Molecular Cloning and Sequencing of Genomic DNA Encoding Aminopeptidase I from Saccharomyces cerevisiae*

Yie-Hwa Chang‡‡ and John A. Smith‡‡¶¶

From the Departments of ‡Molecular Biology and ¶Pathology, Massachusetts General Hospital and the Departments of §§Genetics and ¶¶Pathology, Harvard Medical School, Boston, Massachusetts 02114

Yeast aminopeptidase I is a vacuolar enzyme, which catalyzes the removal of amino acids from the NH₂ terminus of peptides and proteins (Frey, J., and Rohn, K.-H. (1978) Biochim. Biophys. Acta 527, 31–41). A yeast genomic DNA encoding aminopeptidase I was cloned from a yeast EMBL3A library and sequenced. The DNA sequence encodes a precursor protein containing 514 amino acid residues. The “mature” protein, whose NH₂-terminal sequence was confirmed by automated Edman degradation, consists, based on only the DNA sequence, of 469 amino acids. A 45-residue presequence contains positively and negatively charged as well as hydrophobic residues, and its NH₂-terminal residues could be arrayed in an amphiphilic α-helix. This presequence differs from the signal sequences which direct proteins across bacterial plasma membranes and endoplasmic reticulum or into mitochondria. It remains to be established how this unique presequence targets aminopeptidase I to yeast vacuoles and how this sorting utilizes classical protein secretory pathways. Further, the aminopeptidase I gene, localized previously by genetic mapping to yeast chromosome XI and called the LAP4 gene (Trumbly, R. J., and Bradley, G. (1983) J. Bacteriol. 156, 36–48), was determined by DNA blot analyses to be a single copy gene located on chromosome XI.

Yeast aminopeptidase I was first isolated and characterized as a high molecular weight protein by Johnson in 1941 (1). Aminopeptidase I (also called polypeptidase I (1), leucine aminopeptidase IV (2), aminopeptidase III (3), and aminopeptidase yscI (4, 5)) has been localized to yeast vacuoles (equivalent to lysosomes found in higher eukaryotic organisms (6, 7)). The enzyme was established to be a glycoprotein, containing 12% carbohydrate, as well as a Zn²⁺-containing metalloenzyme, which can be inactivated by metal-chelating agents and specifically activated by Zn²⁺ and Cl⁻ (1, 2, 8, 9).

Mutations in the Saccharomyces cerevisiae LAP4 gene, which have been generated by Trumbly and Bradley (2), eliminated the activity of aminopeptidase I, and the LAP4 gene has been genetically mapped to the left arm of chromosome XI. However, since yeast strains with mutations in LAP4 displayed apparently normal growth rates, aminopeptidase I presumably does not have a vital role in cell physiology or its absence can be compensated by other aminopeptidases

1 Portions of this paper (including “Experimental Procedures” and Figs. 1, 2, and 4–6) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

2. Prior to this report, the LAP4 gene has not been cloned and sequenced.

It is known that vacuolar proteases (e.g. proteinase A (10), proteinase B (11), and carboxypeptidase Y (12)) are synthesized as preproteins, and in the case of carboxypeptidase Y, its presequence is proven to contain the vacuolar sorting determinant (13, 14). Although aminopeptidase I has been localized to yeast vacuoles (7), its maturation process and intracellular sorting remain unknown.

This paper reports the cloning, sequencing, and chromosomal location of the gene encoding the aminopeptidase I precursor protein, which contains a presequence differing from other signal sequences which direct proteins across bacterial plasma membranes and endoplasmic reticulum or into mitochondria. This cloned gene will facilitate studies of the processing, catalytic mechanism, and regulation of the biosynthesis of aminopeptidase I.

EXPERIMENTAL PROCEDURES

RESULTS AND DISCUSSION

Cloning and Sequencing of the Yeast Genomic DNA Encoding Aminopeptidase I—Aminopeptidase I was purified by the method of Metz and Rohm (8). The M, of the native enzyme (570,000) and its subunit (50,000) agree with the results of others, who determined that the protein was a multimeric glycoprotein (1, 2, 7, 8). In order to clone the encoding genomic DNA, the NH₂-terminal sequence of the first 18 amino acids of the enzyme was determined, and this sequence formed the basis for the synthesis of oligonucleotide probes (Fig. 1). Fig. 3 shows the complete DNA sequence and deduced protein sequence for yeast aminopeptidase I. The most 5’ ATG, which obeys Kozak’s rule (32) is presumed to be the initiation codon. An open reading frame of 1542 nucleotides (encoding aminopeptidase I and its presequence) extends from this initiation codon ATG to a stop codon ACT (marked End). The open reading frame sequence encodes a protein with 514 amino acid residues and with a calculated molecular weight of 57,175. The NH₂-terminal sequence derived from protein sequence analysis of the first 18 amino acid residues of the purified aminopeptidase I matches residues 1–18 of aminopeptidase I deduced from the DNA sequence.

A TATA-like sequence, TATAAG, is located 102 bases upstream of the initiation codon. CACACCAACA, which is located within 20–30 nucleotides of the initiation codon, may be a ribosomal binding site (33, 34). Furthermore, 150 bases downstream from the stop codon, there is a putative polyad-
Sequence of Yeast Aminopeptidase I

![Sequence of Yeast Aminopeptidase I](image)

**FIG. 7.** Amphiphilic a-helical structure within the presence of yeast aminopeptidase I. The first amino acid residue of the putative signal peptide of the aminopeptidase I precursor is numbered -45. The helix displays residues from -41 to -27 (42). The dotted line demarcates the hydrophobic and hydrophilic faces of the predicted amphiphilic a-helix.

**FIG. 3.** Genomic DNA and deduced amino acid sequence of yeast aminopeptidase I. The clone begins with a noncoding region containing five in-frame stop codons. The most 5' ATG of the open reading frame which obeyed Kozak's rule (32) was assigned as the initiation codon. A TATA-like sequence, TATAAG, is located 102 nucleotides upstream of this initiation codon. The CACAACCA sequence, which is located within 20-30 nucleotides of the initiation enalation signal, AAATAAA (35). In addition, there are four potential N-glycosylation sites (36), indicated by asterisks (Fig. 3).

*Genomic and Chromosomal DNA Blot Analyses—*Taken together, the genomic and chromosomal DNA blot analyses (Figs. 4 and 5, respectively) indicate that there is a single copy of the aminopeptidase I gene located on yeast chromosome XI, which is consistent with previous genetic mapping of the LAP4 gene by Trumbly and Bradley (2).

**Sequence Similarities between Aminopeptidase I and Other Proteins—**The computer program FASTP (37) was used to compare the derived amino acid sequence of yeast aminopeptidase I with the entire National Biomedical Research Foundation protein sequence data bank, and no significant sequence similarity with any other proteins was found. The GAP program (38) was used to compare the sequence of aminopeptidase I with bovine leucine aminopeptidase (39), Escherichia coli peptidase N (40), and E. coli methionine-specific aminopeptidase (41), and no significant sequence similarity was observed.

**Evidence for an Aminopeptidase I Precursor—**Immunoblot analysis of SDS-PAGE purified and electroeluted aminopeptidase I detected a single protein of \( M_r = 50,000 \) that binds rabbit polyclonal anti-API antibody (Fig. 6, lane a), and immunoblot analysis of an extract of total yeast protein...
detects two proteins with $M$ values of 50,000 and 57,000, respectively (Fig. 6, lane b).

The open reading frame of the genomic DNA contains 1542 nucleotides encoding an aminopeptidase I precursor containing 514 amino acid residues. The calculated molecular weight of the precursor is 57,175, which agrees with the $M$ of the precursor estimated by immunoblot analysis (Fig. 6, lane b). A comparison of this deduced protein sequence with the NH$_2$-terminal sequence of the purified aminopeptidase I indicates that the purified protein lacks 45 residues at the NH$_2$ terminus of the precursor protein. Further, the calculated molecular weight of the processed precursor protein (lacking its NH$_2$-terminal 45 amino acid residues) is 51,663, which agrees with the $M$ of the purified enzyme estimated by SDS-PAGE and immunoblot analyses (Fig. 6, lanes a and b). However, Metz and Rohm (8) have determined that aminopeptidase I has a $M$ of 51,000 (including 12% carbohydrates) and calculated from amino acid analysis that the protein lacking carbohydrate had a estimated molecular weight of 44,800. If this is the case, then additional COOH-terminal processing must occur in order to generate the “mature” protein, described by Metz and Rohm (8). How and where such processing occurs and whether or not it is mediated by vacuolar proteases, as has been demonstrated for the maturation of yeast protease B (11), remains to be established. In addition, it was observed that if the protein was purified from yeast cells grown to stationary phase, rather than to log phase as was the case for the protein whose NH$_2$-terminal sequence is shown in Fig. 1, that the NH$_2$-terminal sequence began at the Tyr, located at residue 4, but otherwise matched the deduced protein sequence. It is not yet clear what such additional NH$_2$-terminal processing contributes to the structure or function of aminopeptidase I.

Structure of the Presequence of Yeast Aminopeptidase I—Fig. 7 shows the NH$_2$-terminal residues -45 to -27 of the aminopeptidase I precursor protein displayed in a helical wheel (42). Residues -45 to -42 are not arrayed in the $\alpha$-helix. The presence of both a hydrophobic (consisting of primarily of Leu and Ile residues) and a hydrophilic (consisting of Lys, Arg, Glu, Gln, and Met residues) face in this predicted $\alpha$-helix is characteristic of an amphipathic $\alpha$-helix (43).

Such a presequence containing a stretch of charged (positively and negatively) and hydrophobic residues capable of adopting an amphipathic $\alpha$-helical conformation (Fig. 7) is not typical of the presequences associated with other yeast vacuolar proteins, including protease A (10), protease B (11), and carboxypeptidase Y (12), and other eukaryotic secretory proteins (for a review, see Ref. 44). Their presequences are composed of a stretch of apolar amino acid residues, which are responsible for their targeting to endoplasmic reticulum (for a review, see Refs. 45 and 46). The aminopeptidase I presequence also differs from mitochondrial presequences, which are amphipathic $\alpha$-helices that are primarily containing positively charged residues (for a review, see Ref. 47). In addition, although the presequence of yeast carboxypeptidase Y has been demonstrated to target carboxypeptidase Y to the yeast vacuole (13, 14), there is no obvious equivalent sequence responsible for the sorting of aminopeptidase I into vacuoles. It remains to be established how this unique presequence targets aminopeptidase I to yeast vacuoles and how this sorting utilizes classical protein secretory pathways. Fusion proteins are currently being used to evaluate which elements of the presequence of aminopeptidase I are responsible for its intracellular sorting.

Acknowledgments—We wish to acknowledge the skilled technical assistance of Paul Gallant and Gene Hannigan.

REFERENCES
Sequence of Yeast Aminopeptidase I


Oligonucleosome Synthesis

The oligonucleosomes probes and primers were synthesized on an Applied Biosystems 380 DNA synthesizer and purified by polyethylene glycol electrophoresis to the applied Biosearch Manual. The oligonucleosomes were isolated from solution of oligonucleosomes containing 100 mM NaCl.

Cloning and Sequencing of the Yeast Genetic DNA Encoding Aminopeptidase I

The sheared genomic DNA was transferred onto a nitrocellulose filter and hybridized to radioactive DNA labeled with [32P]dATP. The hybridization was performed at 55°C for 16 h. The filters were washed in 0.2XSSC at 65°C and autoradiographed. The oligonucleosomes probes were separated by size and used to complete the sequencing of both DNA strands (Fig. 3).

DNA Blot Analysis of Yeast Chromosomal DNA

Total DNA was isolated from yeast strain ZY16-1A (21). The DNA (20 μg) was digested with HpaI and HaeIII and electrophoresed on a 0.7% agarose gel. The digested DNA was transferred to a nitrocellulose filter and hybridized to radioactive DNA labeled with [32P]dATP. The hybridization was performed at 55°C for 16 h. The filters were washed in 0.2XSSC at 65°C and autoradiographed. The oligonucleosomes probes were separated by size and used to complete the sequencing of both DNA strands (Fig. 3).

Preparation and Affinity Purification of Polyclonal Antibodies Against Yeast Aminopeptidase I

Antisera were purified, and radioactive antibodies 10°C were isolated by sequential digestion of complete Freund's adjuvant into a new 15 mL syringe and used for the purification of antibodies. The purified antibodies were digested for 1 h at 37°C and dialyzed against PBS. The antibody solution was used for immunization of rabbits.

Inmunological Analysis

Protein concentration at 280 nm was determined using the Bio-Rad protein assay. The protein was digested with trypsin and the digested protein was electrophoresed on a 12% SDS-PAGE gel. The digested protein was transferred to a nitrocellulose filter and hybridized to radioactive DNA labeled with [32P]dATP. The hybridization was performed at 55°C for 16 h. The filters were washed in 0.2XSSC at 65°C and autoradiographed. The oligonucleosomes probes were separated by size and used to complete the sequencing of both DNA strands (Fig. 3).

Experimental Procedures

Materials

Saccharomyces cerevisiae was grown in shake flasks in 1 L Erlemeyer flasks at 30°C. Yeast chromosomal DNA was isolated as described previously (25). The DNA (10 μg) was digested with HpaI and HaeIII and electrophoresed on a 0.7% agarose gel. The digested DNA was transferred to a nitrocellulose filter and hybridized to radioactive DNA labeled with [32P]dATP. The hybridization was performed at 55°C for 16 h. The filters were washed in 0.2XSSC at 65°C and autoradiographed. The oligonucleosomes probes were separated by size and used to complete the sequencing of both DNA strands (Fig. 3).

Sequencing and Amplifying the Yeast Genomic DNA Encoding Aminopeptidase I

The sheared genomic DNA was transferred onto a nitrocellulose filter and hybridized to radioactive DNA labeled with [32P]dATP. The hybridization was performed at 55°C for 16 h. The filters were washed in 0.2XSSC at 65°C and autoradiographed. The oligonucleosomes probes were separated by size and used to complete the sequencing of both DNA strands (Fig. 3).

Preparation and Affinity Purification of Polyclonal Antibodies Against Yeast Aminopeptidase I

Antisera were purified, and radioactive antibodies 10°C were isolated by sequential digestion of complete Freund's adjuvant into a new 15 mL syringe and used for the purification of antibodies. The purified antibodies were digested for 1 h at 37°C and dialyzed against PBS. The antibody solution was used for immunization of rabbits.

Inmunological Analysis

Protein concentration at 280 nm was determined using the Bio-Rad protein assay. The protein was digested with trypsin and the digested protein was electrophoresed on a 12% SDS-PAGE gel. The digested protein was transferred to a nitrocellulose filter and hybridized to radioactive DNA labeled with [32P]dATP. The hybridization was performed at 55°C for 16 h. The filters were washed in 0.2XSSC at 65°C and autoradiographed. The oligonucleosomes probes were separated by size and used to complete the sequencing of both DNA strands (Fig. 3).

Cloning and Sequencing of the Yeast Genetic DNA Encoding Aminopeptidase I

The sheared genomic DNA was transferred onto a nitrocellulose filter and hybridized to radioactive DNA labeled with [32P]dATP. The hybridization was performed at 55°C for 16 h. The filters were washed in 0.2XSSC at 65°C and autoradiographed. The oligonucleosomes probes were separated by size and used to complete the sequencing of both DNA strands (Fig. 3).

DNA Blot Analysis of Yeast Chromosomal DNA

Total DNA was isolated from yeast strain ZY16-1A (21). The DNA (20 μg) was digested with HpaI and HaeIII and electrophoresed on a 0.7% agarose gel. The digested DNA was transferred to a nitrocellulose filter and hybridized to radioactive DNA labeled with [32P]dATP. The hybridization was performed at 55°C for 16 h. The filters were washed in 0.2XSSC at 65°C and autoradiographed. The oligonucleosomes probes were separated by size and used to complete the sequencing of both DNA strands (Fig. 3).

Preparation and Affinity Purification of Polyclonal Antibodies Against Yeast Aminopeptidase I

Antisera were purified, and radioactive antibodies 10°C were isolated by sequential digestion of complete Freund's adjuvant into a new 15 mL syringe and used for the purification of antibodies. The purified antibodies were digested for 1 h at 37°C and dialyzed against PBS. The antibody solution was used for immunization of rabbits.
Sequence of Yeast Aminopeptidase I

Figure 1. NH₂-terminal sequence of yeast aminopeptidase I and design of synthetic oligonucleotide probes. NH₂-terminal sequence of SSH-PAGE purified and electroeluted (18) aminopeptidase I was determined by automated Edman degradation procedures. Degenerate oligonucleotide probes a and b were designed according to the corresponding protein sequence, and these were used sequentially to screen a yeast genomic library.

Figure 2. Restriction map and sequencing strategy for the genomic DNA encoding yeast aminopeptidase I. Solid arrows denote the direction and extent of 5'- or 3'-primer initiated DNA sequencing, and dashed arrows denote synthetic oligonucleotide-pigated DNA sequencing of various subclones contained in recombinant plasmid DNA. The yeast aminopeptidase I precursor, containing 51 amino acid residues, is represented by the thick line, and the flanking regions by the thin lines. Abbreviations for restriction cleavage sites are as follows: E, EcoRI; F, PstI; H, Hind.

Figure 3. DNA blot analysis of yeast chromosomal DNA. The yeast chromosomal aminopeptidase gene (i.e., a "Saccharomyces chrom-α-hybridizer" from Clontech) was probed with the [³²P]-labeled 5.71 restriction fragment (2.2 kb) of yeast aminopeptidase I gene (30)(Fig. 2).

Figure 4. DNA blot analysis of yeast genomic DNA. Yeast total DNA was digested with either EcoRI (lane a) or XbaI (lane b), run on a 0.7% agarose gel, transferred to a nitrocellulose membrane, and probed with [³²P]-labeled PstI restriction fragment (1.2 kb) of the yeast aminopeptidase I gene (29.30)(Fig. 2).

Figure 5. Immuno blot analysis of purified yeast aminopeptidase I and total yeast protein. Lane a shows the immunoblot for the SSH-PAGE purified and electroeluted (18) aminopeptidase I. Lane b shows the immunoblot of an extract of total yeast protein, prepared as described in "Experimental Procedures". Purified aminopeptidase I and the extract of total yeast protein were electrophoresed on a SDS-PAGE gel (8%)(31) and then electroblotted to a nitrocellulose filter (1.2). Rabbit antisera to purified yeast aminopeptidase I antibody and a polyclonal aminopeptidase I antiserum were used as the first antibodies in this procedure. The apparent molecular weights were calculated using protein standards including rabbit muscle phosphorylase (97,400), bovine serum albumin (66,200), egg albumin (45,000), and carbonic anhydrase (29,000).