Biochem. J. (1992) 282, 539–543 (Printed in Great Britain)

On the catalytic mechanism of prokaryotic leader peptidase 1

Michael T. BLACK,* John G. R. MUNN and Aileen E. ALLSOP
SmithKline Beecham Pharmaceuticals, Brockham Park, Betchworth, Surrey RH3 7AJ, U.K.

The catalytic mechanism of leader peptidase 1 (LP1) of the bacterium *Escherichia coli* has been investigated by a combination of site-directed mutagenesis, assays of enzyme activity *in vivo* utilizing a strain of *E. coli* which has a conditional defect in LP1 activity, and gene cloning. The biological activity of mutant forms of *E. coli* LP1 demonstrates that this enzyme belongs to a novel class of proteinases. The possibility that LP1 may be an aspartyl proteinase has been excluded on the basis of primary sequence comparison and mutagenesis. Assignment of LP1 to one of the other three recognized classes of proteinases (metalloproteinases, thiol proteinases and the classical serine proteinases) can also be excluded, as it is clearly demonstrated that none of the histidine or cysteine residues within LP1 are required for catalytic activity. The *Pseudomonas fluorescens lep* gene has been cloned and sequenced and the corresponding amino acid sequence compared with that of *E. coli* LP1. The *E. coli* LP1 and *P. fluorescens* LP1 primary sequences are 50% identical after insertion of gaps. The *P. fluorescens* LP1 has 39 fewer amino acids, a calculated molecular mass of 31903 Da and functions effectively *in vivo* in *E. coli*. None of the cysteine residues and only one of the histidine residues which are present in *E. coli* LP1 are conserved in sequence position in the *P. fluorescens* LP1 enzyme. The possibility that LP1 is a novel type of serine proteinase is discussed.

INTRODUCTION

Targetting of proteins which are synthesized in the cytoplasm to the various extracellular and subcellular locations necessarily involves translocation across one or more lipid bilayers that are normally impermeable to large molecules and to small charged molecules. An essential feature of many translocated proteins involved in overcoming this barrier is that they are initially synthesized as precursor proteins (preproteins) with an *N*-terminal extension known as a leader, or signal, peptide which is proteolytically cleaved from the preprotein during, or shortly after, translocation to yield the mature protein [1]. Whereas eukaryotic leader peptidase activity is associated with protein complexes consisting of between two and six polypeptides [2–4], depending on the species, the prokaryotic leader peptidase (LP1) consists of a single polypeptide [5]. The genes encoding LP1 (*lep*) of *Escherichia coli* [5] and *Salmonella typhimurium* [6] have been cloned and sequenced and the corresponding amino acid sequences share 94% sequence identity. The *E. coli lep* gene encodes an essential protein consisting of 323 amino acids with a calculated molecular mass of 35988 Da, which possesses two hydrophobic domains at the *N*-terminus that anchor the protein to the cytoplasmic membrane exposing the bulk of the protein to the periplasmic space [7].

LP1 activity is highly specific, as it cleaves only preproteins and typically cleaves at a single precise site, even though sequences near cleavage sites are not unique [8]. Despite the lack of sequence conservation among leader peptides, they do share some common features: they possess an *N*-terminal basic region of between two and eight residues and a hydrophobic central core which is followed by a short, less hydrophobic, region. This region often begins at position P6 or P5 with a helix-breaker such as glycine or proline and ends at the cleavage site, with small uncharged side chains such as glycine or alanine commonly occupying the P1 and P3 positions [8]. It therefore appears that leader peptidases may be unusual proteinases, in that the secondary structure of the leader sequence may be the major determinant of enzyme–substrate recognition.

Despite the wealth of knowledge which has been accumulated concerning the substrate specificity, *trans*-membrane topology and interactions with other components of the translocation machinery, the mechanism of the proteolytic cleavage of the scissile bond by LP1 remains enigmatic. Utilizing two main strategies, site-directed mutagenesis and gene cloning, an attempt has been made to determine whether LP1 can be assigned to one of the four groups which represent the known mechanisms of enzyme-catalysed peptide bond hydrolysis. These groups are the serine, thiol-, metallo- and aspartyl proteinases.

MATERIALS AND METHODS

Plasmids and site-directed mutagenesis

A 1.3 kb *Sal*I–*Bam*HI fragment of DNA containing the *lep* gene from *E. coli* strain HJM114 was excised from plasmid pRD8 [9] and cloned into the polylinker site of phasmid vector pTZ19R to generate pTZlep1. This vector was subjected to oligonucleotide-directed site-directed mutagenesis in order to introduce two *Bgl*II sites positioned 25 and 265 nucleotides upstream of the initiator methionine of the *lep* gene, generating pTZlep5. This construct was used as a template for generation of mutations in the *lep* gene. Plasmids pRD8 and pTZlep5 were cut with the restriction endonuclease *Eco*RI, and the 370 bp *Eco*RI fragment from pTZlep5 was cloned into the pRD8 vector fragment to generate pNA1. This had the effect of introducing a *Sac*I site and a unique *Kpn*I site at the 3′ end of the *lep* gene. The 1.3 kb *Sal*I–*Kpn*I fragment containing the *lep* gene was removed from pNA1 and replaced with the equivalent fragment from pTZlep5, thereby introducing the two *Bgl*II sites at the 5′ end of the *lep* gene, and generating pNA2. pNA3 was produced by excising and discarding the 240 bp *Bgl*II fragment from pNA2, thereby removing a sequence of DNA which possesses some promoter activity [10], and self-ligating. LP1 production from

Abbreviation used: LP1, leader peptidase 1.

* To whom all correspondence and reprint requests should be addressed.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession no. X56466.

plasmid pNA3 is thus under control of the *ara* promoter, as in pRD8, and is stimulated by addition of arabinose to the growth medium.

Plasmids were amplified in *E. coli* strain TG1, unless otherwise stated, and cells were grown in Luria broth. The Amersham site-directed mutagenesis kit was used to generate mutants according to the manufacturers' instructions. Mutants were selected by sequencing DNA, derived from randomly selected colonies, by the dideoxynucleotide chain-termination method of Sanger *et al.* [11], and compressions were resolved by sequencing with 7-deaza-deoxyGTP and 7-deaza-deoxyITP. In all cases the entire gene was sequenced to ensure that no secondary unwanted mutations had occurred.

Enzyme activity

The enzymic activity of mutant forms of LP1 was assessed by expressing the protein in *E. coli* strain IT41 [10], in which the genomic copy of the *lep* gene has been replaced with a mutant form of *lep* which has a stop codon inserted at nucleotide positions corresponding to an undefined amino acid position within the *C*-terminal part of the protein. The resultant enzyme possesses activity at 30 °C but is inactive at 42 °C. Cell growth and survival at high temperature is therefore dependent upon complementation of this conditional defect by the presence of plasmid-encoded LP1 activity.

Gene cloning and sequence determination

The *Pseudomonas fluorescens* homologue of the *E. coli lep* gene was cloned from a λ EMBL3 library of genomic DNA isolated from *P. fluorescens* and was a gift from Dr. Neil Clarke and Dr. Chris Franklin, University of Birmingham, Birmingham, U.K. The library was plated out on *E. coli* host strain LE392 to a plaque density of approx. 500/55 cm² plate. The resultant plaques and bacterial lawn were transferred to nylon (Nylon N⁺, Amersham) discs and processed according to the manufacturers' instructions. Discs were examined by hybridization screening using an Enhanced Chemiluminescence (Amersham) detection system to detect hybridization to an *E. coli lep* probe. The probe consisted of the 1.0 kb *Bgl*II–*Kpn*I fragment from pNA2 which was labelled according to the manufacturers' recommendations. All hybridization solutions were used as supplied by the manufacturers, except that the hybridization buffer contained an additional 1.0 M-NaCl and the secondary and tertiary washes were carried out at 28 °C and 35 °C respectively in 4 × SSC (0.6 M-NaCl/0.06 M-sodium citrate, pH 7.0). Plasmid pNA109 was constructed as follows. Plasmid pBR322 was digested with *Eco*RI and *Pvu*II and the 2.3 kb vector fragment was recovered. The chimaeric λ clone PF5, which contains ~ 20 kb of genomic DNA from *P. fluorescens*, was digested with *Eco*RI and *Sma*I, and the resultant 3.4 kb fragment, which encompasses the entire *P. fluorescens lep* gene, was cloned into the pBR322 vector fragment to yield pNA109. Other molecular biological manipulations were performed according to standard procedures [12], and all reagents used were of the highest grade obtainable.

RESULTS AND DISCUSSION

Proteolytic enzymes can be divided into four groups according to mechanism and the nature of the amino acid side chains or prosthetic group which the enzyme employs to effect the hydrolysis of the peptide bond [13]. The characteristic motif of the aspartyl proteinases is the twofold occurrence of the tripeptide Asp-Thr-Gly, usually followed by a serine or threonine residue in most examples, although alanine is a common substitution in examples from the retroviruses (the motif occurs only once in the viral proteinases, which operate as dimers) [14]. The only known exception to the absolute conservation of the active-site triplet is the sequence Asp-Ser-Gly, which occurs in the processing proteinase from Rous Sarcoma Virus [15]. Enzymes which belong to the aspartyl proteinase class should thus be easy to detect provided that the amino acid sequence is known. The remaining three classes all utilize one or more histidine residues which act at or near to the active site. Metalloproteinases contain a Zn^{2+} ion which serves to polarize the carbonyl bond of the substrate and which is co-ordinated by two histidine residues [16]. The serine and thiol (cysteine) proteinases rely on the generation of a powerful nucleophile to effect bond hydrolysis; in these cases a histidine imidazole group abstracts a proton from either a serine hydroxyl or a cysteine thiol group, thereby generating the nucleophile [16].

The obvious first approach to the classification of LP1 is to

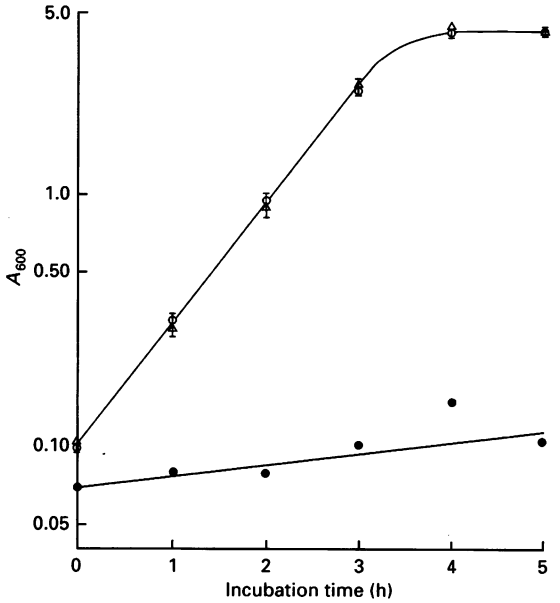

Fig. 1. Effect of the presence or absence of wild-type or mutant forms of LP1 on growth of *E. coli* IT41 at the non-permissive temperature

The increase in cell density of *E. coli* IT41 cells (monitored as absorbance at 600 nm) which were transformed with plasmid pNA3 encoding wild-type or mutant forms of LP1, or with pBR322, is shown. Single bacterial colonies were picked and grown overnight at 42 °C in 2 ml of Luria broth supplemented with ampicillin (100 μg/ml), oxytetracycline (5 μg/ml) and 0.4% arabinose and subsequently diluted 100-fold into 20 ml of Luria broth containing identical supplements and incubated at 42 °C with vigorous shaking. Aliquots of cells were withdrawn at hourly intervals and absorbance at 600 nm was determined. △, pNA3; ○, average of all six mutants; ●, pBR322. The values for pBR322 have been multiplied by 10 for display purposes.

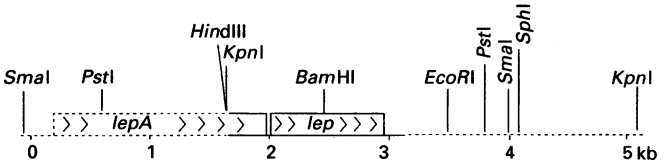

Fig. 2. Partial restriction map of DNA from λPF5

Restriction map of *P. fluorescens* DNA encompassing the *lep* and *lepA* genes. Regions which correspond to unsequenced DNA are shown as broken lines, whereas continuous lines represent regions for which the sequence has been determined. The boundary of the 5′ end of the *lepA* gene is approximate.

```
lepA
ccggaagcgcaagctgcttgagaagcaaaaggccggtaaaaaacgcatgaagcaagtagg
 ArgLysArgLysLeuLeuGluLysGlnLysAlaGlyLysLysArgMetLysGlnValGl
                                                    RBS
taacgtggaagttccacaagaagccttccttgcggtgctcaggttggatagttaggtcct
yAsnValGluValProGlnGluAlaPheLeuAlaValLeuArgLeuAspSerEND

lep
ATGTCACTAAATTTCCCGCTGTTGCTGGTTATCGCCGTCGCCGTTTGCGGCCTCCTGGCG    60
MetSerLeuAsnPheProLeuLeuLeuValIleAlaValAlaValCysGlyLeuLeuAla    20

TTGCTCGATCTGGTGTTCTTCGCCCCGCGTCGGCGGTCGGCCATTGCTTCCTATCAGGGC    120
LeuLeuAspLeuValPhePheAlaProArgArgArgSerAlaIleAlaSerTyrGlnGly    40

AGCGTCAGCCAGCCCGATGCGGTGGTGATCGAGAAGCTGAACAAAGAGCCCTTGCTGGTT    180
SerValSerGlnProAspAlaValValIleGluLysLeuAsnLysGluProLeuLeuVal    60

GAATACGGCAAGTCGTTCTTCCCGGTGTTGTTCATCGTGCTGGTGCTGCGTTCGTTCCTG    240
GluTyrGlyLysSerPhePheProValLeuPheIleValLeuValLeuArgSerPheLeu    80

GTGGAGCCGTTCCAGATCCCTTCCGGGTCGATGAAGCCGACCCTGGACGTGGGCGACTTT    300
ValGluProPheGlnIleProSerGlySerMetLysProThrLeuAspValGlyAspPhe    100

ATCCTGGTGAACAAGTTTTCCTACGGGATCCGCTTGCCGGTGATCGACAAGAAAGTCATC    360
IleLeuValAsnLysPheSerTyrGlyIleArgLeuProValIleAspLysLysValIle    120

GAGGTCGGTGACCCGCAGCGCGGCGATGTGATGGTATTCCGCTACCCGAGCGACCCCAAC    420
GluValGlyAspProGlnArgGlyAspValMetValPheArgTyrProSerAspProAsn    140

GTCAACTACATCAAGCGTGTAGTCGGCCTGCCGGGCGACGTGGTGCGCTACACCAGCGAC    480
ValAsnTyrIleLysArgValValGlyLeuProGlyAspValValArgTyrThrSerAsp    160

AAGCGCCTGTTCATCAACGGTGAATCGGTGGCCGAGAAGCTGCTGGGCGCCGAGCCGAAC    540
LysArgLeuPheIleAsnGlyGluSerValAlaGluLysLeuLeuGlyAlaGluProAsn    180

ACCCTGGGCAGCGCCGAGTTGTACCAGGAAAAACTCGGTGCGGTAGAGCACGAGATCCGC    600
ThrLeuGlySerAlaGluLeuTyrGlnGluLysLeuGlyAlaValGluHisGluIleArg    200

AAGGAAATGAGCCGCTACCGCGCAATGCCTGATGGCCAGTGGAAAGTACCCGCCGGGCAC    660
LysGluMetSerArgTyrArgAlaMetProAspGlyGlnTrpLysValProAlaGlyHis    220

TACTTCATGATGGGCGACAACCGCGACAACTCCAATGACAGCCGTTACTGGGATGACCCC    720
TyrPheMetMetGlyAspAsnArgAspAsnSerAsnAspSerArgTyrTrpAspAspPro    240

AATATTCCCAAGGACCTGCTGGGCATGGTGCCCGACGAGAACATCGTCGGCAAAGCCTTC    780
AsnIleProLysAspLeuLeuGlyMetValProAspGluAsnIleValGlyLysAlaPhe    260

GCGGTCTGGATGAGTTGGCCGGAGCCCAAGCTCAGCCACCTGCCGAACTTCTCGCGGGTC    840
AlaValTrpMetSerTrpProGluProLysLeuSerHisLeuProAsnPheSerArgVal    280

GGGCTGATCAAGTAAtacaggcggcgctgtgaacacagcgccgaatgcttttctggggtt    900
GlyLeuIleLysEnd                                                 300
```

Fig. 3. Nucleotide sequence and deduced amino acid sequence of the *lep* gene from *P. fluorescens*

The nucleotide sequence and the deduced amino acid sequence of the *lep* gene and the 3′ end of the *lepA* gene from *P. fluorescens* are shown. The nucleotide sequence of both strands was determined. The potential ribosome binding site (RBS) and transcription terminator sequences are overlined. Nucleotides in upper case indicate the *lep* structural gene.

compare the amino acid sequence with those of other proteins. A search of the GenEmbl, PIR and Swissprot databases utilizing the FASTA programme [17] resulted in detection of weak similarity to pepsinogen C (an aspartyl proteinase), but nothing of any likely consequence with other sequences. The optimal alignment of pepsinogen C with LP1 resulted in superimposition of two areas of potential significance, including aspartate-158 of the LP1 sequence. Although Asp-158 is not followed by either a Thr-Gly or a Ser-Gly motif, the possible identification of LP1 as an aspartyl proteinase could not be excluded. A second approach is to consider that LP1 has arisen as a consequence of convergent evolution and therefore displays no sequence homology but utilizes a mechanism consistent with identification as a serine, cysteine or metallo-proteinase. Exploration of sequence similarities with known proteinases would be unlikely to yield useful information in such a case.

The *E. coli* LP1 primary sequence contains three cysteine (positions 21, 170 and 176) and three histidine (positions 124, 235 and 323) residues [5]. Published work with LP1 has established that His-323 [10] and Cys-21 [18] are inessential as far as catalytic activity is concerned. It is possible that LP1 may be classified by generating mutants which lack residues which would be essential in particular types of mechanism. Would removal of, for example, Cys-170 result in inactivation of the enzyme? Inactivation would not necessarily indicate that LP1 is a cysteine proteinase, but unaltered activity would certainly exclude identification as a cysteine proteinase. Therefore six mutant forms of LP1 were generated. Histidines at positions 124 and 235 were mutated to asparagine (H124N, H235N), cysteines at positions 170 and 176 were mutated to alanine (C170A, C176A) and the aspartate at position 158 was mutated to asparagine (D158N). The double mutant C170A/C176A was also constructed. *E. coli* strain IT41,

which encodes a temperature-sensitive LP1 enzyme, was transformed with the plasmid pNA3 which produces wild-type LP1, with pNA3 variants which express the various mutant forms of LP1, or with pBR322 as a control. Cell growth was monitored at 42 °C (Fig. 1). It is apparent that the catalytic activity of all the mutant forms of LP1 is physiologically sufficient for cell survival in the absence of wild-type LP1. The growth curves of all mutants are presented as an average, as growth of cells harbouring any of the mutant forms of LP1 was indistinguishable from that of cells expressing the wild-type LP1 enzyme. (Note that the control plasmid pBR322 does not confer viability upon the cells at the non-permissive temperature.) These data also indicate that LP1 does not contain essential disulphide bridges.

The validity of these results is, of course, dependent on both the reliability of the *in vivo* assay used to monitor activity and on the premise that substitution of putatively essential amino acids should render the activity of the enzymes so low as to be incompatible with the requirements for cell survival. Recent work [18] has demonstrated that the *in vivo* assay is in fact a more sensitive monitor of enzyme activity than available *in vitro* assays, as mutant forms of LP1 have been generated which have undetectable activity as isolated enzymes but have significant activity *in vivo*. Experiments with proteinase such as subtilisin and trypsin [19,20] have demonstrated that substitution of key residues at the active site, for example the serine or histidine residues of the catalytic triad, results in decreases in $k_{cat.}$ values of six orders of magnitude. It can therefore be concluded that the data presented in Fig. 1 are inconsistent with identification of LP1 as a member of any of the 'standard' four classes of proteinase, which are defined according to the mechanism of action and to the identity of the amino acid side chains involved in peptide bond hydrolysis.

An additional approach to identifying the amino acid residues which are involved in catalysis, in the absence of a database of related primary structures, is to clone and sequence the homologous gene from another species. Ideally, the amino acid sequence of the chosen gene should display a degree of sequence similarity with LP1 such that the proteins are clearly related but have diverged to a significant degree during the evolutionary process. The target selected was *P. fluorescens*, as the Pseudomonas genus represents some of the most distantly related Gram-negative bacteria to *E. coli*. Southern blots of the *P. fluorescens* genomic library contained within λ EMBL3 identified seven plaques which hybridized relatively strongly with the *E. coli lep* gene probe. One of these plaques (λPF5) was chosen and grown, and the DNA was isolated. The cloned DNA was then subjected to restriction analysis and further Southern blotting to identify DNA fragments containing the *P. fluorescens lep* gene. These fragments were cloned into the phasmid vectors pTZ18R and pTZ19R, single-strand DNA was isolated and the nucleotide sequence of the inserts was determined using M13 reverse primer and synthetic oligonucleotides complementary to the cloned inserts. Fig. 2 shows a partial restriction map of the region of DNA which encompasses the *lep* gene, and Fig. 3 shows the nucleotide sequence and the deduced amino acid sequence of the *lep* gene, including the 5′ and 3′ flanking regions, from *P. fluorescens*. Also shown is a partial sequence of the *lepA* homologue. The *lep* gene encodes a protein consisting of 284 amino acid residues with a calculated molecular mass of 31903 Da. Amino acid sequence similarity clearly indicates that the *lepA* gene resides immediately upstream of the *lep* gene, as it does in both *E. coli* and *S. typhimurium*, although the *lepA* structural gene terminates 5 nucleotides before the initiator methionine codon of LP1 in *P. fluorescens*, whereas the *E. coli lep* and *lepA* genes are separated by 15 nucleotides. It therefore appears likely that the *lepA* and *lep* genes are co-transcribed in *P. fluorescens*, as they are in *E. coli*, under the control of a promoter proximal to the *lepA* gene. Almost immediately downstream of the gene is a possible stem–loop structure, followed by a T-rich region which might function as a rho-independent transcriptional terminator. A potential Shine–Dalgarno sequence is present in the intergenic region.

```
Eco MANMFALILVIATLVTGILWCVDKFFFAPKRRERQAAAQAARDSLDKATL 50
Pfl MSLNFPLLLVIAVAVCGLLALLDLVFFAPRRRSAIASYQGSVSQPDAVVI 50

Eco KKVAPKPGWLETGASVFPVLAIVLIVRSFIYEPFQIPSGSMMPTLLIGDF 100
Pfl EKLNKEPLLVEYGKSFFPVLFIVLVLRSFLVEPFQIPSGSMKPTLDVGDF 100

Eco ILVEKFAYGIKDPIYQKTLIETGHPKRGDIVVFKYPEDPKLDYIKRAVGL 150
Pfl ILVNKFSYGIRLPVIDKKVIEVGDPQRGDVMVFRYPSDPNVNYIKRVVGL 150

Eco PGDKVTYDPVSKELTIQPGCSSGQACENALPVTYSNVEPSDFVQTFSRRN 200
Pfl PGDVVRYTS.DKRLFINGES....VAEKLLGAEPNTLGSAELYQ..... 189

Eco GGEATSGFFEVPKNETKENGIRLSERKETLGDVTHRILTVPIAQDQVGMY 250
Pfl .........................EKLGAVEHEI......RKEMSRY 206

Eco YQQPGQQLATWIVPPGQYFMMGDNRDNSADSRYW.........GFVPEA 290
Pfl RAMPDGQ...WKVPAGHYFMMGDNRDNSNDSRYWDDPNIPKDLLGMVPDE 253

Eco NLVGRATAIWMSFDKQEGEWPTGLRLSRIGGIH 323
Pfl NIVGKAFAVWMSW..PEPKLSHLPNFSRVGLIK 284
```

Fig. 4. Comparison of the amino acid sequences of LP1 from *E. coli* and *P. fluorescens*

Optimal amino acid sequence alignment of the *E. coli* and *P. fluorescens* LP1 proteins resulting from application of the Needleman & Wunsch algorithm [17], with a GAP penalty of 3.0. Bars represent amino acid identities, (:) represents conservative substitutions and (.) represents semi-conservative substitutions. The alignment displays 50% sequence identity and 68% homology when identities and conservative substitutions are summed. Eco, *E. coli* LP1; Pfl, *P. fluorescens* LP1.

A comparison of the protein primary structures of *E. coli* LP1 and *P. fluorescens* LP1 (Fig. 4) reveals a sequence identity of 50%. The *N*-terminal ~ 160 residues of the proteins are highly conserved, as is a prominent sequence towards the *C*-terminus. The most striking differences between the two proteins are a large deletion of ~ 50 residues of the *P. fluorescens* sequence between these two regions, which is followed by the highly conserved *C*-terminal motif, after which an insertion of 10 residues in the *P. fluorescens* enzyme occurs; the significance of these differences is currently unclear.

Despite the differences in primary structure, *P. fluorescens* LP1 is able to function effectively in *E. coli*. When *E. coli* strain IT41 was transformed with the plasmid pNA109, which contains a 3.4 kb *Sma*I–*Eco*RI fragment encompassing the *P. fluorescens lep* gene (see Fig. 2) and incubated at 42 °C, it was found that the cells survived and grew (results not shown). Clearly, both the product and the promoter of the *P. fluorescens lep* gene are active in *E. coli*. It is interesting to note that, with the exception of histidine-235 (numbering according to the *E. coli* gene sequence), none of the amino acid residues which were subjected to mutagenesis are conserved in the *P. fluorescens* LP1 sequence. It was considered a remote possibility that transformed IT41 cells might survive at the restrictive temperature due to a genetic recombination event, such as exchange of coding sequence at the 3′ end of the *lep* gene from plasmid DNA to the genome, or alternatively, of DNA encompassing the wild-type codon from the genome to the mutated plasmid. Although this is unlikely since the transformation efficiency (at 42 °C) of different plasmids was uniform, it was considered appropriate to resequence the

H235N mutant (histidine-235 is the only residue which was subjected to site-directed mutagenesis that is conserved in the *P. fluorescens* sequence) after two growth cycles at 42 °C in IT41. The DNA sequence remained unchanged, confirming the validity of the experimental protocol.

In addition, studies with the isolated enzyme have indicated that *E. coli* LP1 is resistant to inhibition by a wide variety (> 30) of commercially available proteinase inhibitors including *o*-phenanthroline, EDTA, phosphoramidon, 1-chloro-4-phenyl-3-L-tosylamidobutan-2-one ('TPCK'), phenylmethanesulphonyl fluoride, elastinal, pepstatin, diazoacetyl-DL-norleucine methyl ester, 1,2-epoxy-3-(*p*-nitrophenoxy)propane, iodoacetamide, L-*trans*-epoxysuccinyl-leucylamido-(4-guanidino)butane and 2,2-dithiodipyridine (J. Bradley & A. Allsop, unpublished work).

Conclusion

The data presented here are entirely consistent with the proposal that prokaryotic LP1 is representative of a novel class of proteinase. There are currently no data available which might indicate the mechanism by which this class of enzyme effects the hydrolysis of peptide bonds. Perhaps the simplest alteration to a defined mechanism would be replacement of the histidine residue which acts as an active-site base in serine proteinases, in which case likely possibilities would be substitution with a tyrosine residue or an acid group such as an aspartate or a glutamate. Although the pK_a value of tyrosine is considerably higher than that of histidine, it is not unknown for tyrosine to act as a base at the active site of an enzyme that shows optimum activity at pH values close to neutral [21]. A recent publication [22] has indicated that the α-carbon backbone of the class C β-lactamase from *Citrobacter freundii* can be superposed upon that of trypsin when the catalytically active serine residues are aligned, and that under these conditions the hydroxyl oxygen atom of a tyrosine residue at position 150 within the β-lactamase superposes (within < 0.05 nm) upon the ϵ-nitrogen of the imidazole ring of the active-site histidine of trypsin. This observation has led to the proposal that tyrosine-150 functions as a general base during catalysis, in the same way as does histidine-57 in trypsin. The class A β-lactamases also utilize a serine nucleophile to effect hydrolysis of the amide bond of the β-lactam ring, except that in this case a glutamate residue acts as a base by abstracting the serine hydroxyl proton via an intervening water molecule [23]. It is tempting to speculate that the mechanism of action of the prokaryotic LP1 involves the generation of a serine nucleophile, and that this nucleophile is generated by abstraction of the hydroxyl proton by a tyrosine residue, as the phenolate anion, or by an acidic amino acid residue either directly or via an intermediate water molecule. Either of these proposals would represent perhaps the smallest change in enzyme structure and mechanism compared with those which are known to exist. In this respect it is interesting to note that there are seven tyrosine, seven aspartate, four glutamate and seven serine residues which have been conserved in all three prokaryotic LP1 sequences within the periplasmically exposed domain. As the scissile bond of the preprotein is cleaved close to the membrane surface, serine residues at positions 88 and 90 in the *E. coli* LP1 sequence may be considered as likely candidates, as they are located in a region of the enzyme almost immediately *C*-terminal to the second *trans*-membrane sequence. Further experiments will be necessary to determine whether any of these residues are involved in catalysis.

Received 10 July 1991; accepted 19 August 1991

Note added in proof (received 12 December 1991)

The amino acid sequence of the yeast mitochondrial Inner Membrane Protease 1 (IMP1) has been published [24] and displays 30% sequence identity with the *E. coli* LP1 enzyme. We note that only three serine residues are conserved between the two sequences, at positions 90, 278 and 281 (*E. coli* numbering), and that these residues all occur within highly conserved regions of the primary structures. It is also noteworthy that the only histidine residue which is conserved between the *E. coli* LP1 and *P. fluorescens* LP1 sequences is not conserved in IMP1; this supports the view that a histidine residue is not required for catalysis.

We thank Dr. Chris Franklin, Dr. Neil Clarke and Dr. John Hodgson for the DNA library and gratefully acknowledge Dr. Yoshikazu Nakamura for the gift of *E. coli* strain IT41 and Dr. Bill Wickner for plasmid pRD8. Thanks are also due to Isabel Bennett, Janine Bradley, David Knowles, Sara McLean and Robert Southgate for helpful discussions.

REFERENCES

1. J. Bioenerg. Biomembr. (1990) **22**
2. YaDeau, J. T., Klein, C. & Blobel, G. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 517–521
3. Baker, R. K. & Lively, M. O. (1987) Biochemistry **26**, 8561–8567
4. Greenburg, G., Shelness, G. S. & Blobel, G. (1989) J. Biol. Chem. **264**, 15762–15765
5. Wolfe, P. B., Wickner, W. & Goodman, J. M. (1983) J. Biol. Chem. **258**, 12073–12080
6. van Dijl, J. M., van den Bergh, R., Reversma, T., Smith, H., Bron, S. & Venema, G. (1990) Mol. Gen. Genet. **223**, 233–240
7. Moore, K. E. & Miura, S. (1987) J. Biol. Chem. **262**, 8806–8813
8. von Heijne, G. (1983) Eur. J. Biochem. **133**, 17–21
9. Dalbey, R. E. & Wickner, W. (1985) J. Biol. Chem. **260**, 15925–15931
10. Inada, T., Court, D. L. Ito, K. & Nakamura, Y. (1989) J. Bacteriol. **171**, 585–587
11. Sanger, F., Nickeln, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
12. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor
13. Fersht, A. (1985) Enzyme Structure and Mechanism, 2nd edn., W. H. Freeman and Co., New York
14. Davies, D. R. (1990) Annu. Rev. Biophys. Biophys. Chem. **19**, 189–215
15. Miller, M., Jaskolski, M., Mohana, R., Leis, J. & Wlodawer, A. (1989) Nature (London) **337**, 576–579
16. Benyon, R. J. & Bond, J. S. (eds.) (1989) Proteolytic Enzymes: A Practical Approach, IRL Press, Oxford
17. Devereux, J., Haeberli, P. & Smithies, O. (1984) Nucleic Acids Res. **12**, 387–395
18. Bilgin, N., Lee, J. I., Zhu, H.-y., Dalbey, R. & von Heijne, G. (1990) EMBO J. **9**, 2717–2722
19. Carter, P. & Wells, J. A. (1988) Nature (London) **332**, 564–568
20. Higaki, J. N., Evnin, L. B. & Craik, C. S. (1989) Biochemistry **28**, 9256–9263
21. Reid, G. A., White, S., Black, M. T., Lederer, F., Matthews, F. S. & Chapman, S. K. (1988) Eur. J. Biochem. **178**, 329–333
22. Oefner, C., D'Arcy, A., Daly, J. J., Gubernator, K., Charnas, R. L., Heinze, I., Hubschwerlen, C. & Winkler, F. K. (1990) Nature (London) **343**, 284–288
23. Lamotte-Brasseur, J., Dive, G., Dideberg, P., Charlier, P., Frere, J.-M. & Ghuysen, J.-M. (1991) Biochem. J. **279**, 213–221
24. Behrens, M., Michaelis, G. & Pratjè, E. (1991) Mol. Gen. Genet. **228**, 167–176