Identification and characterization of Saccharomyces cerevisiae yapsin 3, a new member of the yapsin family of aspartic proteases encoded by the YPS3 gene

Vicki OLSEN*, Niamh X. CAWLEY*, Jakob BRANDT†, Michi EGEL-MITANI† and Y. Peng LOH*1

*Section of Cellular Neurobiology, Laboratory of Developmental Neurobiology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, U.S.A., and †Molecular Biology, Insulin Research, Novo Nordisk A/S, 2880 Bagsvaerd, Denmark

A new aspartic protease from *Saccharomyces cereisiae*, with a high degree of similarity with yapsin 1 and yapsin 2 and a specificity for basic residue cleavage sites of prohormones, has been cloned. This enzyme was named yapsin 3. Expression of a C-terminally truncated non-membrane anchored yapsin 3 in yeast yielded a heterogeneous protein between 135–200 kDa which, upon treatment with endoglycosidase H, migrated as a 60 kDa form. Amino-acid analysis of the N-terminus of expressed

INTRODUCTION

The yapsin group of enzymes is a recently discovered subfamily of aspartic proteases [1]. Until now this subfamily has included yapsin 1 [also known as yeast aspartic protease 3 (Yap3p)] [2] and yapsin 2 (also known as Mkc7p) [3], both from *Saccharomyces cereisiae*, as well as the mammalian proopiomelanocortin converting enzyme (PCE) (EC 3.4.23.17) found in bovine pituitary intermediate and neural lobe secretory granules [4]. PCE has recently been renamed yapsin A to indicate that it is the first mammalian yapsin characterized with specificity for the basic residues of prohormones [5]. In addition, aspartic proteases with prohormone cleaving properties from bovine chromaffin granules and angler-fish islet secretory granules have been reported [6,7]. So far, the enzymes of the yapsin family have a common specificity for paired or single basic residue cleavage sites of proproteins [2–4,8–11]. This is in contrast with other aspartic proteases, which cleave at hydrophobic residues [12]. Also, the yeast yapsins contain a signal for glycosylphosphatidylinositol (GPI) anchoring [3,13,14], which locates the proteins in the plasma membrane. Although no physiological function has been found for the yeast yapsins to date, it has been suggested that yapsin 1 and yapsin 2 have a role under stress conditions [3,15]. Recent studies have shown that cholecystokinin (CCK) mRNA is co-localized with yapsin-1-like immunoreactivity in rat cortex and hippocampus [16], suggesting that a yapsin-1-related aspartic protease, possibly PCE, may play a role in the processing of pro-CCK and other prohormones in endocrine/neuroendocrine cells. In the present study we have identified and characterized a new yeast member of the yapsin family, yapsin 3, encoded by the *YPS3* gene. Yapsin 3 shows high sequence similarity to yapsin 1 and yapsin 2, and has a specificity for basic residues distinct from that of yapsin 1.

yapsin 3 revealed two different N-terminal residues, serine-48 and phenylalanine-54, which followed a dibasic and a monobasic residue respectively. Cleavage of several prohormones by nonanchored yapsin 3 revealed a specificity distinct from that of yapsin 1.

Keywords: glycosylphosphatidylinositol-anchors, GPI-anchors, proprotein processing, yeast proteinases.

MATERIALS AND METHODS

Materials

Adrenocorticotropic hormone $(ACTH)_{1-39}$, human β -endorphin₁₋₃₁ (β-endorphin₆₁₋₉₁), β-amyloid₁₋₂₈ and CCK₁₃₋₃₃ were purchased from Bachem California (Torrence, CA, U.S.A.). The CCK_{13-33} analogues were custom synthesized by Peptide Technologies Inc. (Gaithersburg, MD, U.S.A.).

Comparison of YPS1 homologues

Homologues of the *YPS1* gene were found using the *Saccharomyces* Genome Database (SGD) of the *S*. *cereisiae* genome (http://genome-www.stanford.edu/Saccharomyces) [17]. The homologues obtained were compared using the 'Genome-wide Protein Similarity' function found in the same database, based on a Smith–Waterman protein sequence comparison [18,19].

Cloning of the YPS3 gene

The *YPS3* gene was cloned from a gene library, as described by Egel-Mitani et al. [2]. From a positive yeast transformant, an 8 kb fragment (pME719), containing *YPS1* [open reading frame (ORF)-YLR120C, previously called *YAP3*] and *YPS3* (ORF-YLR121C), was re-isolated. *YPS3* was subcloned into the expression vector pEMBLyex4 [20] in a truncated form, the 35 amino acid residues nearest the C-terminal having been deleted, creating the construct pYps3–∆35. Transformation into *S*. *cereisiae* strain BJ3501 (MATα, pep4::HIS3 prbl–∆1.6R, his3 ∆200, ura3–52, can1, gal 2) [21] was performed as described by Gietz et al. [22].

Abbreviations used: CCK, cholecystokinin; ACTH, adrenocorticotropic hormone; GPI, glycosylphosphatidyl inositol; PCE, pro-opiomelanocortin

¹ To whom correspondence should be addressed (e-mail ypl@codon.nih.gov).

*Expression of pYps3-***∆***35*

Expression of pYps3–∆35, using the galactose-inducible promoter on pEMBLyex4, was performed as described previously [21]. After incubation for 20 h in galactose medium (0.5 l), the whole of the culture supernatant was concentrated to approx. 100 μ g of protein/ml (\approx 15 ml) by centrifugation filtration using a Filtron 50 kDaMacrosep Omega membrane filter. As a negative control, cells transformed with the empty vector were grown in parallel and treated in the same manner.

Analysis of expressed yapsin 3

Concentrated culture media $(15 \mu g)$ of protein) was run on a Tris/glycine precast SDS/8–16% (w/v) polyacrylamide gel (Novex) and analysed by Coomassie Blue staining. Aliquots containing 15 μ g of protein (approx. 150 μ l) were treated with 0.001 unit of endoglycosidase H (Endo H, Sigma) in 0.05 M sodium phosphate buffer, pH 6.5, containing 4.2 mM 4-(2 aminoethyl)benzenesulphonyl fluoride protein inhibitor at 37 °C for 1 h, and analysed in the same manner. A similar gel was run in parallel except that the proteins, after transfer to a poly(vinylidene difluoride) membrane (Novex), were analysed by N-terminal amino-acid sequencing in 25 mM Tris/200 mM glycine/0.1% SDS buffer containing 20% (v/v) methanol. The proteins were revealed by staining with 0.2% (w/v) Ponceau S in 1% (v/v) acetic acid and the gels were destained in water. Direct N-terminal amino acid sequencing of the bands was carried out by Edman degradation using a Procise 228 Protein Sequencer (model 494A; Perkin-Elmer–Applied Biosystems). β-Lactoglobulin was used to determine the sequence efficiency.

Enzyme assays

Non-anchored yapsin 3 (pYps3–∆35) (1 μg total protein from concentrated medium) was incubated with 100 μ M of substrate in a total volume of 100 μ l (0.1 M sodium acetate, pH 5.3) at 37 °C for 1 h, unless otherwise indicated. The reaction was stopped by the addition of 10 μ l glacial acetic acid. Products and substrate were separated by HPLC (LKB 2150) using a Bio-Rad HiPore RP-318 column (5 mm \times 250 mm). Solution A was 0.1% (w/v) trifluoroacetic acid and solution B was 80% (v/v) acetonitrile/ 0.1% (w/v) trifluoroacetic acid. For β -amyloid₁₋₂₈, a linear gradient of 10–40% solution B in 30 min at 1 ml/min was used. β -Endorphin₁₋₃₁ and its products were separated using a two-step gradient: $10-35\%$ of solution B in 30 min at a rate of 1 ml/min followed by $35-45\%$ of solution B in 30 min at 1 ml/min. ACTH_{1-39} , CCK_{13-33} and the CCK analogues were separated as described previously [23]. Quantification of the products generated from each substrate was performed by measuring the peak height (in mm) at A_{214} and converting into nmol of product using standard curves generated using identical gradient conditions. Identification of the cleavage products was performed by direct N-terminal amino acid sequencing of the collected products as described above.

RESULTS AND DISCUSSION

YPS1, previously named *YAP3* [2], encodes a basic-residuespecific aspartic protease, yapsin 1. A search for genes homologous to *YPS1* in the yeast genome (using the *Saccharomyces* Genome Database) [17] revealed six different genes encoding aspartic proteases. Three have been described previously: *BAR1* [24], *PEP4* [25], and *YPS2* (*MKC7*) [3]. The three unknown homologues, ORF-YLR121C, ORF-YIR039C and ORF-

YDR349C, were designated *YPS3*, *YPS6* and *YPS7* respectively, as they represented potential homologues of *YPS1* and *YPS2*. A pseudogene (ORF-YGL259W) encoding 165 amino acids was also found and designated *YPS5*. Comparison of protein similarity (Table 1) shows that *YPS3* appears to fall within a group composed of the yeast yapsins characterized previously, i.e. *YPS1* and *YPS2*. In this group, sequence identity between the members is \approx 50%. *YPS6* and *BAR1* can be grouped together, based on their identity with the yapsins of $\approx 35\%$, and *YPS7* and *PEP4* make up a third group with $\approx 25\%$ amino-acid identity with the yapsins. These results alone would indicate that yapsin 3, encoded by the *YPS3* gene, is a probable member of the yapsin family. Interestingly, *YPS3* is located next to *YPS1* on chromosome XII. The proteins encoded by *YPS6* and *YPS7* have sequence similarities to yapsin 1, 2 and 3, which are not greater than those encoded by *BAR1* and *PEP4* respectively (Table 1), but do share with yapsins 1, 2 and 3 the property of having a putative GPI-anchoring signal. It would be necessary therefore, to express and characterize the activities of these new aspartic proteases to accurately classify them as yapsins, since the primary criteria for being a member of the yapsin family of aspartic proteases is the ability to specifically cleave substrates at basic amino acids.

An alignment of the protein sequence of yapsin 3 with yapsin 1 and yapsin 2 shows a high degree of similarity throughout the entire protein sequence (Figure 1). Yapsins contain a signal peptide within the N-terminus that directs the newly synthesized proteins to the secretory pathway; this is followed by a propeptide. Unlike yapsin 1 and yapsin 2, yapsin 3 does not contain a large loop insertion almost immediately after the first active-site aspartic acid residue. The function of this loop has not yet been determined, but for yapsin 1 it has been predicted by molecular modelling to be located on the surface of the protein [23]. Removal of the loop from yapsin 1 resulted in no change in the specificity of yapsin 1 (N. X. Cawley and V. Olsen, unpublished work). Cleavage of yapsin 1 into an α - and β -subunit occurs in this loop region [26] and, since yapsin 3 does not contain this loop it probably remains as one polypeptide chain.

The protein sequence of yapsin 3 contains 11 potential Nlinked glycosylation sites as well as a serine/threonine rich domain, which is likely to be O-linked glycosylated as has been suggested for yapsin 1 and Bar1p [13,15,27]. The C-termini of yapsin 1 and yapsin 2 have been shown to function as a GPIanchoring signal, allowing the proteins to be bound to the plasma membrane [3,13,14]. Removal of this tail from yapsin 1 results in secretion of the truncated protein to the medium [13]. The high degree of similarity of yapsin 3 to yapsin 1 and yapsin 2 in this region of the protein, makes it very likely that yapsin 3 is located in the plasma membrane *in io* through the same mechanism, which also has been suggested by Caro et al. [28].

To verify that yapsin 3 is a member of the yapsin family, characterization of the specificity of yapsin 3 was performed. In order to avoid the predicted GPI anchoring of the enzyme to the plasma membrane, and to facilitate the acquisition of an enriched preparation of the expressed protein in the medium, a nonanchored yapsin 3 with the putative GPI-anchoring signal deleted was expressed. Analysis of the gel following SDS/PAGE confirmed that yapsin 3 was overexpressed, secreted into the culture medium and appeared as a diffuse band with an apparent molecular mass of 135–200 kDa (Figure 2, lane 1). Upon treatment with Endo H, yapsin 3 was converted into a major band with a molecular mass of approx. 60 kDa (Figure 2, lane 3), which was not observed in medium from control cells transformed with vector without an insert (negative control) (Figure 2, lane 4). These results were very similar to those obtained for yapsin 1

Table 1 Protein sequence similarity of yeast aspartic proteases

The data shown are based on a Smith–Waterman protein-sequence comparison [18,19] according to the *Saccharomyces* Genome Database (SGD). ¹Percentage alignment indicates the percentage of the gene/ORF shown at the top of the Table which aligns with the gene/ORF shown on the left of the Table. ²Percentage identity indicates the percentage identity within the aligned portion of the sequences. N.D., not determined because the percentage alignment/identity was below the threshold used in the *Saccharomyces* Genome Database, i.e. $P < 0.1$.

Figure 1 Alignment of the deduced amino acid sequence of yapsin 3 with yapsin 1 (Yap3p) and yapsin 2 (Mkc7p)

Asterisks denote identical amino acid residues. Arrows denote the determined N-terminal residues for yapsin 1 and 3 (bold and underlined). The two catalytic asparatyl residues are shown in bold type, the hydrophobic C-termini are shown in italics and the potential Nglycosylated asparagine residues are underlined.

[13], and demonstrated that the 135–200 kDa band of yapsin 3 represented the differentially N-linked glycosylated form of the enzyme.

N-terminal amino-acid sequencing of glycosylated and deglycosylated yapsin 3 resulted in two yapsin 3 sequences: (1)

Figure 2 Analysis of culture medium following SDS/PAGE

Concentrated culture medium (15 μ g) was loaded under reducing conditions into each lane and after PAGE the gel was stained with Coomassie Blue. Lanes 1 and 3 contained medium from cells expressing yapsin 3. Lanes 2 and 4 contained medium from cells transformed with the empty expression vector. The medium used in lanes 3 and 4 was treated with Endo H before gel electrophoresis. The mobilities of molecular mass markers (SeeBlue ; Novex) are shown on the right of each panel.

Figure 3 Influence of pH on the activity of yapsin 3

Protein (4.2 μ g) from concentrated culture medium containing yapsin 3 was incubated with 10 μ g of CCK_{13–33} for 1 h at 37 °C, and at various pH values (attained by adjusting 0.1 M sodium acetate and 0.1 M sodium phosphate buffers).

Table 2 Cleavage specificity of yapsin 3

 1 Sequence of β -endorphin_{1–31}. TGGFMTSEK₉SQTPLVTLFK₁₉NAIIK₂₄NAHK₂₈K₂₉GQ. ²The products generated in the negative control were from cleavage after the P1 arginine and after the P3 methionine residues. N.D., no cleavage activity was observed within the assay time.

S⁴⁸NGHEKFVLANEQSF and (2) F⁵⁴VLANEQSFYSVELA. S⁴⁸ represents the mature N-terminus predicted to result from the activation of proyapsin 3 upon removal of its putative propeptide at $K^{46}R^{47} \downarrow S^{48}$, which is identical to that of proyapsin 1 [26], whereas $K^{53} \downarrow F^{54}$ represents an additional and novel processing site.

Initial characterization of the overexpressed enzyme demonstrated its ability to cleave CCK_{13-33} specifically at lysine-23, with an optimum pH of \approx 5.3 (Figure 3). This activity was completely inhibited by pepstatin A, an active-site-specific aspartic protease inhibitor (results not shown). These results were identical to the specificity [8] and pH optimum (results not shown) of yapsin 1 for this substrate. The low level of protease activity detected in the negative control was most likely due to endogenous yapsins expressed from their natural promoters in the genome and not due to other class-specific proteases. This conclusion was based on the basic residue cleavage specificity of the activity and on the observation that pepstatin A completely inhibited this background activity.

The cleavage specificity of yapsin 3 is shown in Table 2. Without exception, where cleavage had occurred, cleavage only after specific basic residues was observed. In contrast to yapsin 1, which cleaves ACTH_{1-39} almost 200-fold more efficiently than CCK_{13-33} [8], no cleavage of $ACTH_{1-39}$ was observed with yapsin 3, but a 10-fold increase in CCK_{13-33} cleaving activity above that of the negative control was apparent. This represents a distinct difference between the enzymic properties of these two enzymes. A second distinction was found in the preference of yapsin 3 for lysine-9 of β -endorphin₁₋₃₁, whereas yapsin 1 preferred lysine-19 (results not shown). The difference between lysine-9 and lysine-19 is the presence of a lysine residue in the P5' position relative to lysine-19 (see the caption to Table 2 for the β -endorphin₁₋₃₁ sequence). A third and perhaps more dramatic distinction was found with the CCK analogues, described previously by Olsen et al. [23]. Whereas yapsin 1 activity has been shown to be enhanced 21-fold by placing an arginine residue in the $P2'$ position [23] relative to wild-type CCK_{13-33} , yapsin 3 activity was significantly reduced for this substrate $[CCK(P2' Arg)]$.

The common motif for the substrates tested, where a difference between the specificity of yapsin 1 and yapsin 3 was observed, was the presence of additional basic residues flanking the cleavage site. Although these residues have been found to enhance the cleavage efficiency of yapsin 1 [8,23], they appeared to decrease the efficiency of or to prevent yapsin 3 cleavage.

Incubation of $β$ -amyloid₁₋₂₈ with yapsin 3 resulted in cleavage after the single lysine-16 residue which was about 6-fold greater than that of the negative control. Interestingly, it has been reported previously that yapsin 1 and yapsin 2 cleaves the β amyloid precursor *in io* [29,30]. However, disruption of both *YPS1* and *YPS2* reduced the *β*-amyloid cleaving activity by only approx. 85% , suggesting that yeast contains an additional protease capable of cleaving this precursor [29]. The results of the present study suggest that yapsin 3 can perform this task. The role of the yapsins as α-secretase-like enzymes, involved in the processing of cell-associated precursors to secreted forms, is therefore likely, not only in yeast but also in mammals. So far the enzyme(s) responsible for liberating the ectodomain of the β-amyloid peptide precursor *in io* have not been identified, however, one can speculate on the role of mammalian yapsins in this process.

In conclusion, the results of the present work allowed the identification of three new aspartic proteases in *S*. *cereisiae* encoded by the genes *YPS3*, *YPS6* and *YPS7*. However, only *YPS3* showed a higher degree of sequence identity with *YPS1* and *YPS2* than with *BAR1* and *PEP4*, suggesting that this protein is a member of the yapsin family. Compared with the other members of the yapsin family, yapsin 3 showed a high degree of similarity throughout the sequence. The results of this study furthermore demonstrate that the putative propeptide of yapsin 3 is removed at a similar position to that of the propeptide of yapsin 1. Although the sequences of yapsin 1 and yapsin 3 appear to be homologues and both are capable of prohormone cleavage, their specificities are quite distinct, as shown by differences in their affinity for substrates with basic residues governing the cleavage site. Presently, modelling studies of the yapsin family are being carried out in our laboratory in order to understand the overlapping, yet different, specificities of members of the yapsin family.

We thank Dr. Hao-Chia Chen, Endocrinology and Reproduction Research Branch, National Institute of Child Health and Human Development, National Institutes of Health (Bethesda, MD, U.S.A.) for performing N-terminal amino-acid sequencing of yapsin 3 and peptide products. This research was supported by a Danish Natural Science Council Grant (9400095) to V.O.

REFERENCES

Cawley, N. X. and Loh, Y. P. (1998) in Handbook of Proteolytic Enzymes (Barrett, A. J., Rawlings, N. D. and Woessner, J. F., eds.), pp. 905–907, Academic Press, London

- 2 Egel-Mitani, M., Flygenring, H. P. and Trier Hansen, M. (1990) Yeast *6*, 127–137
- 3 Komano, H. and Fuller, R. S. (1995) Proc. Natl. Acad. Sci. U.S.A. *92*, 10752– 10756
- 4 Loh, Y. P., Parish, D. C. and Tuteja, R. (1985) J. Biol. Chem. *260*, 7194–7205
- 5 Cawley, N. X., Cool, D. R., Normant, E., Shen, F.-S., Olsen, V. and Loh, Y. P. (1998) in Proteolytic and Cellular Mechanisms in Prohormone and Proprotein Processing (Hook, V., ed.), pp. 29–48, R. G. Landes Inc., Austin
- 6 Azaryan, A. V., Schiller, M., Mende-Mueller, L. and Hook, V. Y. H. (1995) J. Neurochem. *65*, 1771–1779
- 7 Mackin, R. B., Noe, B. D. and Spiess, J. (1991) Endocrinology *129*, 1951–1957
- 8 Cawley, N. X., Chen, H.-C., Beinfeld, M. C. and Loh, Y. P. (1996) J. Biol. Chem. *271*, 4168–4176
- 9 Cawley, N. X., Noe, B. D. and Loh, Y. P. (1993) FEBS Lett. *332*, 273–276
- 10 Bourbonnais, Y., Germain, D., Ash, J. and Thomas, D. Y. (1994) Biochimie *76*, 226–233
- 11 Loh, Y. P. and Cawley, N. X. (1995) Methods Enzymol. *248*, 136–146
- 12 Dunn, B. M., Jimenez, M., Parten, B. F., Valler, M. J., Rolph, C. E. and Kay, J. (1986) Biochem. J. *237*, 899–906
- 13 Cawley, N. X., Wong, M., Pu, L.-P., Tam, W. and Loh, Y. P. (1995) Biochemistry *34*, 7430–7437
- 14 Ash, J., Dominguez, M., Bergeron, J. J. M., Thomas, D. Y. and Bourbonnais, Y. (1995) J. Biol. Chem. *270*, 20847–20854
- 15 Olsen, V. (1994) Laboratory for Cellular and Molecular Physiology, pp. 1–115, University of Copenhagen, Denmark
- 16 Cawley, N. X., Pu, L.-P. and Loh, Y. P. (1996) Endocrinology *137*, 5135–5143
- 17 Cherry, J. M., Ball, C., Weng, S., Juvik, G., Schmidt, R., Adler, C., Dunn, B., Dwight,

Received 2 November 1998/4 January 1999 ; accepted 8 February 1999

S., Riles, L., Mortimer, R. K. and Botstein, D. (1997) Nature (London) *387* (Suppl. 6632), 67–73

- 18 Smith, T. F. and Waterman, M. S. (1981) J. Mol. Biol. *147*, 195–197
- 19 Pearson, W. R. (1991) Genomics *11*, 635–650
- 20 Baldari, C., Murray, J. A. H., Ghiara, P., Cesareni, G. and Galeotti, C. L. (1987) EMBO J. *6*, 229–234
- 21 Azaryan, A., Wong, M., Friedman, T. C., Cawley, N. X., Estivariz, F. E., Chen, H.-C. and Loh, Y. P. (1993) J. Biol. Chem. *268*, 11968–11975
- 22 Gietz, D., St. Jean, A., Woods, R. A. and Schiestl, R. H. (1992) Nucleic Acids Res. *20*, 1425
- 23 Olsen, V., Guruprasad, K., Cawley, N. X., Chen, H.-C., Blundell, T. L. and Loh, Y. P. (1998) Biochemistry *37*, 2768–2777
- 24 MacKay, V. L., Welch, S. K., Insley, M. Y., Manney, T. R., Holly, J., Saari, G. C. and Parker, M. L. (1988) Proc. Natl. Acad. Sci. U.S.A. *85*, 55–59
- 25 Woolford, C. A., Daniels, L. B., Park, F. J., Jones, E. W., van Arsdell, J. N. and Innis, M. A. (1986) Mol. Cell. Biol. *6*, 2500–2510
- 26 Cawley, N. X., Olsen, V., Zhang, C.-F., Chen, H.-C., Tan, M. and Loh, Y. P. (1998) J. Biol. Chem. *273*, 584–591
- 27 MacKay, V. L., Armstrong, J., Yip, C., Welch, S., Walker, K., Osborn, S., Sheppard, P. and Forstrom, J. (1991) in Structure and Function of the Aspartic Proteinases (Dunn, B. M., ed.), pp. 161–172, Plenum Press, New York
- 28 Caro, L. H. P., Tettelin, H., Vossen, J. H., Ram, A. F. J., van den Ende, H. and Klis, F. M. (1997) Yeast *13*, 1477–1489
- 29 Zhang, W., Espinoza, D., Hines, V., Innis, M., Mehta, P. and Miller, D. L. (1997) Biochim. Biophys. Acta *1359*, 110–122
- 30 Hines, V., Zhang, W., Ramakrishna, N., Styles, J., Mehta, P., Kim, K. S., Innis, M. and Miller, D. L. (1994) Cell. Mol. Biol. Res. *40*, 273–284