Vol. 156, No. 1, 1988 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 14, 1988 Pages 246-254

YEAST KEX2 GENE ENCODES AN ENDOPEPTIDASE HOMOLOGOUS TO SUBTILISIN-LIKE SERINE PROTEASES

Kensaku MIZUNO, Tomoko NAKAMURA, Takehiro OHSHIMA*, Shoji TANAKA* & Hisayuki MATSUO

Department of Biochemistry, Miyazaki Medical College, Kiyotake, Miyazaki 889-16, Japan

*Suntory Institute for Biomedical Research, Shimamoto, Mishima, Osaka 618, Japan

Received August 8, 1988

SUMMARY: Yeast Saccharomyces cerevisiae KEX2 gene previously isolated, was characterized as the gene encoding a calcium-dependent endopeptidase required for processing of precursors of α-factor and killer toxin. In this study, we report the amino acid sequence of the KEX2 gene product deduced from nucleotide sequencing. Our results indicate that the KEX2 gene contains a 2,442-bp open reading frame encoding a polypeptide of 814 amino acids. The deduced amino acid sequence contains a region extensively homologous to the members of subtilisin-like serine protease family near the N-terminus. A putative membrane-spanning domain near the C-terminus was also detected. These facts indicate that the KEX2-encoded protein may function as a membrane-bound, subtilisin-like serine protease.

Many bioactive peptides are processed from their precursors by proteolytic cleavage at paired basic residues. However, little is known about the endopeptidases that are physiologically involved in precursor processing (1). Yeast Saccharomyces cerevisiae cells synthesize and secrete α-factor and killer toxin, which are also processed from larger precursors by specific cleavage at paired basic residues (2,3). Therefore, yeast may provide a simple model system for the study of prohormone processing in higher eukaryotic cells. Yeast kex2 mutants are defective in proteolytic processing of precursors for α-factor and killer toxin (4). Reintroduction of the normal KEX2 gene into deficient strains restored both processing activity and endopeptidase activity specific for cleaving on the carboxyl side of paired basic residues, indicating that the KEX2 gene product may be an endopeptidase involved in precursor processing (4,5). The KEX2 gene product was found to exhibit unique enzymatic properties, including membrane-association, Ca^{2+}-dependency, and substrate specificity toward paired basic residues (4,5).

Abbreviations: MCA, 4-methylcoumaryl-7-amide; AMC, 7-amino-4-methylcoumarin; PMSF, phenylmethylsulfonyl fluoride; DFP, diisopropyl fluorophosphate; pAPMSF, p-amidinophenylmethylsulfonyl fluoride.

0006-291X/88 $1.50

Enzymes with properties similar to those of the KEX2 protease were partially purified by Wolf's group (6,7) and ours (8). To understand the molecular mechanisms of this unique proteolytic action of the enzyme, it was essential to elucidate its primary structure. In a recent review, Fuller, et al. showed the schematic depiction of the structure of the KEX2-encoded protease (9); however, no data regarding its primary structure had as yet been available. In this paper, we report the complete amino acid sequence of the KEX2-encoded protein deduced from nucleotide sequencing.

MATERIALS AND METHODS

Strains and DNA: Saccharomyces cerevisiae K16-57C (MATα leu2 trp1 ura3 kex2-8) was used for the kex2 mutant host. The kex2-8 mutation of K16-57C derived from strain 399 was obtained from Dr. Reed Wickner. R27-7C (MATα leu2 trp1 ura3 his3) was used for the KEX2 wild type strain. The genomic DNA used for isolation of the KEX2 gene was prepared from X2180-1B (10).
Recombinant DNA techniques: Ligation, T4 DNA polymerase and restriction endonuclease treatment were performed according to Maniatis, et al. (11).
Plasmid: pYE vector, carrying the selection marker gene TRP1 and one of the inverted repeat sequence of 2μm DNA which contains the replication origin, was used for the high copy number plasmid. YCpLe vector, carrying the selection marker gene LEU2, CEN4 and ARS1, was used for the single copy number plasmid. Both vectors have pBR322 replication origin and ampicillin-resistant gene.
Transformation: Transformation of yeast was carried out by using the lithium acetate technique (12). Transformation of E.coli was performed as described before (13).
Bioassay for killer activity: To isolate the KEX2 gene, the transformants were screened for the secretion of active killer toxin. Agar diffusion assay for killer activity was performed on a lawn of the sensitive indicator 5 x 47 (MATa/MATα his1/+ trp1/+ ura3/+) (14), according to the procedure of Wickner and Leibowitz (15).
Assay for protease activity: Permeabilization of yeast cells with Brij 58 was carried out by the methods described previously (4). Protease activity was measured as follows (8): Twenty nanomoles of Boc-Gln-Arg-Arg-MCA were incubated for 1 hr at 37°C with the permeabilized cells in 250 μl of 0.4 M Tris-HCl buffer (pH 7.0) containing 0.1% Lubrol, 1 mM EGTA, pepstatin (1mg/ml), and bestatin (1mg/ml) in the presence or absence of 2 mM $CaCl_2$. The amounts of AMC released from the substrate were measured by a fluorescence spectrophotometer with excitation at 380 nm and emission at 460 nm. Calcium-dependent endopeptidase activity was evaluated after the addition of $CaCl_2$.
Sequencing: DNA sequence analysis was carried out by the dideoxy method after subcloning suitable restriction fragments into M13mp18 and M13mp19 vectors (16). The DNA sequence was determined for both strands, and across all restriction sites used for subcloning.

RESULTS AND DISCUSSION

The KEX2 gene was previously isolated by Julius, et al. (4). They described that the functional KEX2 gene was located on a 3.5Kb EcoRI fragment. According to their description, we isolated the 3.5Kb EcoRI fragment capable of complementing the kex2-8 mutation from X2180-1B genomic DNA (Fig.1A). However, the nucleotide sequence analysis revealed that the 3.5 Kb EcoRI fragment contained one long open reading frame without a stop codon. To determine the complete coding sequence of the KEX2 gene product, the 5.0 Kb EcoRI-SalI fragment capable of complementing the kex2-8 mutation was newly

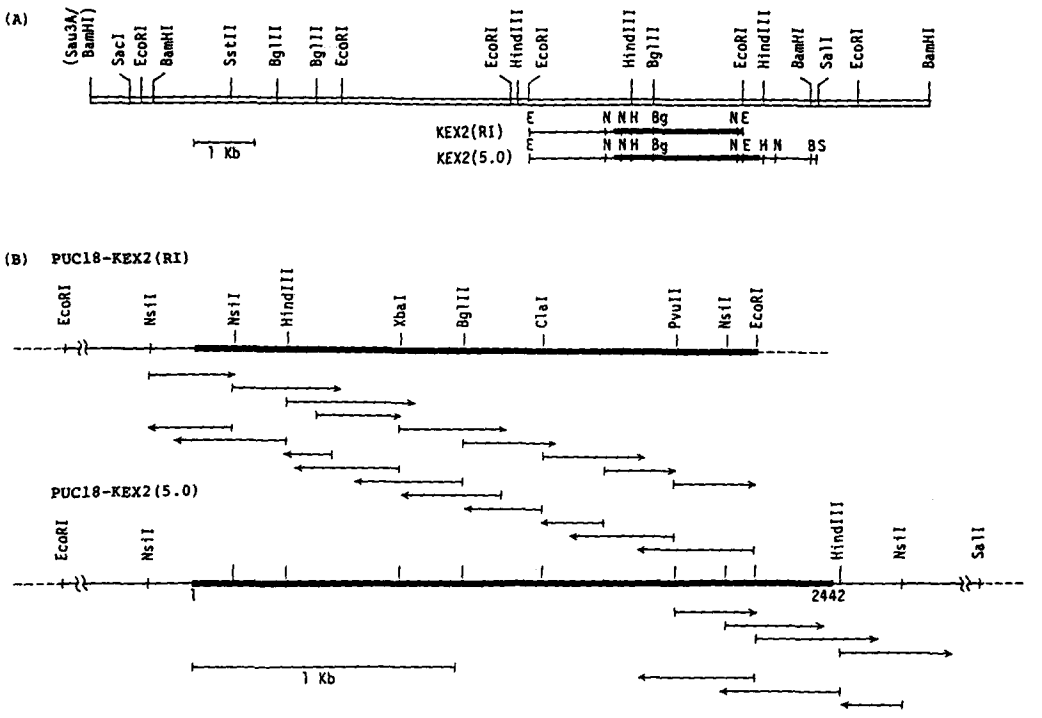

Fig.1 (A) Upper: Physical map of KEX2 gene previously reported (4). Middle and lower: Physical maps of KEX2(RI) and KEX2(5.0) isolated in this study. E;EcoRI, N;NsiI, H;HindIII, B;BamHI, S;SalI, Bg;BglII. (B) Strategy for sequencing cloned DNA. The restriction map shows only the relevant restriction sites. Protein coding region is indicated by a closed box. Arrows under each clone indicate direction and extent of sequence determination. Dashed line indicates the region of PUC18.

isolated. As shown in Fig.1A, the restriction maps of these two isolated genes (named KEX2(RI) and KEX2(5.0)) were identical to that of the KEX2 gene previously reported (4). Transformants of the kex2-8 mutants containing the cloned genes on a single- or multi-copy number plasmid regained both killer activity and Ca^{2+}-dependent endopeptidase activity (Table 1). Hence, these genes probably encode a Ca^{2+}-dependent endopeptidase required for maturation

Table 1. Killer and endoprotease activity of the normal and kex2 mutant cells

Strain	Inserting direction of KEX2 gene	Killer activity	Endoprotease activity
K16-57C (non)	———	-	0.5
K16-57C[pYE-KEX2(RI)a]	KEX2(RI) → IRI	+	22.9
K16-57C[pYE-KEX2(RI)b]	← KEX2(RI) IRI	++	50.3
K16-57C[pYE-KEX2(5.0)b]	← KEX2(5.0) IRI	++	78.8
K16-57C[YCpLe-KEX2(RI)a]	KEX2(RI) → CEN4	+	5.0
K16-57C[YCpLe-KEX2(RI)b]	← KEX2(RI) CEN4	+	11.1
R27-7C (non)	———	+	7.2

of active killer toxin. Two plasmids containing overlapping segments of the yeast genome were analyzed for nucleotide sequence by the dideoxy method, according to the strategy indicated in Fig.1B.

Fig.2 shows the 2,848-nucleotide sequence of the KEX2 gene and the deduced amino acid sequence of the KEX2 gene product. The sequence contains one long open reading frame corresponding to 814 amino acid residues. The translational initiation site is tentatively assigned to the methionine codon at nucleotides 1 to 3, because this is the first ATG triplet that appears downstream from the nonsense codon TAA (nucleotides -57 to -55) found in-frame. A translation stop codon (TGA) occurs in-frame after the 814th codon specifying serine. A putative polyadenylation site AATAAA is present at position 2,616 to 2,621. Thus the KEX2 gene appears to encode a polypeptide of 814 amino acid residues with a calculated molecular weight of 89,997.

The hydropathicity profile of the deduced amino acid sequence revealed two hydrophobic segments. One is at the N-terminus of the protein and may function as a signal sequence to direct the KEX2 protein into endoplasmic reticulum. The other segment, consisting of 21 hydrophobic residues (residue 679 to 699) near the C-terminus of the protein, may function as a membrane-spanning domain, as was generally found in many transmembrane proteins. There are five potential sites for N-linked glycosylation, that is, Asn-X-Ser/Thr, where X can be any amino acid except proline (Fig. 2).

When the amino acid sequence deduced for the KEX2 gene product was compared with the known polypeptide sequences stored in the NBRF data bank, significant homology was found with members of subtilisin-like serine protease family (17). Fig.3A shows an alignment of the respective amino acid sequences of the KEX2 gene product and subtilisin BPN' (18). When the region of amino acid residues [152-410] of KEX2 gene product was compared with the sequence [9-246] of subtilisin BPN', 28% was found to be occupied by identical residues and 22% by conservative residues. The active site of subtilisin BPN' was assigned to the triad consisting of residues Asp-32, His-64, and Ser-221 (18). The KEX2 protein also contains these three residues at homologous positions (Asp-175, His-213 and Ser-385). The extensive homology of the amino acid sequence observed between KEX2 protein and subtilisins, especially in the region around the active site residues (Fig.3B), strongly suggests that the KEX2 protein may physiologically function as a serine protease in a manner similar to that of the subtilisins. Subtilisins are classified into a serine-

---

Fig.2 Nucleotide sequence of KEX2 gene and predicted amino acid sequence of KEX2 gene product. Nucleotides are numbered from the presumed initiator ATG. Amino acid residues are numbered starting at Met-1. The potential N-glycosylation sites are underlined. The residues corresponding to the active site of subtilisin are boxed. Two hydrophobic domains situated at the N-terminus and near the C-terminus are indicated by dashed lines. The AATAAA sequence is doubly underlined.

```
                                                T GCA TAA TTC TGT CAT AAG CCT GTT   -145
-144  CTT TTT CCT GGC TTA AAC ATC CCG TTT TGT AAA AGA GAA ATC TAT TCC               -97
 -96  ACA TAT TTC ATT CAT TCG GCT ACC ATA CTA AGG ATA AAC TAA TCC CGT               -49
 -48  TGT TTT TTG GCC TCG TCA CAT AAT TAT AAA CTA CTA ACC CAT TAT CAG                -1

   1  Met Lys Val Arg Lys Tyr Ile Thr Leu Cys Phe Trp Trp Ala Phe Ser                16
   1  ATG AAA GTG AGG AAA TAT ATT ACT TTA TGC TTT TGG TGG GCC TTT TCA                48

  17  Thr Ser Ala Leu Val Ser Ser Gln Gln Ile Pro Leu Lys Asp His Thr                32
  49  ACA TCC GCT CTT GTA TCA TCA CAA CAA ATT CCA TTG AAG GAC CAT ACG                96

  33  Ser Arg Gln Tyr Phe Ala Val Glu Ser Asn Glu Thr Leu Ser Arg Leu                48
  97  TCA CGA CAG TAT TTT GCT GTA GAA AGC AAT GAA ACA TTA TCC CGC TTG               144

  49  Glu Glu Met His Pro Asn Trp Lys Tyr Glu His Asp Val Arg Gly Leu                64
 145  GAG GAA ATG CAT CCA AAT TGG AAA TAT GAA CAT GAT GTT CGA GGG CTA               192

  65  Pro Asn His Tyr Val Phe Ser Lys Glu Leu Leu Lys Leu Gly Lys Arg                80
 193  CCA AAC CAT TAT GTT TTT TCA AAA GAG TTG CTA AAA TTG GGC AAA AGA               240

  81  Ser Ser Leu Glu Glu Leu Gln Gly Asp Asn Asn Asp His Ile Leu Ser                96
 241  TCA TCA TTA GAA GAG TTA CAG GGG GAT AAC AAC GAC CAC ATA TTA TCT               288

  97  Val His Asp Leu Phe Pro Arg Asn Asp Leu Phe Lys Arg Leu Pro Val               112
 289  GTC CAT GAT TTA TTC CCG CGT AAC GAC CTA TTT AAG AGA CTA CCG GTG               336

 113  Pro Ala Pro Pro Met Asp Ser Ser Leu Leu Pro Val Lys Glu Ala Glu               128
 337  CCT GCT CCA CCA ATG GAC TCA AGC TTG TTA CCG GTA AAA GAA GCT GAG               384

 129  Asp Lys Leu Ser Ile Asn Asp Pro Leu Phe Glu Arg Gln Trp His Leu               144
 385  GAT AAA CTC AGC ATA AAT GAT CCG CTT TTT GAG AGG CAG TGG CAC TTG               432

 145  Val Asn Pro Ser Phe Pro Gly Ser Asp Ile Asn Val Leu Asp Leu Trp               160
 433  GTC AAT CCA AGT TTT CCT GGC AGT GAT ATA AAT GTT CTT GAT CTG TGG               480

 161  Tyr Asn Asn Ile Thr Gly Ala Gly Val Val Ala Ala Ile Val Asp Asp               176
 481  TAC AAT AAT ATT ACA GGC GCA GGG GTC GTG GCT GCC ATT GTT GAT GAT               528

 177  Gly Leu Asp Tyr Glu Asn Glu Asp Leu Lys Asp Asn Phe Cys Ala Glu               192
 529  GGC CTT GAC TAC GAA AAT GAA GAC TTG AAG GAT AAT TTT TGC GCT GAA               576

 193  Gly Ser Trp Asp Phe Asn Asp Asn Thr Asn Leu Pro Lys Pro Arg Leu               208
 577  GGT TCT TGG GAT TTC AAC GAC AAT ACC AAT TTA CCT AAA CCA AGA TTA               624

 209  Ser Asp Asp Tyr His Gly Thr Arg Cys Ala Gly Glu Ile Ala Ala Lys               224
 625  TCT GAT GAC TAC CAT GGT ACG AGA TGT GCA GGT GAA ATA GCT GCC AAA               672

 225  Lys Gly Asn Asn Phe Cys Gly Val Gly Val Gly Tyr Asn Ala Lys Ile               240
 673  AAA GGT AAC AAT TTT TGC GGT GTC GGG GTA GGT TAC AAC GCT AAA ATC               720

 241  Ser Gly Ile Arg Ile Leu Ser Gly Asp Ile Thr Thr Glu Asp Glu Ala               256
 721  TCA GGC ATA AGA ATC TTA TCC GGT GAT ATC ACT ACG GAA GAT GAA GCT               768

 257  Ala Ser Leu Ile Tyr Gly Leu Asp Val Asn Asp Ile Tyr Ser Cys Ser               272
 769  GCG TCC TTG ATT TAT GGT CTA GAC GTA AAC GAT ATA TAT TCA TGC TCA               816

 273  Trp Gly Pro Ala Asp Asp Gly Arg His Leu Gln Gly Pro Ser Asp Leu               288
 817  TGG GGT CCC GCT GAT GAC GGA AGA CAT TTA CAA GGC CCT AGT GAC CTG               864

 289  Val Lys Lys Ala Leu Val Lys Gly Val Thr Glu Gly Arg Asp Ser Lys               304
 865  GTG AAA AAG GCT TTA GTA AAA GGT GTT ACT GAG GGA AGA GAT TCC AAA               912

 305  Gly Ala Ile Tyr Val Phe Ala Ser Gly Asn Gly Gly Thr Arg Gly Asp               320
 913  GGA GCG ATT TAC GTT TTT GCC AGT GGA AAT GGT GGA ACT CGT GGT GAT               960

 321  Asn Cys Asn Tyr Asp Gly Tyr Thr Asn Ser Ile Tyr Ser Ile Thr Ile               336
 961  AAT TGC AAT TAC GAC GGC TAT ACT AAT TCC ATA TAT TCT ATT ACT ATT              1008

 337  Gly Ala Ile Asp His Lys Asp Leu His Pro Pro Tyr Ser Glu Gly Cys               352
1009  GGG GCT ATT GAT CAC AAA GAT CTA CAT CCT CCT TAT TCC GAA GGT TGT              1056

 353  Ser Ala Val Met Ala Val Thr Tyr Ser Ser Gly Ser Gly Glu Tyr Ile               368
1057  TCC GCC GTC ATG GCA GTC ACG TAT TCT TCA GGT TCA GGC GAA TAT ATT              1104

 369  His Ser Ser Asp Ile Asn Gly Arg Cys Ser Asn Ser His Gly Gly Thr               384
1105  CAT TCG AGT GAT ATC AAC GGC AGA TGC AGT AAT AGC CAC GGT GGA ACG              1152

 385  Ser Ala Ala Ala Pro Leu Ala Ala Gly Val Tyr Thr Leu Leu Leu Glu               400
1153  TCT GCG GCT GCT CCA TTA GCT GCC GGT GTT TAC ACT TTG TTA CTA GAA              1200

 401  Ala Asn Pro Asn Leu Thr Trp Arg Asp Val Gln Tyr Leu Ser Ile Leu               416
1201  GCC AAC CCA AAC CTA ACT TGG AGA GAC GTA CAG TAT TTA TCA ATC TTG              1248
```

```
417  Ser Ala Val Gly Leu Glu Lys Asn Ala Asp Gly Asp Trp Arg Asp Ser   432
1249 TCT GCG GTA GGG TTA GAA AAG AAC GCT GAC GGA GAT TGG AGA GAT AGC  1296

433  Ala Met Gly Lys Lys Tyr Ser His Arg Tyr Gly Phe Gly Lys Ile Asp   448
1297 GCC ATG GGG AAG AAA TAC TCT CAT CGC TAT GGC TTT GGT AAA ATC GAT  1344

449  Ala His Lys Leu Ile Glu Met Ser Lys Thr Trp Glu Asn Val Asn Ala   464
1345 GCC CAT AAG TTA ATT GAA ATG TCC AAG ACC TGG GAG AAT GTT AAC GCA  1392

465  Glu Thr Trp Phe Tyr Leu Pro Thr Leu Tyr Val Ser Gln Ser Thr Asn   480
1393 CAA ACC TGG TTT TAC CTG CCA ACA TTG TAT GTT TCC CAG TCC ACA AAC  1440

481  Ser Thr Glu Glu Thr Leu Glu Ser Val Ile Thr Ile Ser Glu Lys Ser   496
1441 TCC ACG GAA GAG ACA TTA GAA TCC GTC ATA ACC ATA TCA GAA AAA AGT  1488

497  Leu Gln Asp Ala Asn Phe Lys Arg Ile Glu His Val Thr Val Thr Val   512
1489 CTT CAA GAT GCT AAC TTC AAG AGA ATT GAG CAC GTC ACG GTA ACT GTA  1536

513  Asp Ile Asp Thr Glu Ile Arg Gly Thr Thr Thr Val Asp Leu Ile Ser   528
1537 GAT ATT GAT ACA GAA ATT AGG GGA ACT ACG ACT GTC GAT TTA ATA TCA  1584

529  Pro Ala Gly Ile Ile Ser Asn Leu Gly Val Val Arg Pro Arg Asp Val   544
1585 CCA GCG GGG ATA ATT TCA AAC CTT GGC GTT GTA AGA CCA AGA GAT GTT  1632

545  Ser Ser Glu Gly Phe Lys Asp Trp Thr Phe Met Ser Val Ala His Trp   560
1633 TCA TCA GAG GGA TTC AAA GAC TGG ACA TTC ATG TCT GTA GCA CAT TGG  1680

561  Gly Glu Asn Gly Val Gly Asp Trp Lys Ile Lys Val Lys Thr Thr Glu   576
1681 GGT GAG AAC GGC GTA GGT GAT TGG AAA ATC AAG GTT AAG ACA ACA GAA  1728

577  Asn Gly His Arg Ile Asp Phe His Ser Trp Arg Leu Lys Leu Phe Gly   592
1729 AAT GGA CAC AGG ATT GAC TTC CAC AGT TGG AGG CTG AAG CTC TTT GGG  1776

593  Glu Ser Ile Asp Ser Ser Lys Thr Glu Thr Phe Val Phe Gly Asn Asp   608
1777 GAA TCC ATT GAT TCA TCT AAA ACA GAA ACT TTC GTC TTT GGA AAC GAT  1824

609  Lys Glu Glu Val Glu Pro Ala Ala Thr Glu Ser Thr Val Ser Gln Tyr   624
1825 AAA GAG GAG GTT GAA CCA GCT GCT ACA GAA AGT ACC GTA TCA CAA TAT  1872

625  Ser Ala Ser Ser Thr Ser Ile Ser Ile Ser Ala Thr Ser Thr Ser Ser   640
1873 TCT GCC AGT TCA ACT TCT ATT TCC ATC AGC GCT ACT TCT ACA TCT TCT  1920

641  Ile Ser Ile Gly Val Glu Thr Ser Ala Ile Pro Gln Thr Thr Thr Ala   656
1921 ATC TCA ATT GGT GTG GAA ACG TCG GCC ATT CCC CAA ACG ACT ACT GCG  1968

657  Ser Thr Asp Pro Asp Ser Asp Pro Asn Thr Pro Lys Lys Leu Ser Ser   672
1969 AGT ACC GAT CCT GAT TCT GAT CCA AAC ACT CCT AAA AAA CTT TCC TCT  2016

673  Pro Arg Gln Ala Met His Tyr Phe Leu Thr Ile Phe Leu Ile Gly Ala   688
2017 CCT AGG CAA GCC ATG CAT TAT TTT TTA ACA ATA TTT TTG ATT GGC GCC  2064

689  Thr Phe Leu Val Leu Tyr Phe Met Phe Phe Met Lys Ser Arg Arg Arg   704
2065 ACA TTT TTG GTG TTA TAC TTC ATG TTT TTT ATG AAA TCA AGG AGA AGG  2112

705  Ile Arg Arg Ser Arg Ala Glu Thr Tyr Glu Phe Asp Ile Ile Asp Thr   720
2113 ATC AGA AGG TCA AGA GCG GAA ACG TAT GAA TTC GAT ATC ATT GAT ACA  2160

721  Asp Ser Glu Tyr Asp Ser Thr Leu Asp Asn Gly Thr Ser Gly Ile Thr   736
2161 GAC TCT GAG TAC GAT TCT ACT TTG GAC AAT GGA ACT TCC GGA ATT ACT  2208

737  Glu Pro Glu Glu Val Glu Asp Phe Asp Phe Asp Leu Ser Asp Glu Asp   752
2209 GAG CCC GAA GAG GTT GAG GAC TTC GAT TTT GAT TTG TCC GAT GAA GAC  2256

753  His Leu Ala Ser Leu Ser Ser Ser Glu Asn Gly Asp Ala Glu His Thr   768
2257 CAT CTT GCA AGT TTG TCT TCA TCA GAA AAC GGT GAT GCT GAA CAT ACA  2304

769  Ile Asp Ser Val Leu Thr Asn Glu Asn Pro Phe Ser Asp Pro Ile Lys   784
2305 ATT GAT AGT GTA CTA ACA AAC GAA AAT CCA TTT AGT GAC CCT ATA AAG  2352

785  Glu Lys Phe Pro Asn Asp Ala Asn Ala Glu Ser Ala Ser Asn Lys Leu   800
2353 CAA AAG TTC CCA AAT GAC GCC AAC GCA GAA TCT GCT TCC AAT AAA TTA  2400

801  Gln Glu Leu Gln Pro Asp Val Pro Pro Ser Ser Gly Arg Ser ***
2401 CAA GAA TTA CAG CCT GAT GTT CCT CCA TCT TCC GGA CGA TCG TGA TTC  2448

2449 GAT ATG TAC AGA AAG CTT CAA ATT ACA AAA TAG CAT TTT TTT CTT ATA  2496
2497 GAT TAT AAT ACT CTC TCA TAC GTA TAC GTA TAT GTG TAT ATG ATA TAT  2544
2545 AAA CAA ACA TTA ATA TCC TAT TCC TTC CGT TTG AAA TCC CTA TGA TGT  2592
2593 ACT TTG CAT TGT TTG CAC CCG CGA ATA AAA TGA AAA CTC CGA ACC GAT  2640
2641 ATA TCA AGC ACA TAA AAG GGG AGG GTC CAA TTA ATG CAT
```

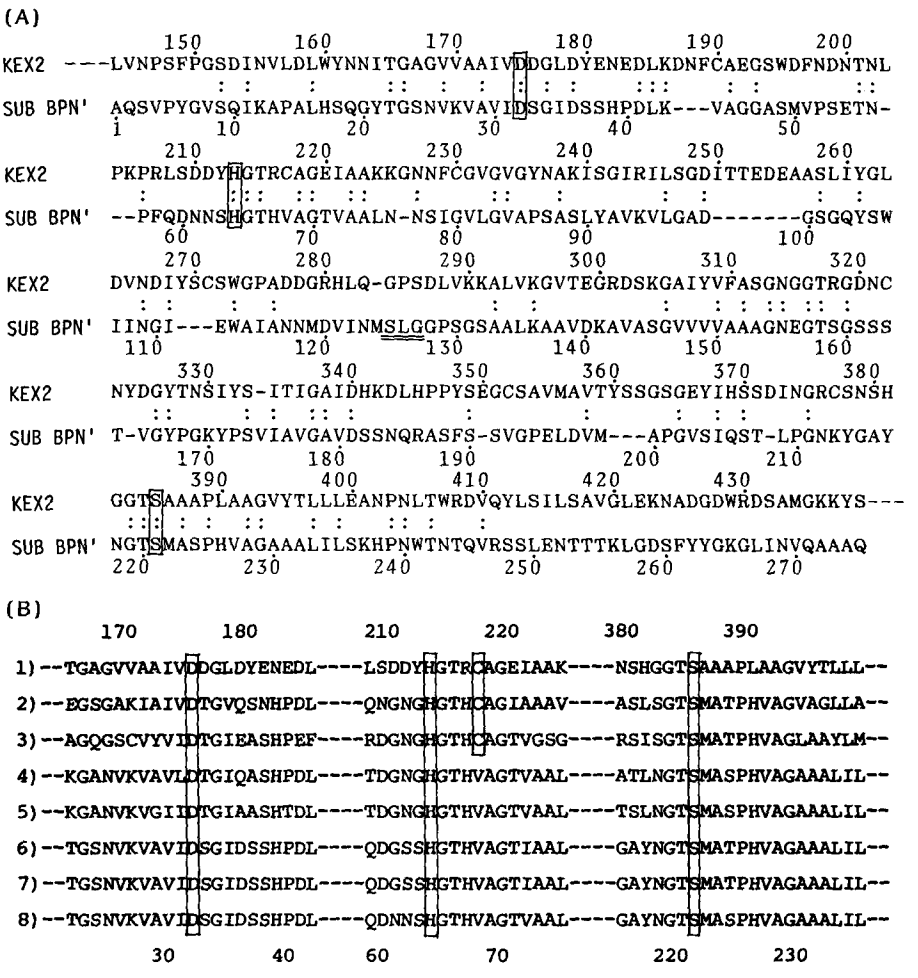

Fig.3. (A) Alignment of the amino acid sequences of the KEX2 protein (upper) and subtilisin BPN' (lower) (19). Sets of identical residues are marked with two dots. Gaps (-) have been inserted to achieve maximum homology. The residues constructing the active site are boxed and the substrate binding site of subtilisin is doubly underlined. (B) Comparison of the amino acid sequences around the active site residues. The residues constructing the active site and the putative free cysteine residues are boxed. Only the residue numbers of KEX2 protease and subtilisin BPN' are specified in the figure. 1) KEX2 protease, 2) thermitase, 3) proteinase K, 4) subtilisin Carlsberg, 5) subtilisin DY, 6) subtilisin I168, 7) subtilisin from Bacillus amylosaccariticus, 8) subtilisin BPN'.

protease superfamily, but are not related to the chymotrypsin-like serine-protease family, the latter being considered to have evolved independently. Sequence comparison led us to classify the KEX2 protease as a subtilisin-like serine-protease.

It has been reported that KEX2 protease activity was inhibited by thiol-directed reagents such as monoiodoacetate and heavy metal ions but not by PMSF, a general serine-protease inhibitor (4,5). Similar results were obtained with membrane-bound proteases in yeast, as reported by Wolf's group

and by ours (6-8). As had been known, two fungal subtilisin-like serine proteases, proteinase K and thermitase, are inactivated by thiol-directed reagents. They have a free cysteine residue located immediately adjacent to a histidine residue in the active site (19). KEX2 protein contains eight Cys residues, one of which (Cys-217) is situated in a homologous position (Fig.3B). Inactivation of KEX2 protease by thiol-reagents may be due to the blockage of this free Cys residue. Our preliminary study showed that KEX2 protease was inactivated by inhibitors for serine protease, such as PMSF, DFP, and pAPMSF, at a high concentration of 10 mM. (Details will be reported elsewhere.) These results, taken together, further support the conclusion that KEX2 protease belongs to the serine protease family, as also suggested by Fuller, et al. (9).

KEX2 protease activity exhibits Ca^{2+}-dependency (5-8), analogous to calcium-activated neutral protease (calpain) found in mammalian tissues. However, the KEX2 product does not contain sequences with obvious resemblance to calpains or other calcium-binding proteins. Further studies are necessary to understand the molecular mechanisms by which calcium regulates the enzyme.

The structural organization of KEX2 protein is similar to those of other processing enzymes ever characterized, including yeast KEX1-encoded carboxypeptidase B-like protease (20), and peptide C-terminal α-amidating enzymes from Xenopus laevis skin (21) and bovine pituitary (22). These enzymes have in common two hydrophobic regions, a putative signal sequence at the N-terminus and a putative membrane-spanning domain near the C-terminus, both of which may function to localize and anchor them to the intracelluler target organelles. The active domain of each enzyme, situated at the N-terminal side of the membrane-spanning domain, should be oriented inside the secretory vesicles.

Paired basic residues are conserved on various bioactive peptide precursors to serve as sites of proteolytic processing in a wide range of eukaryotic species from yeast to mammals. This fact suggests a possibility that endopeptidases similar to KEX2 gene product may be involved in precursor processing in mammalian tissues (23,24). The cloned KEX2 gene may serve as a probe for searching these enzymes, which have not yet been well-characterized.

ACKNOWLEDGEMENTS

We thank Dr. Reed Wickner for providing the strain carrying kex2-8 mutation, Miss Kumiko Matsubara for technical assistance. We also thank Drs. Tatsuo Nakayama, Yoshiko Setoguchi, and Naoya Akizuki for valuable comments.

REFERENCES

1. Docherty,K. & Steiner,D.F. (1982) Ann.Rev.Physiol. 44, 625-638.
2. Kurjan,J. & Herskowitz,I. (1982) Cell 30, 933-943.

3. Bostian,K.A., Ellott,Q., Bussey,H., Burn,V., Smith,A. & Tipper,D.J. (1984) Cell 36, 741-751.
4. Julius,D., Brake,A., Blair,L., Kunisawa,R. & Thorner,J. (1984) Cell 37, 1075-1089.
5. Fuller,R., Brake,A. & Thorner,J.(1986) in Microbiology 1986, ed. Levine,L. (American Society for Microbiology, Washington,DC), pp. 273-278.
6. Achstetter,T. & Wolf,D.H. (1985) EMBO J. 4, 173-177.
7. Wagner,J.C., Escher,C. & Wolf,D.H. (1987) FEBS Lett. 218, 31-34.
8. Mizuno,K., Nakamura,T., Takada,K., Sakakibara,S. & Matsuo,H. (1987) Biochem.Biophys.Res.Commun. 144, 807-814.
9. Fuller,R.S., Sterne,R.E. & Thorner,J. (1988) Ann.Rev.Physiol. 50, 345-362.
10. Davis,R.W. et al. (1980) Methods in Enzymology 65, 404-411.
11. Maniatis,T., Fritsch,E.F. & Sambrook,J. (1982) in Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratories, New York).
12. Ito,H., Fukuda,Y., Murata,K. & Kimura,A. (1983) J.Bacteriol. 153, 163-168.
13. Hanahan,D. (1983) J.Mol.Biol. 166, 557-580.
14. Wickner,R.B. (1974) Genetics 76, 423-432.
15. Wickner,R.B. & Leibowitz,M.J. (1976) Genetics 82, 429-442.
16. Messing,J. (1983) Methods in Enzymology 101, 20-78.
17. Kraut,J. (1971) in "The Enzymes", vol.3, ed. Boyer, P.D., (Academic Press, New York), pp. 547-560.
18. Wells,J.A., Ferrari,E., Henner,D.J., Estell,D.A. & Chen,E.Y. (1983) Nucl. Acids Res. 11, 7911-7925.
19. Betzel,C., Pal,G.P., Struck,M., Jany,K-D. & Saenger,W. (1986) FEBS Lett. 197, 105-110.
20. Dmochowska,A., Dignard,D., Henning,D., Thomas,D.Y. & Bussey,H. (1987) Cell 50, 573-584.
21. Ohsuye,K., Kitano,K., Wada,Y., Fuchimura,K., Tanaka,S., Mizuno,K. & Matsuo,H. (1988) Biochem. Biophys. Res. Commun. 150, 1275-1281.
22. Eipper,B.A. et al. (1987) Mol. Endocrinol. 1, 777-790.
23. Davidson,H.W., Peshavaria,M. & Hutton,J.C. (1987) Biochem. J. 246, 279-286
24. Brennan,S.O. & Peach, R.J. (1988) FEBS Lett. 229, 167-170.