peptidylglutamyl-peptide hydrolase activity

Will L.H. Gerards*, Frank W.H. Hop, Ine L.A.M. Hendriks, Hans Bloemendal

Department of Biochemistry, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands

Received 30 March 1994; revised version received 22 April 1994

Abstract

The cDNA encoding a human prosome β-subunit (HSBpros26) was isolated from a lymphoma library using the cDNA of the *Xenopus* homologue as a probe. The cDNA contains an open reading frame encoding a protein of 233 amino acids and a calculated molecular weight of 25,909. Comparison with interspecies homologues of HSBpros26 from *Xenopus* (XLB), rat (RN3) and yeast (PRE4) reveals a high degree of identity between the β-subunits except for the N-terminal end, which is probably cleaved post-translationally. The complete coding sequence of HSBpros26 has been expressed in *E. coli*. The produced protein of about 27 kDa reacts with the prosomal monoclonal antibody MCP205, kindly provided by Dr. K. Hendil. The molecular weight of the native protein is about 28 kDa indicating that the protein is present as monomers. Finally partially purified HSBpros26 preparations do not contain any proteolytical activity.

Key words: Prosome; Proteasome; Multicatalytic proteinase; β-Subunit; cDNA cloning; Expression

1. Introduction

Prosomes are ubiquitous, cylindrical multicatalytic proteinase complexes with a molecular mass of 720 kDa consisting of 15–25 different subunits. These complexes, which are found in all eukaryotic cells analyzed to date, are thought to be important in non-lysosomal protein breakdown [1–2]. In some Archaebacteria a comparable particle is present containing only two different subunits named α and β [3]. cDNA cloning and analysis of several eukaryotic subunits revealed that they can be subdivided into two related families of α-type and β-type respectively, according to the homology with the archaebacterial α- or β-subunit [4,4a]. Between different eukaryotes very homologous subunits of over 95% sequence identity were found [5]. For higher eukaryotic prosomes little detailed information is available on the properties of single subunits, which could give information on subunit–subunit interaction and thus assembly and subunit composition of the particle. Expression of cDNAs in *E. coli* [6] is a way to produce large amounts of prosome subunits. The α- and β-subunits of *Thermoplasma acidophilum* were readily synthesized in *E. coli* and simultaneous expression of both subunits resulted in the formation of functional prosomes [7].

Here we report the cloning and expression in *E. coli* of the human β-subunit (HSBpros26) cDNA in order to enable the study of the properties of this subunit in more detail.

2. Materials and methods

2.1. Screening procedure and analysis of the HSBpros26 cDNA

A lymphoma cDNA library in the λZAP (Stratagene) was screened by Southern hybridization at 60°C with the complete *Xenopus laevis* β-subunit (XLB) cDNA [8] as a probe. Recombinant pBluescript SK-phagemids were rescued from positive bacteriophage clones by in vivo excision according to the procedure provided by the manufacturer. 5'-Ends of the cDNA's in this vector were sequenced using single-stranded DNA and the KiloBase Sequencing System (Biorad; Fig. 1, hatched arrows). The longest HSB clone (number 17) was completely sequenced (Fig. 1, filled arrows) using double-stranded DNA and Sequenase (United States Biochemical). Blotting procedures, radiolabeling of probes and subcloning were performed using standard methods [9]. Homology searches and characterization of the open reading frame were performed with the CAOS/CAMM computer facility of the University of Nijmegen, which includes Fasta, Clustal V, Motifs and Pepstats [10].

2.2. Expression of the HSB protein

The HSBpros26 cDNA was expressed in the pET expression system of Studier et al. [6]. To subclone the complete open reading frame of the HSB properly in the pET3c expression vector we mutated the ATG startcodon to generate a *Nde*I restriction site. This was achieved by site-directed mutagenesis using the oligonucleotide-directed in vitro mutagenesis system (Amersham) and the mutagenic primer (Eurogentec):

5'-CTCCCGATTCCTCCATGGATCCGGC-3'
$\overline{\quad\quad}$
*Nde*I

The mutated *Nde*I–*Xho*I HSB cDNA fragment was cloned into the *Nde*I–*Bam*HI sites of the pET3c vector and transformed to *E. coli* BL21(DE3). For production of the HSB protein, bacteria were grown in LB medium containing 100 μg/ml ampicillin. At OD_{550} 0.6 expression was induced by adding IPTG to a final concentration of 1 mM. Bacteria were harvested 4 h after induction by centrifugation at 5,000 × *g* for 20 min. After resuspension of the pellet in lysis buffer, 50 mM Tris-HCl, pH 7.6, 1 mM EDTA, 0.3 mM NaCl bacteria were lysed with chicken

*Corresponding author. Fax: (31) (80) 54 05 25.

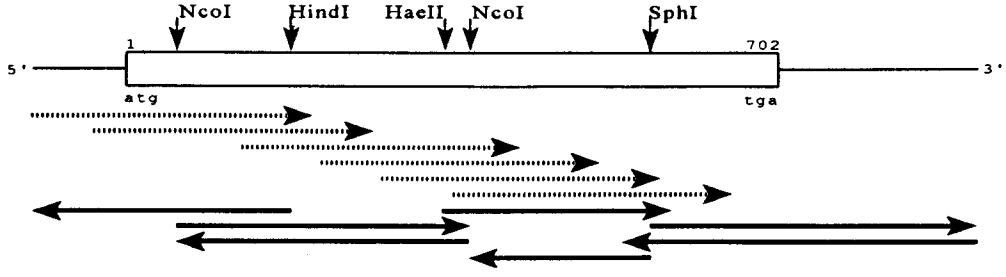
Abbreviations: PGHP, peptidylglutamyl-peptide hydrolase; HSB, Homo sapiens prosome β-subunit; Cbz, benzoyloxycarbonyl; NA, β-naphtylamide; MNA, 4-methoxy-β-naphtylamide; AMC, 7-amido-4-methylcoumarin; WSF, water-soluble fraction.

SSDI 0014-5793(94)00454-4

egg white lysozyme (Sigma). The water-soluble fraction (WSF) containing most of the HSB protein was obtained after centrifugation at 10,000 × *g* for 30 min. Expression of the HSB protein was analyzed by SDS-PAGE and Western blot. The molecular weight of the native HSB protein present in the bacterial WSF was determined using HPLC gel filtration on a Superdex 75 column (Pharmacia). Proteolytical activity of the HPLC-sample containing HSB (1 μg) was determined with the fluorogenic peptides Cbz-LLE-NA, benzoyl-FVR-MNA and succinyl-LLVY-AMC according to [11].

3. Results and discussion

The screening of $1 \cdot 10^6$ plaques with the XLB-cDNA resulted in the isolation of 24 clear positives including 14 clones containing HSB cDNAs of different lengths. Partial 5′-sequences of these cDNAs and the total sequence of clone 17, the longest cDNA containing the complete coding sequence, were obtained. The overall homology of XLB and HSB cDNAs turned out to be 66%. The HSB nucleotide sequence and the deduced amino acid sequence are shown in Fig. 1. The nucleotide at position −19 is variable, in two of three clones analyzed a guanidine and in one an adenosine was found. A partial sequence in the EMBL-EST database encoding the 5′-end of HSB (EM_EST:T08798) [12], identical to the 5′-end of our clone, also has an adenosine at that position. This polymorphism reflects most likely the existence of two functionally identical HSB alleles. The open reading frame of 233 amino acids encodes a protein with a molec-




```
                                                                gaattcggcacgaggttttt -80
ggggtcgcggtccggactttgggcgggggtccggccccaggacagttttaccgcattccgtccactcccgattccttc  -1

atg gat ccg gcg tct gca ctt tac aga ggt cca atc acg cgg acc cag aac ccc atg gtg   60
MET Asp Pro Ala Ser Ala Leu Tyr Arg Gly Pro Ile Thr Arg Thr Gln Asn Pro Met Val  20

acc ggg acc tca gtc ctc ggc gtt aag ttc gag ggc gga gtg gtg att gcc gca gac atg 120
Thr Gly Thr Ser Val Leu Gly Val Lys Phe Glu Gly Gly Val Val Ile Ala Ala Asp Met  40

ctg gga tcc tac ggc tcc ttg gct cgt ttc cgc aac atc tct cgc att atg cga gtc aac 180
Leu Gly Ser Tyr Gly Ser Leu Ala Arg Phe Arg Asn Ile Ser Arg Ile Met Arg Val Asn  60

aac agt acc atg ctg ggt gcc tct ggc gac tac gct gat ttc cag tat ttg aag caa gtt 240
Asn Ser Thr Met Leu Gly Ala Ser Gly Asp Tyr Ala Asp Phe Gln Tyr Leu Lys Gln Val  80

ctc ggc cag atg gtg att gat gag gag ctt ctg gga gat gga cac agc tat agt cct aga 300
Leu Gly Gln Met Val Ile Asp Glu Glu Leu Leu Gly Asp Gly His Ser Tyr Ser Pro Arg 100

gct att cat tca tgg ctg acc agg gcc atg tac agc cgg cgc tcg aag atg aac cct ttg 360
Ala Ile His Ser Trp Leu Thr Arg Ala Met Tyr Ser Arg Arg Ser Lys Met Asn Pro Leu 120

tgg aac acc atg gtc atc gga ggc tat gct gat gga gag agc ttc ctc ggt tat gtg gac 420
Trp Asn Thr Met Val Ile Gly Gly Tyr Ala Asp Gly Glu Ser Phe Leu Gly Tyr Val Asp 140

atg ctt ggt gta gcc tat gaa gcc cct tcg ctg gcc act ggt tat ggt gca tac ttg gct 480
Met Leu Gly Val Ala Tyr Glu Ala Pro Ser Leu Ala Thr Gly Tyr Gly Ala Tyr Leu Ala 160

cag cct ctg ctg cga gaa gtt ctg gag aag cag cca gtg cta agc cag acc gag gcc cgc 540
Gln Pro Leu Leu Arg Glu Val Leu Glu Lys Gln Pro Val Leu Ser Gln Thr Glu Ala Arg 180

gac tta gta gaa cgc tgc atg cga gtg ctg tac tac cga gat gcc cgt tct tac aac cgg 600
Asp Leu Val Glu Arg Cys Met Arg Val Leu Tyr Tyr Arg Asp Ala Arg Ser Tyr Asn Arg 200

ttt caa acc gcc act gtc acc gaa aaa ggt gtt gaa ata gag gga cca ttg tct aca gag 660
Phe Gln Thr Ala Thr Val Thr Glu Lys Gly Val Glu Ile Glu Gly Pro Leu Ser Thr Glu 220

acc aac tgg gat att gcc cac atg atc agt ggc ttt gaa tga aatacagatgcattatccagaac 725
Thr Asn Trp Asp Ile Ala His Met Ile Ser Gly Phe Glu ***                         233

tgaagttgccctacttttaactttgaacttggctagttcaaagatagactcttcttttgtaaagtaaataaattcttca 804
aaatgaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcgag                       860
```

Fig. 1. Nucleotide sequence and predicted amino acid sequence of human prosome β-subunit (lower panel). Underlined is the amino acid sequence of HSBpros26 obtained by direct protein sequencing [12]. Double underlined is the consensus sequence for β-type prosome subunits. The putative signal for polyadenylation (AATAAA) is overlined. The upper panel shows the sequence strategy and restricion sites used herein. Numbers indicate the start and stop nucleotide of the coding region (box). Hatched and filled arrows indicate 5′-sequenced regions of different clones and sequenced regions of clone 17, respectively.

ular mass of 26 kDa (hence the indication HSBpros26) and a pI of 5.8. The putative product contains only the consensus sequence for the family of prosomal β-type subunits, defined in the Motifs program, and the N-terminal sequence of the human β-subunit determined by direct protein sequencing [13]. This suggests that the N-terminal part of the primary product is cleaved resulting in a mature protein containing 220 amino acids, with a molecular mass of 24,5 kDa and a pI of 5.8. Post-translational cleavage of the N-terminus also occurs in the rat homologue of HSB, RN3 [14], β-subunit of *Thermoplasma* [7] and probably most β-type subunits. Northern blot analysis of poly(A)$^+$ RNA derived from human melanoma cell lines 530 and M14 [15] using the HSB cDNA as a probe revealed one single band of about 1.04 kb indicating that HSB clone 17 is nearly full length (not shown).

Homology searches in the EMBL nucleotide sequence database resulted in the identification of the partial sequence Mmorfa (M74556), most likely of the mouse homologue of HSBpros26. This cDNA sequence of 166 nucleotides shows 88.6% identity to nucleotide 78–243 of HSB. The authors assign the start of the coding sequence to nucleotide 111. However, upon our deduction of the amino acid sequence, starting at nucleotide 2 of Mmorfa and applying 2 insertions (Fig. 2, positions 81 and 90) we found at least 89% identity and 96% similarity with HSB. In view of this very high homology we assume that two nucleotides at positions 81 and 90 in the published sequence are missing (Fig. 2). Our evaluation shows that the complete sequence of the fragment (166 nt and not 55 nt) is coding for part of the mouse homologue of HSB.

In a phylogenetic tree, based on the Clustal V multiple alignment of all available prosomal amino acid sequences in the PIR and SWISSPROT databases, HSB, RN3, XLB and yeast PRE4 [16] formed a separate subbranch within the β-subunit branch, indicating that they are interspecies homologues (not shown). Interestingly it has been reported that the yeast PRE4 subunit is essential for PGHP activity [16]. Multiple alignment of the four sequences (Fig. 3) clearly reveals the high degree of conservation between this type of subunit of different species, which has also been shown for other subunits [5,17]. RN3, XLB and PRE4 are, respectively, 93%, 81% and 45% identical to HSBpros26. The leader is not conserved in sequence and length, e.g. XLB seems to have no leader at all, suggesting that this track is not important for assembly or any other function. The sequence surrounding the potential nuclear targeting signal RRxK, as proposed previously [16] and which is conserved in yeast and *Xenopus* is also highly conserved in rat and human (Fig. 3), but the functionality hereof remains to be established.

The pET expression system turned out to be a good system for high level expression of HSB. Fig. 4 shows that, after induction with IPTG, bacteria containing the HSB expression vector for 4 h, the HSB protein is expressed at higher levels than the bacterial proteins (lane 3). Western blot analysis was performed with the monoclonal antibody MCP205 that reacts with the human equivalent of rat RN3 (Kristensen et al., in preparation). Fig. 4 shows that only in bacteria with the HSB cDNA yields staining of a 27 kDa band (lanes 7, 8 and 9). This correlates well with calculated molecular mass of the primary HSB protein. In a purified liver prosome preparation MCP205 reacts with a protein of approximately

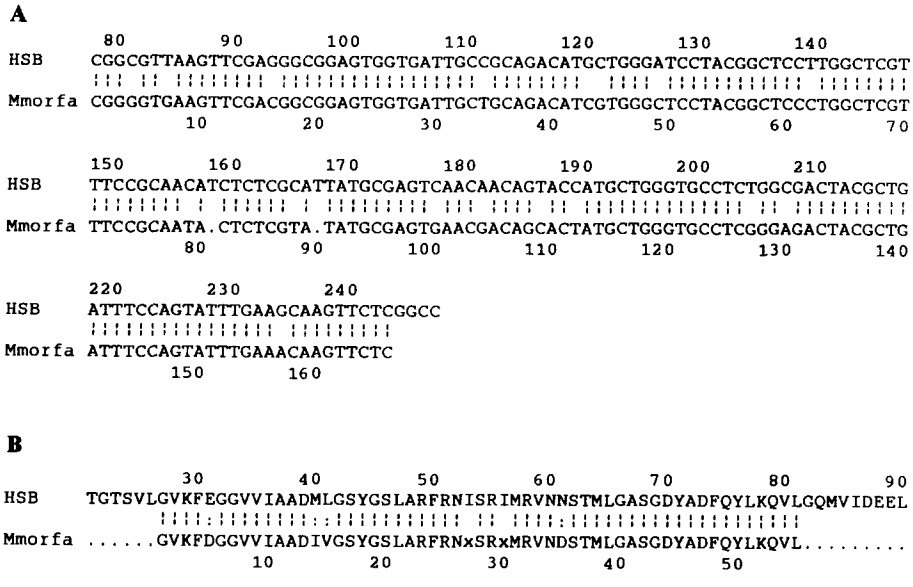


Fig. 2. Alignment of the nucleic acid (A) and amino acid (B) sequence of HSBpros26 and Mmorfa. The amino acid sequence was deduced from the nucleic acid sequence starting at position 2. The two codons containing the gaps (.), introduced in the nucleic acid sequence to obtain maximal homology, are translated as 'x'.

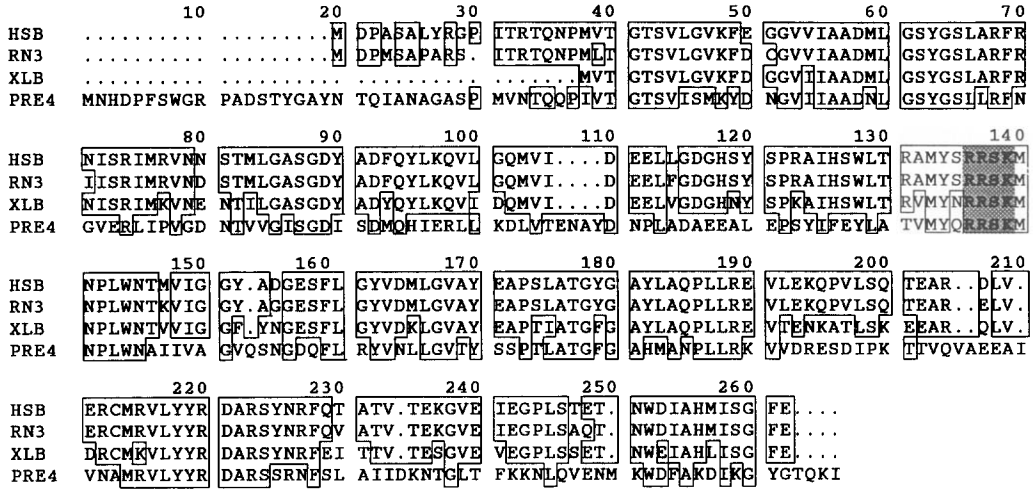


Fig. 3. Alignment of the amino acid sequences of HSB homologues. Boxed residues are identical in at least two sequences. Gaps (.) have been introduced to obtain maximal homology. The potential nuclear targeting signal (RRxK) comprises amino acids 136–139 (shaded).

25 kDa (lane 10), most likely representing the processed calf homologue of HSB. However, we do not know the identity of the reacting protein of about 57 kDa (lanes 3 and 4).

Most of the HSBpros26 protein was water-soluble. To determine the molecular mass of native HSBpros26 we applied bacterial WSF on a Superdex 75 gel filtration column. All HSB protein eluted just before chymotrypsinogen A (25 kDa), indicating that HSB is present as monomers. This suggests that in the prosome complex one HSB subunit is not associated to another one of the same type or that such an association is prevented by the leader sequence. The HPLC fraction containing the partially purified HSBpros26 protein was tested for proteolytical activity using the test peptides Cbz-LLE-NA, benzoyl-FVR-MNA and succinyl-LLVY-AMC. No activity towards these peptides could be detected. This is in agreement with the findings that dissociated prosome particles no longer contain proteolytical activity [1,18,19].

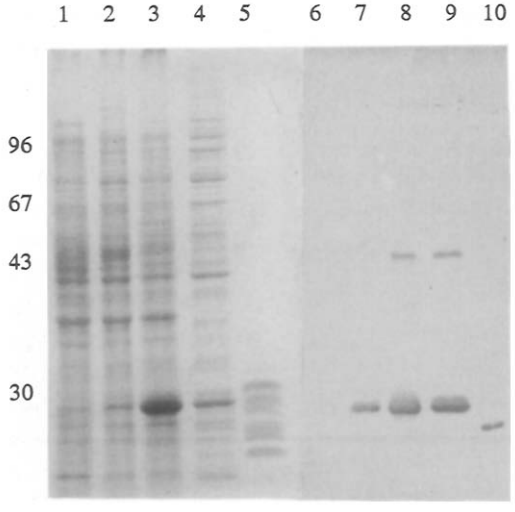


Fig. 4. Expression of the HSB protein in *E. coli*. Lanes (1–5) Coomassie brilliant blue stained gel, lanes (6–10) immunoblot of the same samples with MCP205. Lanes (1) and (6), total *E. coli* BL21(DE3) without expression vector, 4 h after induction with IPTG; lanes (2) and (7), total *E. coli* BL21(DE3) containing the HSB expression vector, just before induction and lanes (3) and (8) 4 h after induction; lanes (4) and (9), WSF of bacteria of lanes 3 and 8; lanes (5) and (10), purified calf liver prosomes. SDS-PAGE molecular mass markers are indicated in kDa.

Acknowledgements: We thank Dr. G. Martens (University of Nijmegen) for providing the XLB cDNA, R. Nelissen for help in the screening procedure and Dr. K. Hendil (University of Copenhagen) for providing us with the MCP205 antibody. The work was supported by the Technology foundation (S.T.W.).

References

[1] Rivett, A.J. (1993) Biochem. J. 291, 1–10.
[2] Scherrer, K. (1990) Mol. Biol. Rep. 14, 1–9.
[3] Dahlmann, B., Kopp, F., Kuehn, L., Niedel, B., Pfeifer, G., Hegerl, R. and Baumeister, W. (1989) FEBS Lett. 251, 125–131.
[4] Zwickl, P., Grziwa, A., Pühler G., Dahlmann, B., Lottspeich, F. and Baumeister, W. (1992) Biochemistry 31, 964–972.
[4a] Bey, F., Silva-Pereira, I., Coux, O., Viegas-Péquinot, E., Recillas Targa, F., Nothwang, H.G., Dutrillaux, B. and Sherrer, K. (1993) Mol. Gen. Genet. 237, 193–205.
[5] Tanaka, K., Tamura, T., Yoshimura, T. and Ichihara, A. (1992) The New Biologist 4, 173–187.
[6] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods Enzymol. 185, 60–89.
[7] Zwickl, P., Lottspeich, F. and Baumeister, W. (1992) FEBS Lett. 312, 157–160.
[8] Van Riel, M.C.H.M. and Martens, G.J.M. (1991) FEBS Lett. 291, 37–40.
[9] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn., Cold Spring Harbor Labatory, Cold Spring Harbor, NY.
[10] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acid Res. 12, 387–395.
[11] Nothwang, H.G., Coux, O., Bey, F. and Scherrer, K. (1992) Eur. J. Biochem. 207, 621–630.
[12] Adams, M.D., Soares, M.B., Kerlavage, A.R., Fields, C. and Venter, J.C. (1993) Nature Genet. 4, 373–380.

[13] Lee, L.W., Moomaw, C.R., Orth, K., McGuire, M.J. DeMartino, G.N. and Slaughter, C.A. (1990) Biochim. Biophys. Acta 1037, 178–185.
[14] Thomson, S., Balson, D.F. and Rivett, A.J. (1993) FEBS Lett. 322, 135–138.
[15] Weterman, M.A.J., Stoopen, G.M., Muijen, G.N.P., Kuznicki, J., Ruiter, D.J. and Bloemers, H.P.J. (1992) Cancer Res. 52, 1291–1296.
[16] Hilt, W., Enenkel, C., Gruhler, A., Singer, T. and Wolf, D.H. (1993) J. Biol. Chem. 268, 3479–3486.
[17] Tamura, T., Shimbara, N., Aki, M., Ishida, N., Bey, F., Scherrer, K., Tanaka, K. and Ichihara, A. (1992) J. Biochem. 112, 530–534.
[18] Yu, B., Pereira, M.E. and Wilk, S. (1993) J. Biol. Chem. 268, 2029.
[19] Nothwang, H.G., Coux, O., Bey, F. and Scherrer, K. (1992) Biochem. J. 287, 733–739.