STRUCTURE-FUNCTION STUDIES OF SERINE HYDROLASES: SYNTHESIS OF A GENE FOR α-LYTIC PROTEASE AND THE PURIFICATION AND CHARACTERIZATION OF A MUTANT β-LACTAMASE

Thesis by

Dianne M. Perez

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1988

(Submitted November 23, 1987)

Acknowledgement

I wish to thank my research adviser, John H. Richards, for his guidance and support during my graduate study at Caltech. I also wish to thank the entire Richards' research group for creating an atmosphere that made graduate study an enjoyable experience.

Above all, I wish to thank my husband, Gil, for encouraging me during these five years. Though many times he thought he was hampering my study, the truth is that without him none of this would be possible. Thank you for loving me and believing in me.

NIH training grant GM 07616 provided financial support.

Abstract

The author has constructed a synthetic gene for α-lytic protease. Since the DNA sequence of the protein is not known, the gene was designed by using the reverse translation of α-lytic protease's amino acid sequence . Unique restriction sites are carefully sought in the degenerate DNA sequence to aid in future mutagenesis studies. The unique restriction sites are designed approximately 50 base pairs apart and their appropriate codons used in the DNA sequence. The codons used to construct the DNA sequence of α-lytic protease are preferred codons in E. coli or used in the production of β-lactamase. Codon usage is also distributed evenly to ensure that one particular codon is not heavily used. The gene is essentially constructed from the outside in. The gene is built in a stepwise fashion using plasmids as the vehicles for the α-lytic oligomers. The use of plasmids allows the replication and isolation of large quantities of the intermediates during gene synthesis. The α-lytic DNA is a double-stranded oligomer that has sufficient overhang and sticky ends to anneal correctly in the vector. After six steps of incorporating α-lytic DNA, the gene is completed and sequenced to ensure that the correct DNA sequence is present and that no mutations occurred in the structural gene.

β-lactamase is the other serine hydrolase studied in this thesis. The author used the class A RTEM-1 β-lactamase encoded on the plasmid pBR322 to investigate the roll of the conserved threonine residue at position 71. Cassette mutagenesis was previously used to generate all possible amnio acid substitutions at position 71. The work presented here describes the purification and kinetic characterization of a T71H mutant previously constructed by S.Schultz. The mutated gene was transferred into plasmid pJN for expression and induced with IPTG. The enzyme is purified by column chromatography and FPLC to homogeneity. Kinetic studies reveal that the mutant has lower k_{cat} values on benzylpenicillin, cephalothin and 6-aminopenicillanic acid but no changes in k_m except for cephalothin which is approximately 4 times higher. The mutant did not change siginificantly in its pH profile compared to the wild-type enzyme. Also, the mutant is more sensitive to thermal denaturation as compared to the wild-type enzyme. However, experimental evidence indicates that the probable generation of a positive charge at position 71 thermally stabilized the mutant.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 3

LIST OF TABLES

Nomenclature and Abbreviations

All of the amino acids used are abbreviated according to the standard one amino acid base code. DTT is dithiothreitol. DNA stands for deoxyribonucleic acid. ATP is adenosine triphosphate. T71H stands for the amino acid threonine at position 73 in the protein being replaced by the amino acid histidine at the same position.

CHAPTER 1

INTRODUCTION

Enzymes are proteins that catalyze and regulate most of the molecular interconversions that occur in living organisms. As with all other proteins, their linear sequences of amino acids are encoded in the DNA of the appropriate structural genes. To understand the relationship between the linear sequence of amino acids in a protein and its three-dimensional structure and biological function forms one of the major challenges of present day protein biochemistry. An important approach to gain insights into these relationships has recently been provided by the ability to change, in a precise way, the amino acid sequence in a given protein and then to analyze the way in which this specific change alters the three-dimensional structure and function. This has become possible by virtue of developments in molecular biology that allow site specific mutations to be made in DNA. Since it is not yet possible to predict, with any reliability, the three-dimensional structure of a protein based only on the knowledge of its amino acid sequence, realistic goals in structure-function studies generally involve the alteration of relatively few residues at a time, with the expectation that these relatively conservative alterations will not grossly affect the three-dimensional structure of the protein so that, for example, it no longer folds into an active form. Some realistic goals for possible improvements in

protein function using presently available knowledge include alterations that can enhance the stability of enzymes so that they remain active under conditions sufficiently harsh to denature the wild-type analogue.Longer term goals include the ability to design an enzyme with a particular activity or specificity not encountered naturally (1).

Before the advent of site specific mutagenesis to alter the DNA that encodes an enzyme, chemical procedures had been used to change particular amino acid functional groups within a preformed protein. An early example was the chemical modification in the protease subtilisin of an active site serine residue into cysteine (2). In general, such chemical approaches are fraught with difficulties as the procedures necessary to make the changes generally involve harsh conditions and non-specific reactions that affect the protein in undesired ways. Site specific mutagenesis, however, provides a powerful and clean approach to allow a specific change of a particular amino acid to any one of 19 other possibilities. Site specific mutagenesis of a DNA sequence was first accomplished in the single stranded phage ϕX-174 (3,4). The first changes made specifically in proteins to study structure-function relationships were in β-lactamase (5) and tyrosyl tRNA synthetase (6).

Specific mutations can be introduced into

structural genes incorporated in either single or double stranded vectors. Single stranded vectors have proved more generally useful in this regard. In a common procedure, the gene for the protein of interest is subcloned into a single stranded M13 vector. DNA replication of this vector is primed by an oligonucleotide that binds specifically at the site of the intended mutation, and this primer is synthesized with bases that do not complement the sequence in the template. These mismatched bases define the site and the nature of the resulting mutations. In vitro chain extension from the 3′ end of the primer can be accomplished, using enzymes such as Klenow fragment of DNA polymerase, to produce a double stranded vector, the template strand of which incorporates the wild-type protein, and the newly synthesized strand of which encodes the desired mutation. This heteroduplex is transformed into an appropriate host such as E. coli to produce colonies of mixed genotype due to semiconservative replication , one daughter vector encodes the wild-type protein while the other encodes the mutant. A plasmid preparation and a second transformation produce pure genotypic colonies which can be screened for the mutation by hybridization with the radioactive primer used earlier to generate the mutant (7).

Recently, a new powerful method has developed in site specific mutagenesis called site saturation mutagenesis (Figure 1) (8,9).In a typical procedure, the DNA encoding the amino acid of interest is enclosed between two unique restriction sites. These unique sites either exist naturally or can be generated if needed using standard mutagenic techniques. The unique restriction sites are desired near the mutagenic site to allow a synthetic oligonucleotide to be inserted with ease. The DNA of interest between the two restriction sites is now removed and oligomers containing a mixture of oligonucleotides encoding all 20 amino acids is ligated back into the gene. The 20 different plasmids now encoding the wild-type protein and 19 mutations is now transformed into E. coli and the colonies screened phenotypically to select active mutants. Cassette mutagenesis is very powerful because it allows the rapid study of many mutants. Furthermore, only mutants are selected through the phenotypic screening that display significant activity or an interesting function. Therefore, site saturation mutagenesis enables a thorough study of the role an important amino acid residue plays in the structure-function of a protein by generating all possible amino acid substitutions at the desired site.

During recent years, many interesting mutants have been generated that allow insight into the

structure-function relationships in proteins. An example is the study of the specificity pocket of trypsin (10). Both of the trypsin mutants generated, G216A and G226A, have lower k_{cat} values than the wild-type enzyme. However, both mutants also significantly changed in their preference for substrates in comparison to the wild-type enzyme.

The generation and study of mutant enzymes with amino acid changes that influence or involve the catalytic mechanism are producing interesting conclusions. For example, changing asparagine 155 in the serine protease subtilisin reduces k_{cat} by 200 to 4000 fold depending upon the amino acid substitution (11).Asparagine 155 stabilizes the tetrahendral intermediate in the "oxyanion hole" of subtilisin. The value of k_m does not significantly change in any of these asparagine mutants . In the serine hydrolase β-lactamase, the active site serine 70 residue was changed to threonine. The mutant has no catalytic activity (12). Furthermore, the active site lysine 73 in β-lactamase was replaced with all other 19 amino acids by cassette mutagenesis. None of the lysine mutants are phenotypically resistant to penicillin. However, kinetic analysis of the K73R mutant does show residual activity (0.01%) while the K_M remained unchanged. In addition, the K73C mutant was also kinetically studied and the pH-activity profile is drastically altered. The K73C

mutant has maximum activity at pH 8.3 while the wild-type enzyme has a maximum activity at pH 6.5. (13).

The generation and study of mutants in the regions of protein not involved directly with the three-dimensional structure or catalysis are providing insights into the pH dependence of an enzyme. Previous chemical modification studies indicated that the overall surface charge of serine proteases is related to the pH dependence of catalysis (14). Therefore, site specific mutagenesis of subtilisin would provide a precise method for the study of the protein's surface residues. A site specific mutant of subtilisin was generated in which the Asp 99 surface residue was changed to Ser 99 (15). This mutation should not affect the structure or catalytic properties of the enzyme . The D99S mutant did shift the pK_a of the active site from 7.17 to 6.88.

Serine hydrolases, such as subtilisin and trypsin, are the objects for mutations because of their known three-dimensional structure and catalytic function. The family of serine hydrolases have in common an active site serine residue which is essential for activity. The mode of catalysis is a nucleophilic attack of the hydroxyl group of the serine. Members of the family include : acetylcholine esterases, diisopropyl phosphorofluoridate hydrolyzing enzymes, serine proteases and β-lactamases of which the last two are studied extensively in this thesis.

Serine proteases are found extensively in organisms, being utilized in blood clotting, digestion, cell lysis and defense mechanisms. Their functions have great diversification yet there are analogies both in amino acid sequence and three-dimensional structure. The digestive enzymes which include trypsin, chymotrypsin and elastase are well studied both in structure and function .An interesting serine protease is α-lytic protease, which although a microbial serine protease, is homologous to these mammalian digestive enzymes.

The author has constructed a synthetic gene for α-lytic protease. Since the DNA sequence of the protein is not known, the gene was designed by using the reverse translation of α-lytic protease's amino acid sequence . Unique restriction sites are carefully sought in the degenerate DNA sequence to aid in future mutagenesis studies. The unique restriction sites are designed approximately 50 base pairs apart and their appropriate codons used in the DNA sequence. The codons used to construct the DNA sequence of α-lytic protease are preferred codons in E. coli or used in the production of β-lactamase. Codon usage is also distributed evenly to ensure that one particular codon is not heavily used. The gene is essentially constructed from the outside in. The gene is built in a stepwise fashion using plasmids as the vehicles for the α-lytic oligomers. The use of plasmids allowsthe replication and isolation of large

quantities of the intermediates during gene synthesis. The α-lytic DNA is a double-stranded oligomer that has sufficient overhang and sticky ends to anneal correctly in the vector. After six steps of incorporating α-lytic DNA, the gene is completed and sequenced to ensure that the correct DNA sequence is present and that no mutations occurred in the structural gene. After the gene is expressed, it will be utilized for structure-function studies with the use of site-specific or cassette mutagenesis.

β-lactamase is the other serine hydrolase studied in this thesis. The author used the class A RTEM-1 β-lactamase encoded on the plasmid pBR322 to investigate the roll of the conserved threonine residue at position 71. Cassette mutagenesis was previously used to generate all possible amnio acid substitutions at position 71. The work presented here describes the purification and kinetic characterization of a T71H mutant previously constructed by S.Schultz (8). The mutated gene was transferred into plasmid pJN for expression and induced with IPTG. The enzyme is purified by column chromatography and FPLC to homogeneity. Kinetic studies reveal that the mutant has lower k_{cat} values on benzylpenicillin, cephalothin and 6-aminopenicillanic acid but no changes in k_m except for cephalothin which is approximately 4 times higher. The mutant did not change siginificantly in its pH profile compared to the

wild-type enzyme. Also, the mutant is more sensitive to thermal denaturation as compared to the wild-type enzyme. However, experimental evidence indicates that the probable generation of a positive charge at position 71 thermally stabilized the mutant.

REFERENCES

(1) Maugh, T.H. (1984) Science, 223, 269-271

(2) Neet, K. E. and Koshland, Jr.,D.E. (1966) Proc. Natl. Acad. Sci. USA, 56, 1606

(3) Razin, A, Hirose, T.,Itakura, K., and Riggs, A.D. (1978) Proc. Natl. Acad. Sci. USA,75, 4268

(4) Shortle, D., and Nathans, D. (1978) Proc. Natl..Acad. Sci. USA,75, 2170

D.-McFarland,G.,Cohen,L.W.,Riggs,A.D.,Morin,-C.,Itakura,K.and Richards,J.H.(1982) Proc. Natl. Acad. Sci. USA, 79, 6409-6413

(6) Wilkinson,A.J.,Fersht,A.R.,Blow,D.M.and Winter,G.(1983) Biochemistry,22, 3581-3586

(7) Zoller, M.J. and Smith, M. (1983) Methods Enzymol.,100,468-500

(8) Schultz, S.C. and Richards, J.H.(1986) Proc. Natl. Acad. Sci. USA,83, 1588-1592

(9) Wells, J.A., Vasser,M.,and Powers , D.B.,(1985) Gene, 34,315-323

(10) Craik, C.S.,Largman,C.,Fletcher,T.,Rocznіak,S., Barr, P.,Flettenick,R. and Rutter, W.(1985) Science ,228,291-297

(11) Rochester, D.E., Winer, J.A., and Shah, D.M.(1986) Embo. J.,5, 541

(12) D.-McFarland,G.,Neitzel, J.J. and Richards ,J.H. (1986) Biochemistry,25, 332-338

(13) Carroll, S.S. (1987) Ph.D.thesis, California

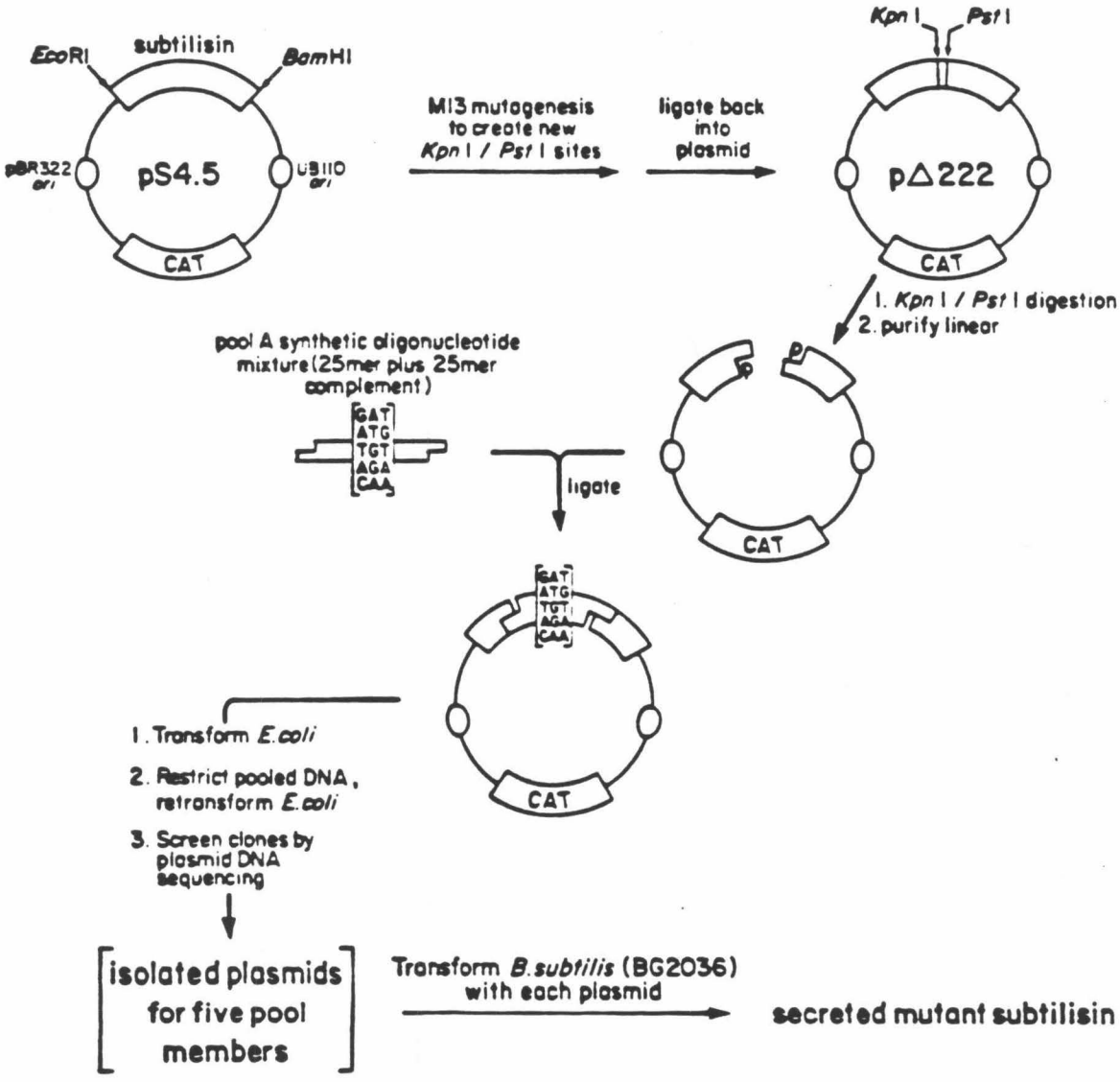
Institute of Technology.

(14) Valenzuela, P. and Bender, M.L.(1971) Biochim. Biophys. Acta., 250, 538-548

(15) Thomas, P.G., Russell, A.J. and Fersht, A.R.(1985) Nature, 318, 375-376.

FIGURE 1

The site saturation mutagenesis method (9).


EcoRI
subtilisin
BamHI
pBR322 ori
pS4.5
UB110 ori
CAT
M13 mutagenesis to create new Kpn I / Pst I sites
ligate back into plasmid
Kpn I
Pst I
pΔ222
CAT
1. Kpn I / Pst I digestion
2. purify linear
pool A synthetic oligonucleotide mixture (25mer plus 25mer complement)
GAT
ATG
TGT
AGA
CAA
ligate
CAT
GAT
ATG
TGT
AGA
CAA
CAT
1. Transform E.coli
2. Restrict pooled DNA, retransform E.coli
3. Screen clones by plasmid DNA sequencing
isolated plasmids for five pool members
Transform B.subtilis (BG2036) with each plasmid
secreted mutant subtilisin

CHAPTER 2

Synthesis of a Gene for α-Lytic Protease

Site specific mutagenesis can provide a wealth of information on the structure-function relationships of a protein. Site specific mutagenesis first involves the isolation, cloning and expression of the gene encoding the protein of interest. To obtain the gene, three procedures are known:(i) synthesis of the cDNA from the messenger RNA using reverse transcriptase (ii) isolation of the gene from nature using restriction enzyme digestions ,library construction and screening (iii) synthesis of the DNA by chemical methods.

The chemistry for the synthesis of oligonucleotides began to make strides when the structures of the nucleic acids were first determined (1,2). However, pioneering work had a severe limitation on the size of the polydeoxyribonucleotides that could be synthesized (3). The first chemical methods employed the phosphodiester approach (4), the phosphotriester approach (5,6) and then the phosphite triester method (7). The development of the phosphite triester method allowed research using synthetic genes since the chemistry was reliable and the process relatively easy. Finally, Caruthers developed the deoxynucleotide phosphoramidite chemistry on a solid support (8,9,10) and virtually all synthetic DNA is now made in this fashion. Figure 1 outlines the phosphoramidite chemistry based upon Caruther's work and the synthesis that is used in Caltech's Applied Biosystems DNA synthesizer (11).

Today, now that the chemistry has been developed to a point of high yields, reliability and automation, the field of gene synthesis is ever expanding and no limitations exist currently on the chemical aspects of the designing of a gene of interest. One of the first goals of gene synthesis was to synthesize the yeast alanine transfer RNA (tRNA) by ligating small oligomers together. This was accomplished by 1970 (12) even though the chemical synthesis was early in development. The oligomers were 10 to 12 bases long corresponding to the two strands of DNA. Twenty-six oligomers were annealed together with 5 base overlaps and then covalently linked using DNA ligase. The 26 oligomers were first subgrouped into 4 duplexes and then these 4 duplexes ligated to form the structural gene. Khorana (13) designed the gene from the natural sequence which was known at that time except for changing one codon to insert an EcoR1 restriction site at the end of the gene to facilitate easy removal of the gene.

During the 1970's, the genes for the medically useful human insulin (14) and somatostatin (15) were also synthesized. These two examples demonstrate that a gene could be designed from the amino acid sequence and be could expressed in E. coli despite being human peptide hormones. The oligomers used in building these genes were also small, in the 10-15 base range and were then annealed and ligated together. The genes were

expressed by being fused to the E. coli β-galactosidase gene on plasmid pBR322.

In the early 1980's, synthesis of genes had developed to a point where the time required to synthesize a gene could be similar to that required for isolating the gene from natural sources. This was due mostly to the solid phase approach which increased the rate of synthesis 10-fold over the solution phase. These improvements allowed the synthesis of longer genes. Human leukocyte interferon gene which was 514 bases long is typical of those synthesized during this period (16).

By the mid 1980's, gene were synthesized with longer olgonucleotides due to the higher efficiency of the reactions involved. Typical of the genes produced were human complement C5a, which is 253 base long and ligated together with 16 oligonucleotides approximately 32 bases long (17). Bovine rhodopsin which is 1057 bases in length is the longest gene constructed to date. A total of 72 oligomers 15-40 bases long were annealed to form 3 synthetic fragments which were then cloned and assembled (18). Fifty-two changes were made in the synthetic rhodopsin as compared to the natural gene. Twenty-six unique restriction sites were created though no attempt was made to optimize codon usage in E. coli.

The design, synthesis and study of synthetic genes offers many advantages over the cloning the natural gene. First, the synthetic gene can be designed with

unique restriction sites that can facilitate mutagenesis. In the future, even whole protein domains could be removed or replaced and the effects of these changes studied. Secondly, the synthetic gene is a tool to study some interesting molecular biological problems. For example, one can study the use of a particular codon twice in a row for the same amino acid and its effects on expression. Moreover, one can study the effects of complementary sequences on expression. Thirdly, many genes are only encoded as a single copy in an immense population of DNA and could be very hard to isolate and clone. Moreover, the gene can be degraded easily or be lethal to the cell. Synthetic genes can be made rapidly on the same time scale as trying to isolate the gene from natural sources and the DNA sequence can be altered in the synthetic gene to make it inactive so that no degradation or lethality occurs.

This chapter describes the construction of a synthetic gene for α-lytic protease. Since the DNA sequence of the protein is not known, the gene was designed using the reverse translation of α-lytic protease's amino acid sequence. Unique restriction sites are carefully sought in the degenerate DNA sequence to aid in future mutagenesis studies. The unique restriction sites are designed approximately 50 bases apart and their appropriate codons used in the DNA sequence. The codons used to construct the DNA sequence of α-lytic

protease are preferred codons in E. coli or used in the production of β-lactamase. Codon usage is also distributed evenly to ensure that one particular codon is not heavily used. The gene is constructed in a stepwise fashion using plasmids as vehicles for the synthetic oligomers. After six steps of incorporating DNA encoding regions of α-lytic protease, the synthetic gene was complete and then sequenced to ensure that no mutations occurred in the structural gene.

Materials and Methods

Reagents and Enzymes

T_4 DNA ligase and T_4 polynucleotide kinase were purchased from Boehringer Mannheim Biochemicals (BMB). Acrylamide and bis-acrylamide were electrophoresis grade and also were from BMB. NACS Prepac columns were obtained from Bethesda Research Laboratory (BRL). Sephadex G-25 was from Sigma. All restriction enzymes were purchased from BMB except for AsuII which was obtained from Promega Biotech. pBR322 was purchased from BRL as was electrophoresis grade agarose. [γ-^{32}P]-dATP,-dGTP and -dTTP along with [γ-^{32}P]-ATP (>5000 Ci/mmol) were purchased from Amersham. Potassium tetrachloropalladate (II) was a generous gift from Brent Iverson but can be purchased from Aldrich. Ampillicin was obtained from Sigma. All other chemicals used were of reagent grade or better.

DNA

All oligonucleotides were synthesized on the Applied Biosystems DNA synthesizer using the phosphoramidite method (10,11). The oligomers were purified on a preparative 15% polyacrylamide gel. After elution of the DNA from the acrylamide gel,the oligomers were further purified by ion-exchange chromatography using the NACS Prepac columns. The oligomers were then

desalted on a G-25 Sephadex spin column. DNA quantitation was done according to Maniatis (19) by measuring the optical density of the DNA at 260nm and 280nm.

Plasmid DNA was isolated according to the alkaline lysis method (20). DNA fragments and restriction digests were electrophoresed on a 1.2% agarose gels. DNA was isolated from preparative agarose gels by removing the desired band and using an electroluter from IBI.

Bacterial Strains

E. coli strain LS1 (21) (a derivative of HB101) was used as a host for vectors of modified pBR322.

Computer Program

The computer program used to perform reverse translations, restriction enzyme site searches and other manipulations was the DNA MASTER program written by Tim Hunkapillar (22).

DNA Sequencing

DNA sequencing is a modified version of the Maxam and Gilbert method (23). The modification was done according to the method of Brent Iverson (24). Essentially, the modification includes performing only the "G" and "A" reactions on a double stranded DNA fragment. The "A" reaction uses potassium

tetrachloropalladate(II) in 0.1M acid to modify adenine residues.

Kination and Annealing of Oligomers

After the oligomers are purified as previously described, 100pmoles of oligomer is added to 2.5ul 10x kinase buffer (0.5M Tris,pH 7.6, 0.1M $MgCl_2$, 1mM EDTA), 2ul of 0.1M DTT, 2ul of 10mM ATP, 2ul of T_4 polynucleotide kinase and water up to 25ul. The mixture is incubated at 37 °C for 30 minutes, then incubated at 65 °C for 10 minutes to kill the enzyme. 5ul of each kinased oligomer (20 pmoles) that is to be annealed together is added to 5ul of 10x ligase buffer (0.5M Tris, pH 7.4, 0.1M $MgCl_2$) and water is added to 50 ul. The eppendorf tube containing the mixture is heated to boiling in a water filled beaker 3 minutes, then slow cooled for approximately 1 hour to allow annealing to proceed.

Preparation of Vector

pBR322 (approximately 10 ug) is added to 5 ul of EcoR1 (50 units) and 3 ul of Aval(24 units) with 10 ul of 10x EcoR1 buffer (1M Tris, pH 7.5, 0.5M NaCl, 0.1M $MgCl_2$) and 71ul of water. The sample is incubated at 37 °C for 2 hours, then incubated at 65 C for 10 minutes to kill the enzymes. The digested sample is then electrophoresed on a 1.2% agarose gel and stained with

ethidium bromide. Two bands are evident, one at 2937bp and the other at 1424bp. The band at 2937 bp is isolated and used as the modified pBR322 vector for the synthesis of the gene.

RESULTS

Design of the Synthetic Gene

The amino acid sequence for α-lytic protease from Lysobacter enzymogenes (formerly Myxobacter 495) was determined in 1970 by Olson et al (26). The reverse translation of the amino acids with their degeneracy in codon usage is shown in Figure 2. The letter N stands for any of the 4 bases, P for the purines adenine and guanine, Y for the pyrimidines cytosine and thymine, R for cytosine and guanine and Z for adenine and cytosine. The gene will be 594 bases long.

The next step in the design of the gene was to determine unique restriction sites in the degenerate DNA sequence. Unique restriction sites should cleave the gene only once and also not be found in the vector. The restriction sites were placed about 50 bases apart to ensure the ease of making synthetic oligomers for mutagenesis. Once the restriction sites are determined, only those codons that are preferred in E. coli or are used in the production of β-lactamase are chosen. Table 1 shows the preferred codons in E. coli. Most of the codons that were used in the synthetic gene are E. coli preferred except in a few cases where a restriction site dictated a particular unpreferred codon. However, in all cases, the unpreferred codon was used frequently in β-lactamase production. Table 2 lists the codon usage and

Table 3 the codon analysis in the synthetic gene that was finally used. If one were given a choice between two or more preferred codons to use, codon usage was distributed evenly to ensure that one particular codon is not heavily used.

The design of the gene is now complete with the DNA sequence and restriction sites as shown in Figure 3. In addition to designing the structural gene, a start codon and the EcoR1 restriction site were placed at the beginning of the gene. At the end of the gene, a stop codon and a BglII site were designed. Both the EcoR1 and BglII sites were designed in the synthetic gene for easy placement and removal of the structural gene from the vector. Moreover, the design of the BglII site with the sequence AGATCT destroys the AvaI site (CPyCGPuG) upon ligation of the oligomer into the vector. In the structural gene, an important design was changing the active site serine to alanine at amino acid position 143 as shown in Figure 3. This was designed so that if any α-lytic protease was being expressed after completion of the gene, it would be inactive and not destroy the cell. Also, the inactive enzyme would be useful until a good expression system was found for the gene. An inactive enzyme also prevents selection pressures that the cell uses to mutate the gene.

Construction of the Synthetic Gene

Each oligomer that was synthesized is shown in Figure 4. They are labeled A through O with the A'and O' oligomers being the complementary sequences. The oligomers range in length from 38 to 69 bases long. The oligomers were designed upon annealing to have at least an eleven base overlap. The oligomers were also designed in length so that when the oligomers were inserted into the vector at each step, the total addition of base pairs per step was 100 bases.

Steps 1 through 3

The plasmid chosen as the vector in construction of the synthetic gene is a modified pBR322 plasmid. The modification involves the deletion of nucleotides 6 through 1424 in the wild-type vector. This essentially removes all of the restriction sites between EcoR1 and AvaI which are contained in the gene for tetracycline resistance. Therefore, insertion of α-lytic oligomers result in a plasmid that is only ampicillin resistant and not tetracycline resistant. The removal of the nucleotides and , thus, restriction sites in pBR322 enables several of these restriction sites to be utilized in the design of the α-lytic protease gene. The modified vector will be referred to as pBR32.

A schematic of the construction of the synthetic gene is shown in Figures 5a,b and c. The procedure in

constructing the gene essentially involves building the gene from the outside in. To begin, oligomers O,O' and A,A', representing the beginning and end of the structural gene, are kinased and annealed together to form a double stranded fragment.This double stranded fragment now contains ends compatible for ligation into EcoR1 and Ava1 restriction sites. This double stranded fragment is then ligated into the EcoR1 and Ava1 sites of the modified pBR322 vector resulting in the formation of plasmid pBR32S1 (S1 for step 1). Correct ligation of the OO'and AA' oligomers into pBR32 is determined by the presence of unique restriction sites Mlu1, AsuII and BglII. A simple plasmid preparation, restriction digest and analysis on an agarose gel will reveal if the restriction sites are present.

Step 1 is an important step in the construction of the synthetic gene because inserted here are the oligomers representing the beginning and the end of the structural gene. After step 1 was completed, it was discovered that a mutation occurred at the BglII site resulting in the destruction of the restriction site. However, the BglII site is located after the stop codon outside the structural gene and is not used in the construction of the synthetic gene. Therefore, it was decided to continue building the gene as planned and to correct the mutation after the gene is completed.

The insertion of the oligomers for step 2 was done

in an analogous manner as step 1 (Figure 5a). However, for step 2, the plasmid pBR32S1 is digested with restriction enzymes Mlu1 and AsuII to reopen the vector. The oligomers B,B' and C,C' are annealed forming a double stranded fragment with ends compatible to correctly ligate into Mlu1 and AsuII restriction sites. The double stranded fragment is then ligated into the Mlu1 and AsuII restriction sites of plasmid PBR32S1 forming plasmid pBR32S2. The 8 bases inserted between restriction sites Mlu1 and AsuII in the plasmid pBR32S1 are designed to facilitate the binding and cutting of the restriction enzymes. The 8 bases are removed upon cleavage by Mlu1 and AsuII. Evidence indicates that a certain number of nucleotides are needed surrounding the restriction site to enable the restriction endonuclease to cleave properly (28). In fact, it was observed that some restriction enzymes associated with the 8 base region did not cut as well as restriction sites located in the middle of the structural gene. Perhaps this is an indication of a sequence requirement needed by the endonuclease for the bases adjacent to the cleavage site.

The insertion of the oligomers for step 3 is done in precisely the same way as step 2 except different restriction enzymes are used to reopen the vector (Figure 5a). Opening, inserting oligomers and closing the vector would ideally be done six times to

incorporate the full structural gene. However, ligation of oligomers in step 4 failed after several repeated attempts. It was then decided that steps 4, 5 and 6 should be constructed in another vector, thereby, bypassing the use of restriction enzymes SacII and BssHII which gave sporadic and incomplete cleavage patterns in step 4.

Steps 4 through 6

The schematic for constructing steps 4 through 6 is shown in Figure 5b. The vector used in constructing steps 4,5 and 6 of the synthetic gene is the same modified pBR322 vector used in step 1. Synthetic linkers were made to incorporate the BssHII and SacII restriction sites of the oligomers for step 4 into the EcoR1 and AvaI restriction sites of the modified vector. The synthetic linkers were short oligomers ranging from 9 to 15 base pairs. Incorporation of the step 4 oligomers DD'EE' and LL' into the modified vector using the synthetic linkers was sucessful. Steps 5 and 6 were ligated into the vector in an analogous manner to steps 1 and 2 with no further problems.

Incorporating Steps 4 through 6 to Steps 1 through 3

A schematic describing the incorporation of the two halves of the synthetic gene is shown in Figure 5c. At this point in the construction of the synthetic gene,

one-half (steps 1 through 3) of the gene is contained in plasmid pBR32S3 and the other half (steps 4 through 6) is contained in plasmid pBR32S6. Attempts at cleaving plasmid pBR32S6 with restriction enzymes SacII and BssHII were unsuccessful, thus, a recurrence of the same difficulties encountered in step 4. However, since the difficulty is known to involve the cleavage of the enzymes SacII and BssHII, linkers were again constructed connecting the restriction sites BssHII to HindIII and connecting SacII to Sph1. Therefore, plasmid pBR32S6 containing steps 4 through 6 is cut with restriction enzymes HindIII and Sph1 and the gene fragment isolated and ligated using the synthetic linkers into plasmid pBR32S3 (steps 1 through 3) previously reopened with restriction enzymes SacII and BssHII.This technique proved successful and the entire structural gene for - lytic protease is now complete. All of the unique restriction sites in the completed gene were tested to ensure their presence and complete cleavage patterns (Figure 6). The entire gene was sequenced (Figures 7,8,9,10,11) and the DNA sequence of the structural gene was found intact with no mutations occurring. However, the mutation that occurred in the BglII site during step 1 of the construction is still present (Figure 7). The BglII site is important in the synthetic gene design to facilitate the removal of the gene from the vector for expression or for later mutagenesis studies. Therefore,

the mutation at the BglII site needs to be corrected.

Correction of the BglII Restriction Site

Correction of the mutation at the BglII site is a fairly simple procedure using the three fragment ligation strategy of S. Schultz (27) (Figure 12). Plasmid pBR32LP containing the completed gene is digested with restriction enzymes EcoR1 and AsuII and the 550 bp fragment containing most of the structural gene is isolated on a agarose gel. A new synthetic oligomer containing the DNA sequence from the AsuII site of the α-lytic protease gene to the Aval site is made. This synthetic oligomer encodes the correct BglII restriction site. The 550bp fragment of the synthetic gene is ligated with the new synthetic oligomer into the same modified pBR322 vector previously used in step 1. After transformation, colonies were screened by restriction digests for the corrected BglII site. Figure 13 shows the restriction map of the corrected α-lytic protease gene and its successful digestion with BglII.

DISCUSSION

The stepwise addition of oligomers in the construction of the synthetic gene for ∝-lytic protease proves to be a highly useful and flexible method for the general construction of synthetic genes. The stepwise addition of oligomers into a plasmid vehicle allows the amplification and isolation of the plasmid to provide large quantities of the intermediates during gene construction. Therefore, one can determine at each step of incorporation that the oligomer ligated correctly. This method of construction is highly useful for graduate students or other relatively inexperienced molecular biologists to monitor the growth of the synthetic gene. The stepwise addition of oligomers also allows the successful construction of very long genes. Previous work in the construction of synthetic genes implemented a method using a large scale ligation of the oligomers. In a general procedure, 20 to 30 oligomers of short length are ligated together in one step. If one did not isolate the correct gene from this ligation mixture, the reason for the failure could not be pinpointed to a particular oligomer or region of the synthetic gene. Therefore, one could not correct or circumvent the area of difficulty. The construction of the ∝-lytic protease gene described here is an excellent

example of the flexibility of this stepwise procedure to circumvent areas of difficulty. It was discovered that the restriction enzymes SacII and BssHII used for the incorporation of step 4 oligomers did not cleave fully. Since the problem was traced to these two restriction enzymes, the use of synthetic linkers to bypass those restriction sites circumvented the problem and gene construction was able to be continued.

An interesting observation in the construction of the synthetic gene is the dependence of the restriction enzymes SacII and BssHII on the nature of the bases adjacent to the cleavage site. The effect of the adjacent sequences on the rate of restriction endonuclease cleavage has been reported for several enzymes (28). Substituting nucleotide analogues at the adjacent sequences of restriction sites indicated that adjacent sequences do not influence directly the binding or cleavage of the enzyme but causes changes in the DNA conformation near the restriction site (29). This change in DNA conformation then sterically inhibits enzyme binding. This effect of the sequences adjacent to the restriction site, however, varied from enzyme to enzyme. SacII and BssHII proved the most troublesome, however, HindIII and BamHI were also noted to only cleave fully when flanked by the DNA sequence of the α-lytic protease structural gene. Cleavage by HindIII and BamHI was partially inhibited when the enzymes were adjacent to

the 8 base region which was designed between restriction sites to facilitate cleavage. Therefore, the use of a particular restriction enzyme in the design of a synthetic gene should be carefully considered. Some enzymes, such as AsuII are difficult to purchase commercially and very expensive. Other enzymes such as XmaI, HpaI and NarI are very sensitive to cleavage conditions and care must be taken to ensure the correct salt concentration and the purity of the plasmid preparation.

Spontaneous mutation in a DNA sequence is a generally rare event, the probablity of occurrencebeing about 1 in 10^{10} for E. coli (30). Since the synthetic gene for ∝-lytic protease contained an alanine in place of the active site serine, there also were no selection pressures to mutate the synthetic gene. However, at labile sites such as a ligation site, the chance of a mutation is much higher. This is what occurred at the BglII site of the synthetic gene for ∝-lytic protease. The mutation that occurred consisted of two transversions and one deletion. However, this mutation was easily corrected after obtaining the oligonucleotides. Use of the three fragment ligation method proved the ease at which mutagenesis can be accomplished in a synthetic gene.

SUMMARY

The author has designed and constructed a synthetic gene for α-lytic protease. The gene was designed by using the reverse translation of the amino acid sequence of α-lytic protease. Unique restriction sites were designed in the sequence approximately 50 bases apart. The gene was constructed in a stepwise fashion using plasmids as a vehicle for the synthetic oligomers. The use of plasmids in the construction allows the amplification and isolation of large quantities of each intermediate in the synthesis. This method of gene synthesis is highly useful for inexperienced students and very flexible in circumventing areas of difficulty.

REFERENCES

1. D.M.Brown and A.R.Todd (1955) in "Nucleic Acids", E.Chargaff and J.N.Davidson, Eds. (Academic Press, New York) Vol. I.
2. J.D.Watson and F.H.C.Crick (1953) **Nature** ,**171**, 737.
3. H.G.Khorana (1968) **Pure Appl. Chem.**,**17**, 349.
4. P.T.Gilham and H.G.Khorana (1958) **J. Amer. Chem. Soc.**, **80**, 6212.
5. R.L.Letsinger and V.Mahaderan (1965) **J. Amer. Chem. Soc.**, **87**, 3526.
6. J.C.Catlin and F.Cramer (1973) **J. Org. Chem.** , **38**, 245.
7. R.L.Letsinger and W.B.Lunsford (1976) **J. Amer. Chem. Soc.**, **98**, 3655.
8. M.D.Matteucci and M.H.Caruthers (1981) **J. Amer. Chem. Soc.**, **103**, 3185.
9. S.L.Beaucage and M.H.Caruthers (1981) **Tetrahedron Lett.**, **22**, 1859.
10. L.J.McBride and M.H.Caruthers (1983) **Tetrahedron Lett.** ,**24**, 245.
11. S.J.Horvath, J.Firca, T.Hunkapillar ,M. Hunkapillar and L.E.Hood (1987) "Recombinant DNA", Part E, Vol 154 of "Methods in Enzymology "(Academic Press,

Inc).

12. H.G.Khorana, K.L.Agarwal, H.Buchi, M.H.Caruthers, N.K.Gupta, K.Kleppe, A.Kumar, E.Ohtsuka, U.L.RajBhandary, J.H.Van de Sande, V. Sgaramella, T.Terao, H.Weber and T.Yamada (1972) J. Mol. Biol., 72, 209.

13. H.G.Khorana (1979) Science, 203, 614.

14. R.Crea, A.Kraszeuski, T.Hirose and K.Itakura (1978) Proc. Natl. Acad. Sci. USA, 75, 5765.

15. K.Itakura, T.Hirose, R.Crea, A.D.Riggs, H.L.Heyneker, F.Bolivar and H.W.Boyer (1977) Science ,198, 1056.

16. M.D.Edge, A.R.Greene, G.R.Heathcliffe, P.A.Meacock, W.Schuck, D.B.Scanlon, T.C.Atkinson, C.R.Newton and A.F.Markham (1981) Nature , 292, 756.

17. W.Mandicki, K.W.Mollison, T.J.Bolling, B.S.Powell, G.W.Carter and J.L.Fox (1985) Proc. Natl. Acad. Sci. USA, 82, 3543.

18. S.S-Karnik, H.G.Khorana, M.Nassal and D.D.Oprean (1986) Proc. Natl. Acad. Sci. USA , 83, 599.

19. T.Maniatis,E.F.Fritsch and J.Sambrook (1982) "Molecular Cloning, a Laboratory Manual " (Cold Spring Harbor Laborartory).

20. D.Ish-Horawicz and J.F.Burke (1981) Nucleic Acids Res. , 9, 2989.

21. H.W.Boyer and D.Roulland-Dussoix (1969) J. Mol. Biol., 41, 459.

22. T.Hunkapillar, personal communication

23. A.M.Maxam and W.Gilbert (1980) Meth. Enzymol. , 65, 499.

24. B.Iverson (1987) Ph.D. Thesis, California Institute of Technology.

25. D.Hanahan (1983) J. Mol. Biol., 166, 557.

26. M.O.J.Olson, N.Nagabhushan, M.Dziviniel, L.B.Smillie and D.R. Whitaker (1970) Nature , 228, 438.

27. S.Schultz (1986) Ph.D.Thesis, California Institute of Technology.

28. M.Thomas and R.W.Davis (1975) J. Mol. Biol. ,91, 315.

29. J.W.Bodnar, W.Zempsky, D.Warden, C.Bergson and D.C. Ward (1983) J. Biol. Chem. , 258, 15206.

30. L. Stryer (1981) "Biochemistry" 2nd. ed.,W.H.Freeman and Company, San Francisco, p.635.

FIGURE 1

Synthesis of DNA using the phosphoramidite chemistry (11).

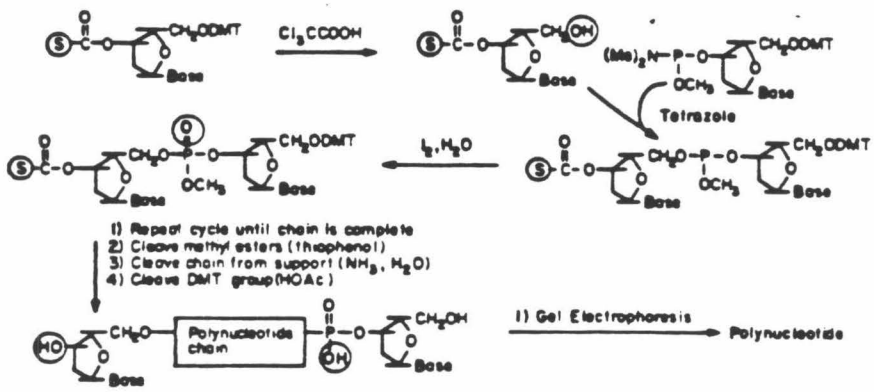

Cl3CCOOH
Tetrazole
I2, H2O
1) Repeat cycle until chain is complete
2) Cleave methyl esters (thiophenol)
3) Cleave chain from support (NH3, H2O)
4) Cleave DMT group(HOAc)
Polynucleotide chain
1) Gel Electrophoresis
Polynucleotide

FIGURE 2

Reverse translation of α-lytic protease. The letter N stands for any of the 4 bases, P for the purines adenine and guanine, Y for the pyrimidines cytosine and thymine, R for cytosine and guanine and Z for adenine and cytosine.

AMBIGUOUS BASE CODE: N=A,T,C,G; P=A,G; Y=C,T; R=A,T; S=C,G; Z=A,T,C

```
1     A N I V G G I E Y S I N N A S L C S V G     20
1     GCNAAYATZGTNGGNGGNATZGAPTAYRSNATZAAYAAYGCNRSNYTNTGYRSNGTNGGN     60

21    F S V T R G A T K G F V T A G H C G T V     40
61    TTYRSNGTNACNZGNGGNGCNACNAAPGGNTTYGTNACNGCNGGNCAYTGYGGNACNGTN     120

41    N A T A R I G G A V V G T F A A R V F P     60
121   AAYGCNACNGCNZGNATZGGNGGNGCNGTNGTNGGNACNTTYGCNGCNZGNGTNTTYCCN     180

61    G N D R A W V S L T S A Q T L L P R V A     80
181   GGNAAYGAYZGNGCNTGGGTNRSNYTNACNRSNGCNCAPACNYTNYTNCCNZGNGTNGCN     240

81    N G S S F V T V R G S T E A A V G A A V     100
241   AAYGGNRSNRSNTTYGTNACNGTNZGNGGNRSNACNGAPGCNGCNGTNGGNGCNGCNGTN     300

101   C P S G R T T G Y Q C G T T T A K N V T     120
301   TGYZGNRSNGGNZGNACNACNGGNTAYCAPTGYGGNACNATZACNGCNAAPAAYGTNACN     380

121   A N Y A E G A V R G L T Q G N A C M G R     140
381   GCNAAYTAYGCNGAPGGNGCNGTNZGNGGNYTNACNCAPGGNAAYGCNTGYATGGGNZGN     420

141   G P S G G S W I T S A G Q A Q G V M S G     160
421   GGNGAYRSNGGNGGNRSNTGGATZACNRSNGCNGGNCAPGCNCAPGGNGTNATGRSNGGG     480

161   G N V Q S N G N N C G I P A S Q R S S L     180
481   GGNAAYGTNCAPRSNAAYGGNAAYAAYTGYGGNATZCCNGCNRSNCAPZGNRSNRSNYTN     540

181   F E R L Q P I L S Q Y G L S L V T G     198
541   TTYGAPZGNYTNCAPCCNATZYTNRSNCAPTAYGGNYTNRSNYTNGTNACNGGN     594
```

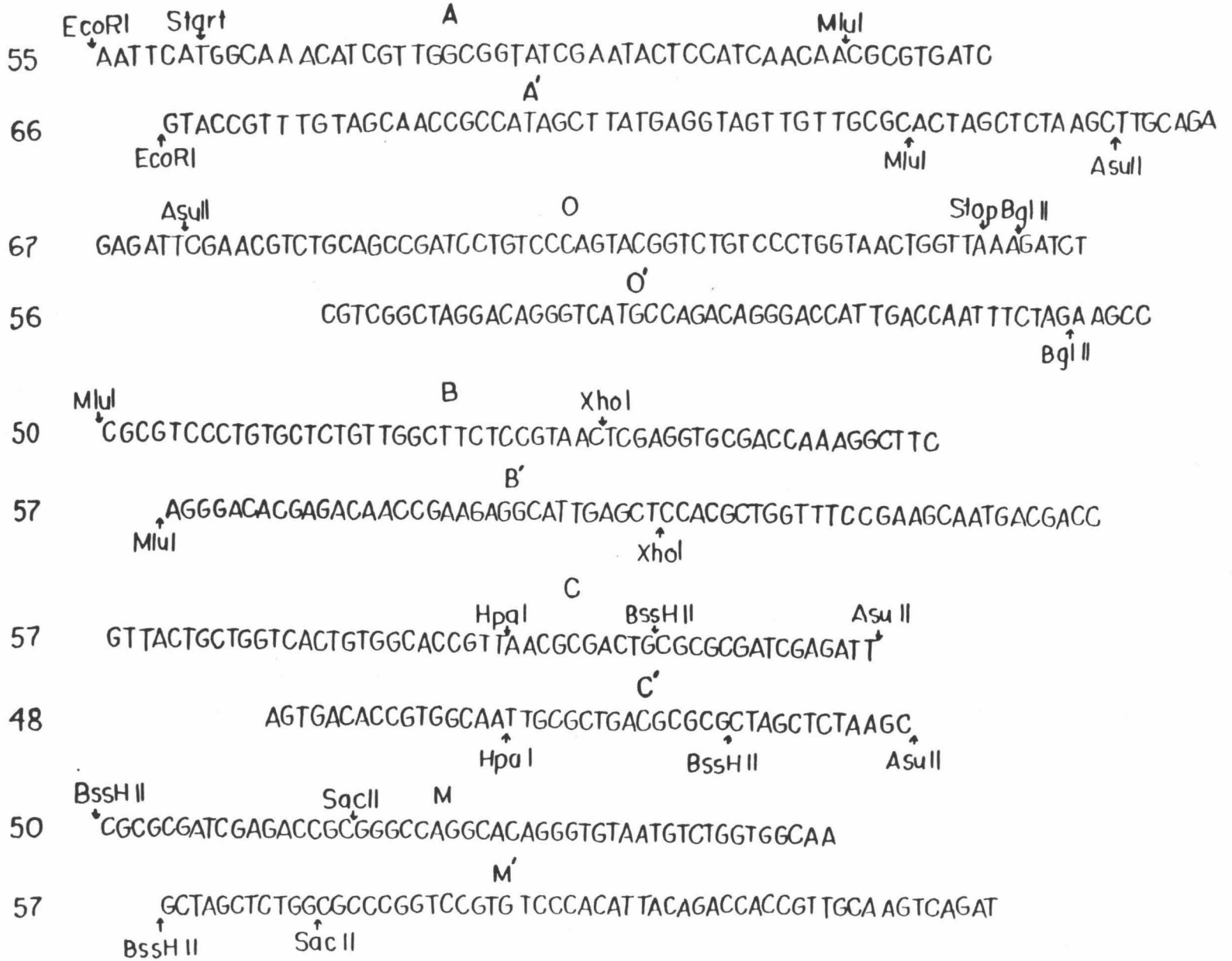
FIGURE 3

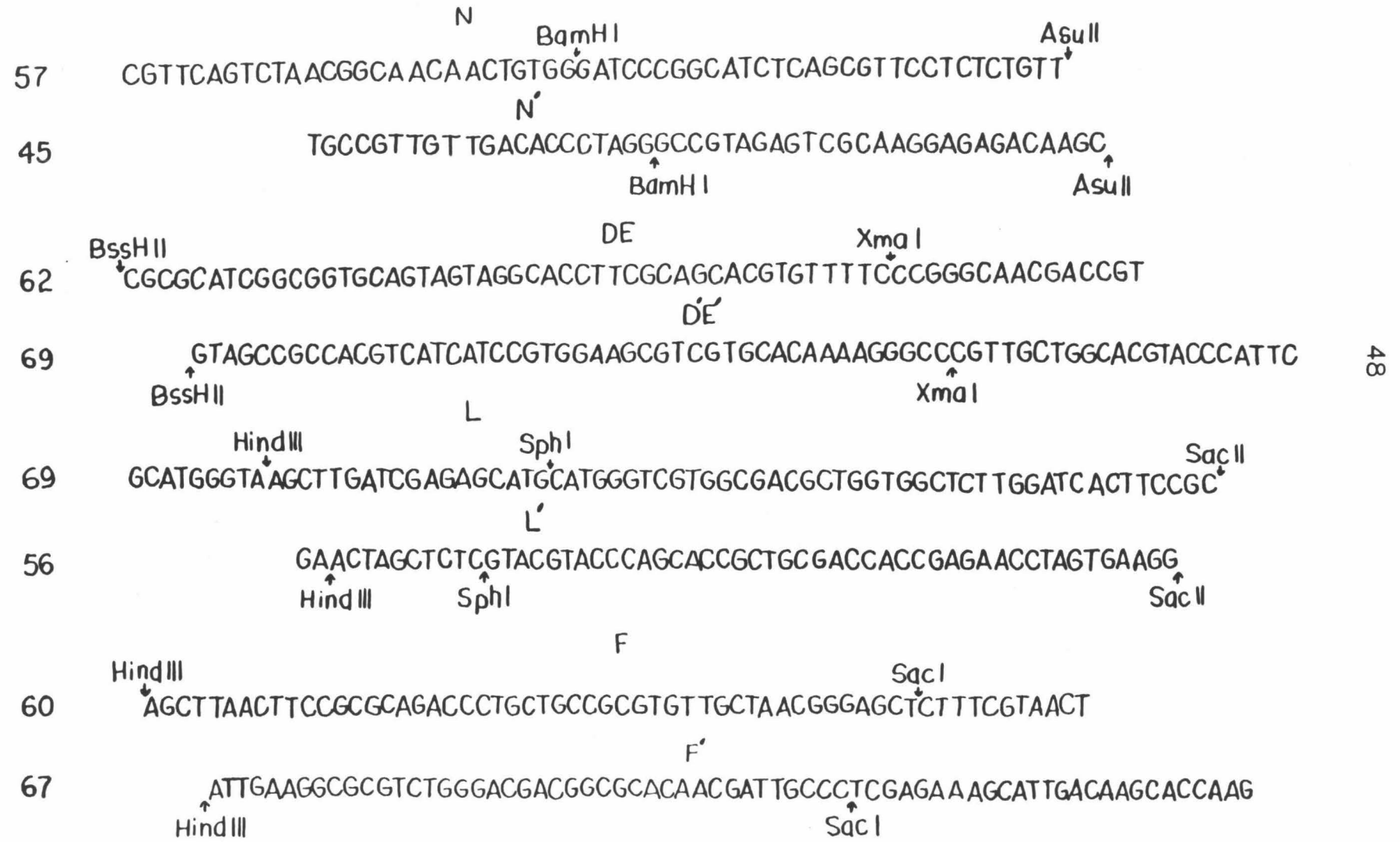
A synthetic gene for α-lytic protease. Restriction enzyme sites are boxed. The synthesized oligonucleotides are labeled A through O.


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
EcoRI
Start
MluI
XhoI
HpaI
BsaHII
XmaI
HindIII
SacI
XmaIII
BsiEII
KpnI
NaeI
StuI
SphI
SacII
BamHI
AsuII
Stop
BglII

FIGURE 4

Oligonucleotides synthesized for the gene of α-lytic protease. Restriction sites are indicated by arrows. Oligomers are labelled A through O with A′ through O′ being the complementary sequence.


EcoRI
StgrI
A
MluI
55
AATTCATGGCAAACATCGTTGGCGGTATCGAATACTCCATCAACAACGCGTGATC
A'
66
GTACCGTTTGTAGCAACCGCCATAGCTTATGAGGTAGTTGTTGCGCACTAGCTCTAAGCTTGCAGA
EcoRI
MluI
AsuII
AsuII
O
StopBglII
67
GAGATTCGAACGTCTGCAGCCGATCCTGTCCCAGTACGGTCTGTCCCTGGTAACTGGTTAAAGATCT
O'
56
CGTCGGCTAGGACAGGGTCATGCCAGACAGGGACCATTGACCAATTTCTAGAAGCC
BglII
MluI
B
XhoI
50
CGCGTCCCTGTGCTCTGTTGGCTTCTCCGTAACTCGAGGTGCGACCAAAGGCTTC
B'
57
AGGGACACGAGACAACCGAAGAGGCATTGAGCTCCACGCTGGTTTCCGAAGCAATGACGACC
MluI
XhoI
C
HpaI
BssHII
AsuII
57
GTTACTGCTGGTCACTGTGGCACCGTTAACGCGACTGCGCGCGATCGAGATT
C'
48
AGTGACACCGTGGCAATTGCGCTGACGCGCGCTAGCTCTAAGC
HpaI
BssHII
AsuII
BssHII
SacII
M
50
CGCGCGATCGAGACCGCGGGCCAGGCACAGGGTGTAATGTCTGGTGGCAA
M'
57
GCTAGCTCTGGCGCCCGGTCCGTGTCCCACATTACAGACCACCGTTGCAAGTCAGAT
BssHII
SacII


N
BamH I
Asu II
57 CGTTCAGTCTAACGGCAACAACTGTGGGATCCCGGCATCTCAGCGTTCCTCTCTGTT
N'
45 TGCCGTTGTTGACACCCTAGGGCCGTAGAGTCGCAAGGAGAGACAAGC
BamH I
Asu II
DE
Xma I
BssH II
62 CGCGCATCGGCGGTGCAGTAGTAGGCACCTTCGCAGCACGTGTTTTCCCGGGCAACGACCGT
DE'
69 GTAGCCGCCACGTCATCATCCGTGGAAGCGTCGTGCACAAAAGGGCCCGTTGCTGGCACGTACCCATTC
BssH II
Xma I
L
Hind III
Sph I
Sac II
69 GCATGGGTAAGCTTGATCGAGAGCATGCATGGGTCGTGGCGACGCTGGTGGCTCTTGGATCACTTCCGC
L'
56 GAACTAGCTCTCGTACGTACCCAGCACCGCTGCGACCACCGAGAACCTAGTGAAGG
Hind III
Sph I
Sac II
F
Hind III
Sac I
60 AGCTTAACTTCCGCGCAGACCCTGCTGCCGCGTGTTGCTAACGGGAGCTCTTTCGTAACT
F'
67 ATTGAAGGCGCGTCTGGGACGACGGCGCACAACGATTGCCCTCGAGAAAGCATTGACAAGCACCAAG
Hind III
Sac I

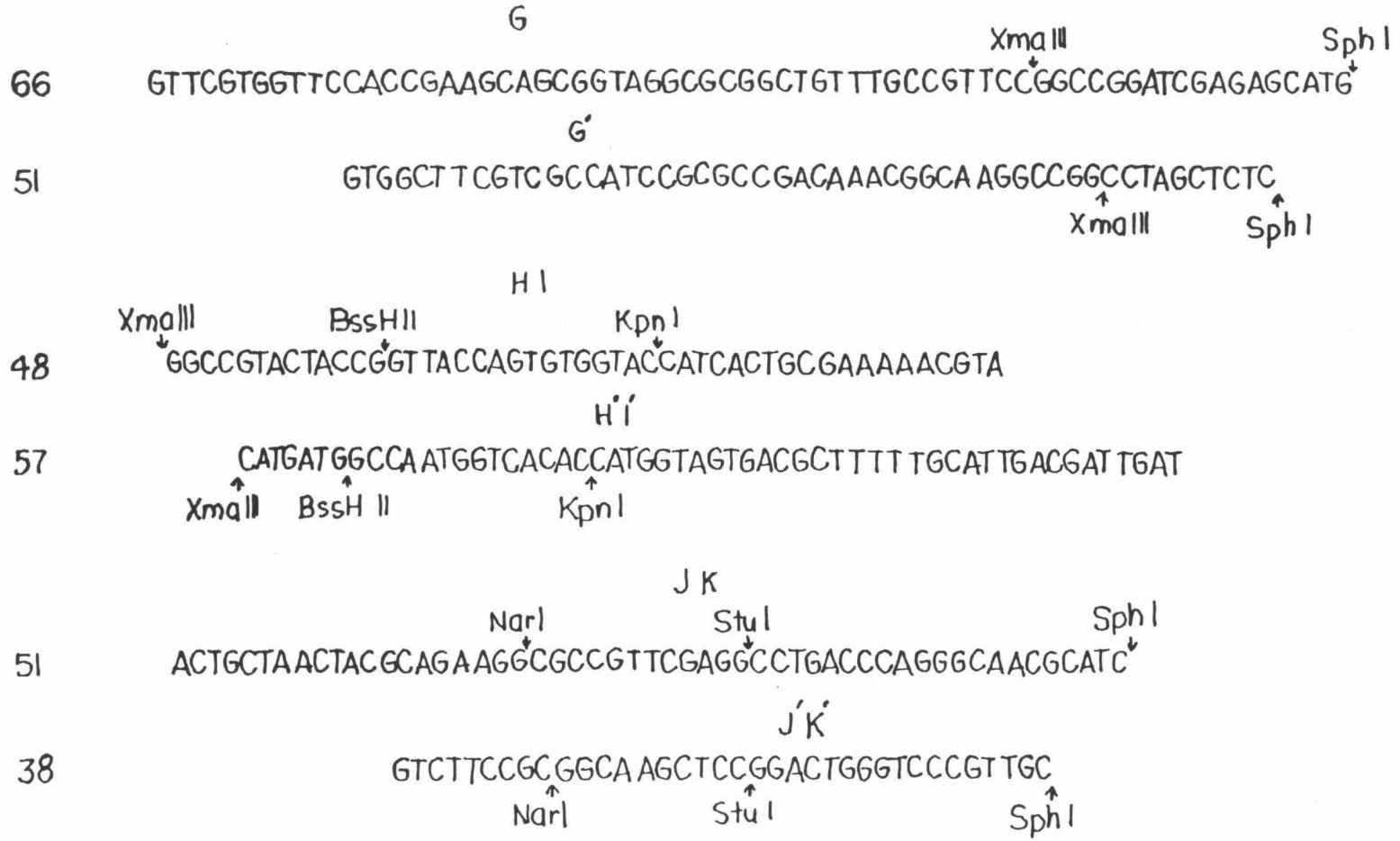

G
Xma III
Sph I
66
GTTCGTGGTTCCACCGAAGCAGCGGTAGGCGCGGCTGTTTGCCGTTCCGGCCGGATCGAGAGCATG
G'
51
GTGGCTTCGTCGCCATCCGCGCCGACAAACGGCAAGGCCGGCCTAGCTCTC
Xma III
Sph I
H I
Xma III
BssH II
Kpn I
48
GGCCGTACTACCGGTTACCAGTGTGGTACCATCACTGCGAAAAACGTA
H' I'
57
CATGATGGCCAATGGTCACACCATGGTAGTGACGCTTTTTGCATTGACGATTGAT
Xma III
BssH II
Kpn I
J K
Nar I
Stu I
Sph I
51
ACTGCTAACTACGCAGAAGGCGCCGTTCGAGGCCTGACCCAGGGCAACGCATC
J' K'
38
GTCTTCCGCGGCAAGCTCCGGACTGGGTCCCGTTGC
Nar I
Stu I
Sph I

FIGURE 5a

Construction of steps 1 through 3 of the synthetic gene of α-lytic protease.

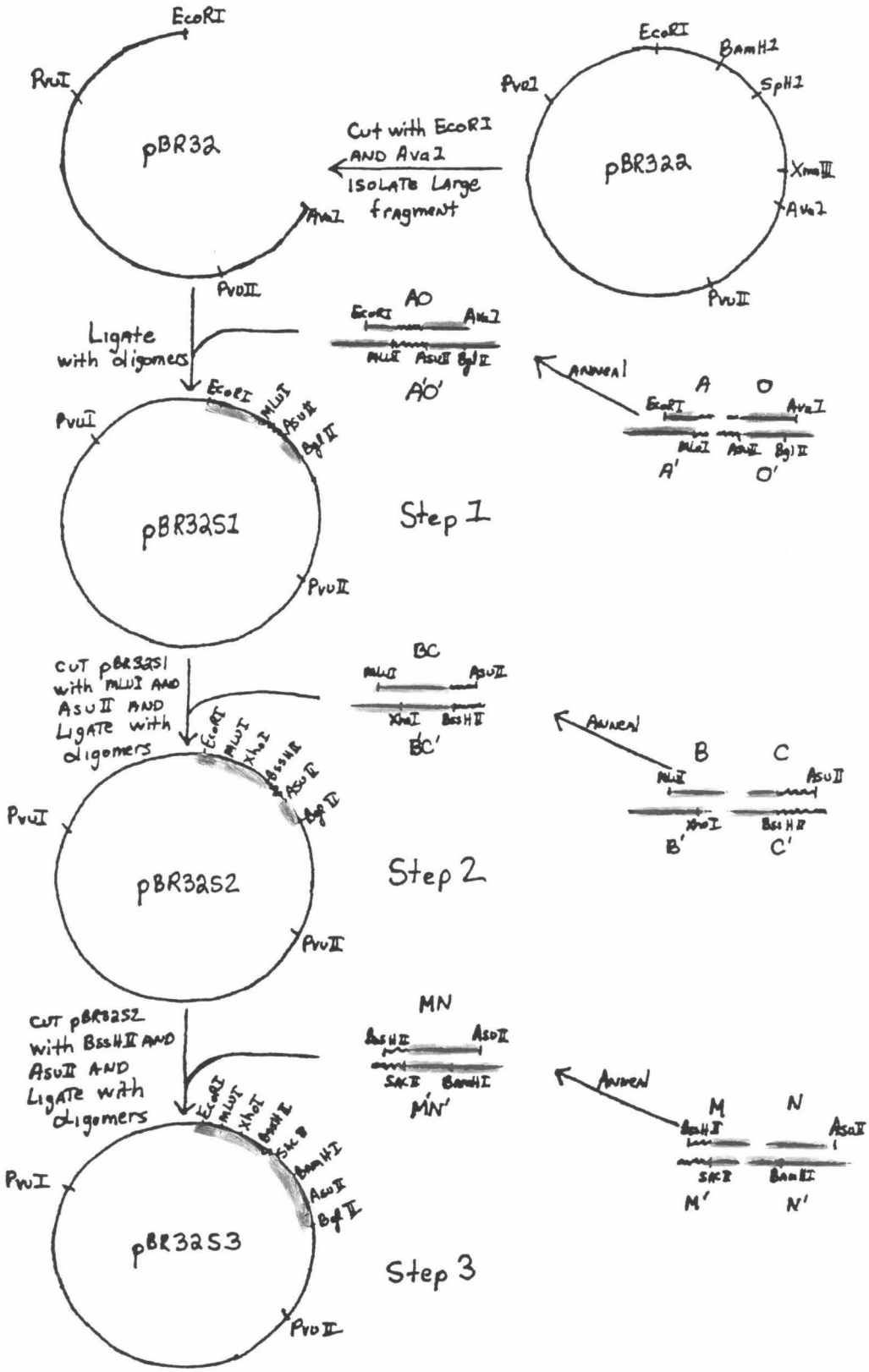

EcoRI
PvuI
pBR32
AvaI
PvuII
Cut with EcoRI
AND AvaI
ISOLATE Large
fragment
EcoRI
BamHI
PvuI
SphI
pBR322
XmnIII
AvaI
PvuII
Ligate
with oligomers
AO
EcoRI
AvaI
AluI AsuII BglII
A'O'
ANNEAL
A
O
EcoRI
AvaI
AluI
AsuII
BglII
A'
O'
pBR32S1
Step 1
CUT pBR32S1
with AluI AND
AsuII AND
Ligate with
oligomers
BC
AluI
AsuII
XhoI
BssHII
B'C'
ANNEAL
B
C
AluI
AsuII
XhoI
BssHII
B'
C'
pBR32S2
Step 2
CUT pBR32S2
with BssHII AND
AsuII AND
Ligate with
oligomers
MN
BssHII
AsuII
SacI
BamHI
M'N'
ANNEAL
M
N
BssHII
AsuII
SacI
BamHI
M'
N'
pBR32S3
Step 3

FIGURE 5b

Construction of steps 4 through 6 in the synthetic gene for α-lytic protease.

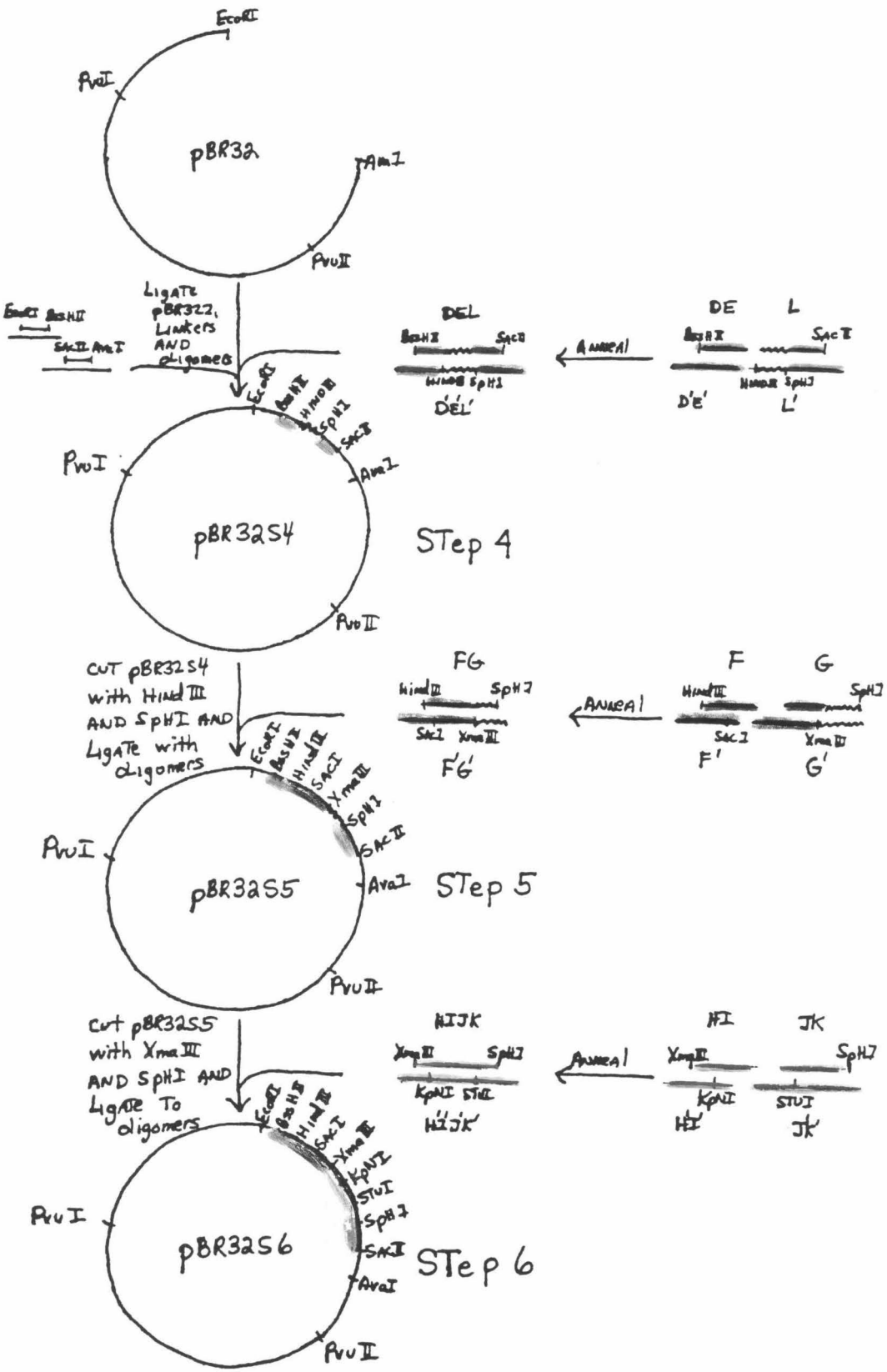

EcoRI
AvaI
pBR32
AmaI
PvuII
Ligate pBR322, Linkers AND oligomers
DEL
BssHII SacII
HindIII SphI
DÉL'
ANNEAL
DE L
D'E' L'
pBR32S4
STep 4
PvuI
AvaI
PvuII
CUT pBR32S4 with HindIII AND SphI AND Ligate with oligomers
FG
HindIII SphI
SacI XmaIII
FG'
F G
F' G'
pBR32S5
STep 5
CUT pBR32S5 with XmaIII AND SphI AND Ligate To oligomers
HIJK
XmaIII SphI
KpnI StuI
HI JK
H'I' J'K'
pBR32S6
STep 6

FIGURE 5c

Incorporating steps 4 through 6 into steps 1 through 3 in the construction of the synthetic gene of α-lytic protease.

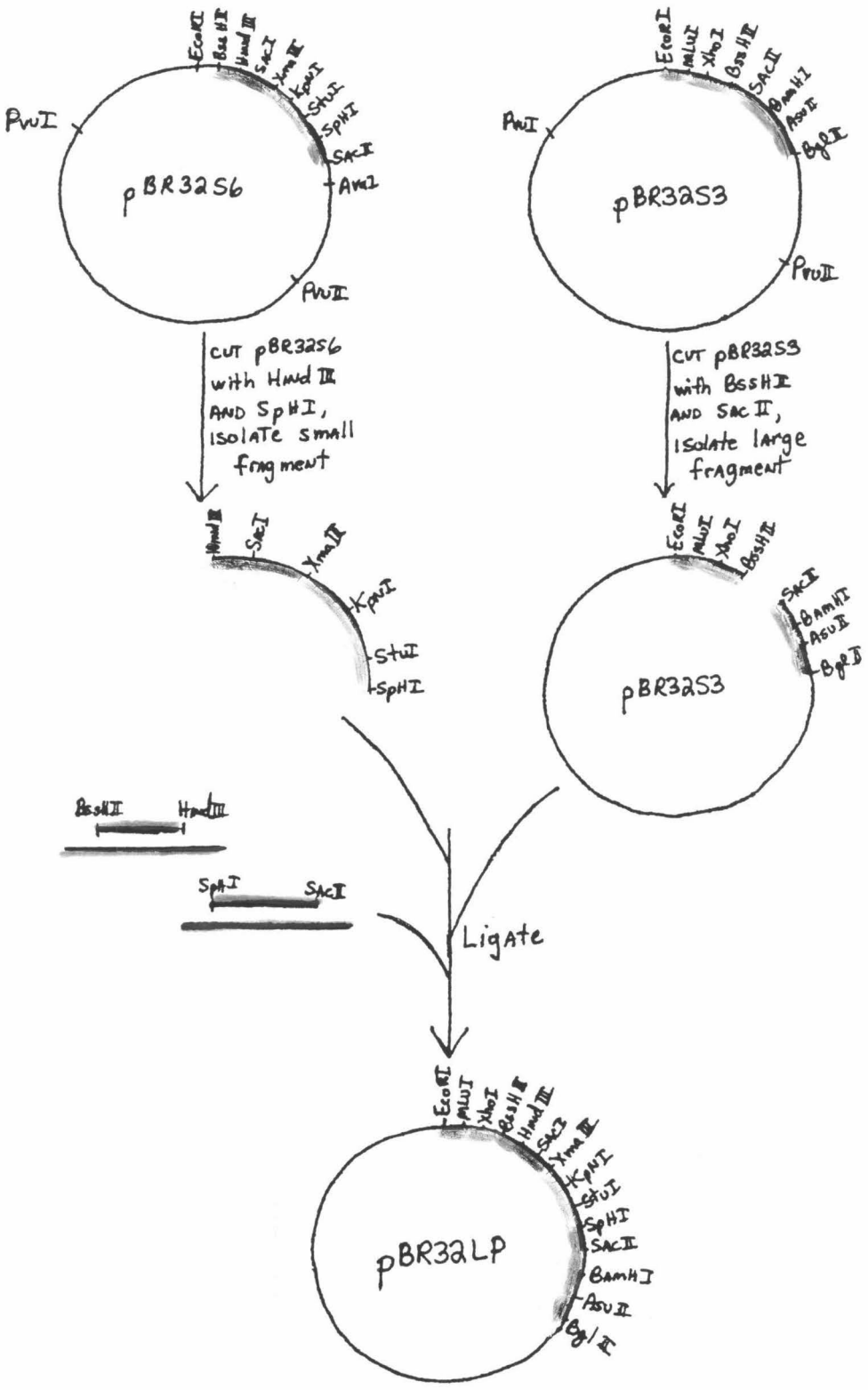

pBR32S6
PvuI
PvuII
AvaI
SpHI
SacI
KpnI
StuI
XmaIII
HindIII
BssHII
EcoRI
pBR32S3
PvuII
SacII
BamHI
AsuII
BglII
MluI
XhoI
CUT pBR32S6 with HindIII AND SpHI, ISOLATE small fragment
CUT pBR32S3 with BssHII AND SacII, ISOLATE large fragment
BssHII
HindIII
SpHI
SacI
Ligate
pBR32LP

FIGURE 6

Restriction enzyme cleavages of the synthetic gene for α-lytic protease. Lane 1 is the uncut plasmid pBR32LP. Lane 2, plasmid pBR32LP cleaved by EcoR1; lane 3, cleavage by Mlu1; lane 4, cleavage by Xho1; lane 5, cleavage by BssHII; lane 6, cleavage by Xma1; lane 7, cleavage by HindIII; lane 8, cleavage by Sac1; lane 9,cleavage by Eag1(XmaIII); lane 10, cleavage by BsteII; lane 11, cleavage by Kpn1; lane 12, cleavage by Nar1; lane 13, cleavage by Stu1; lane 14, cleavage by Sph1; lane 15, cleavage by SacII; lane 16, cleavage by BamHI; lane 17, cleavage by AsuII; lane 18, cleavage by Ava1.

LANES

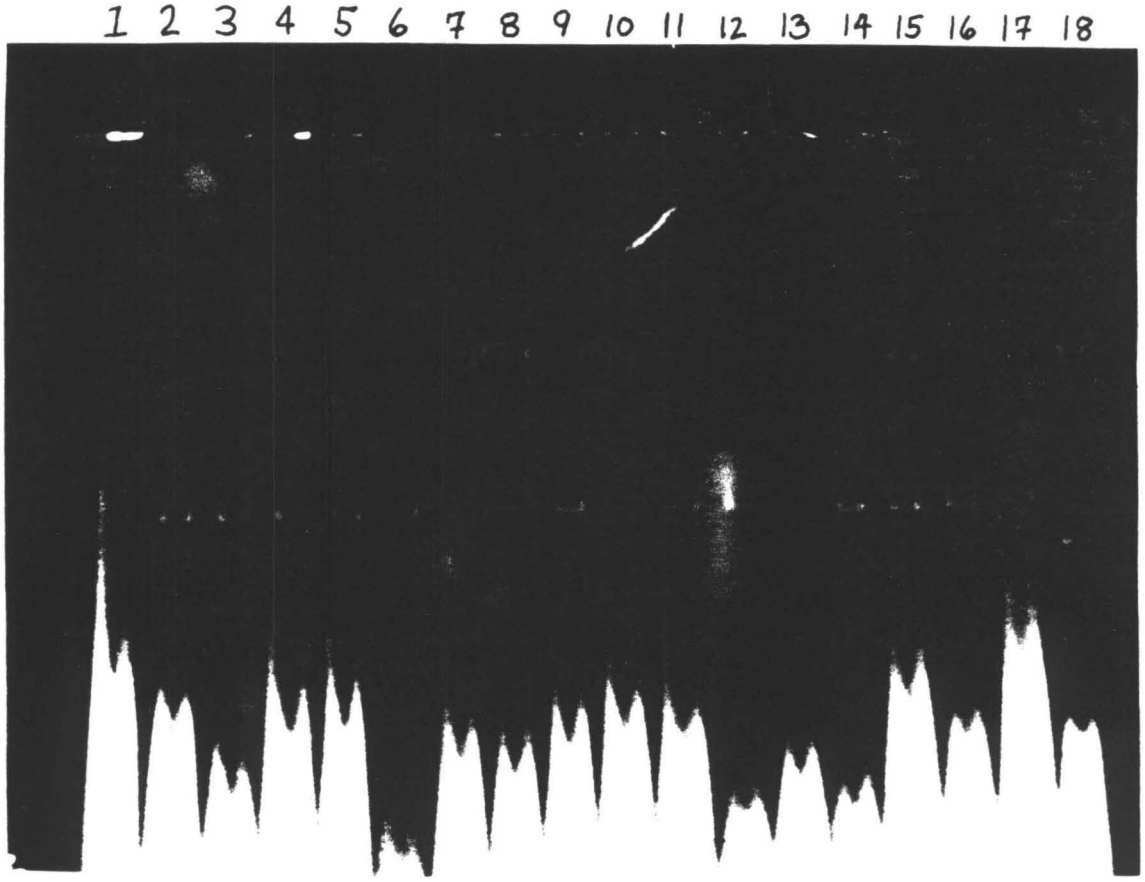


FIGURE 7

DNA sequencing of the synthetic gene for α-lytic protease. Left column letters indicate the DNA sequence of the gene mutated at the BglII site and the right column letters indicate the correct sequence.

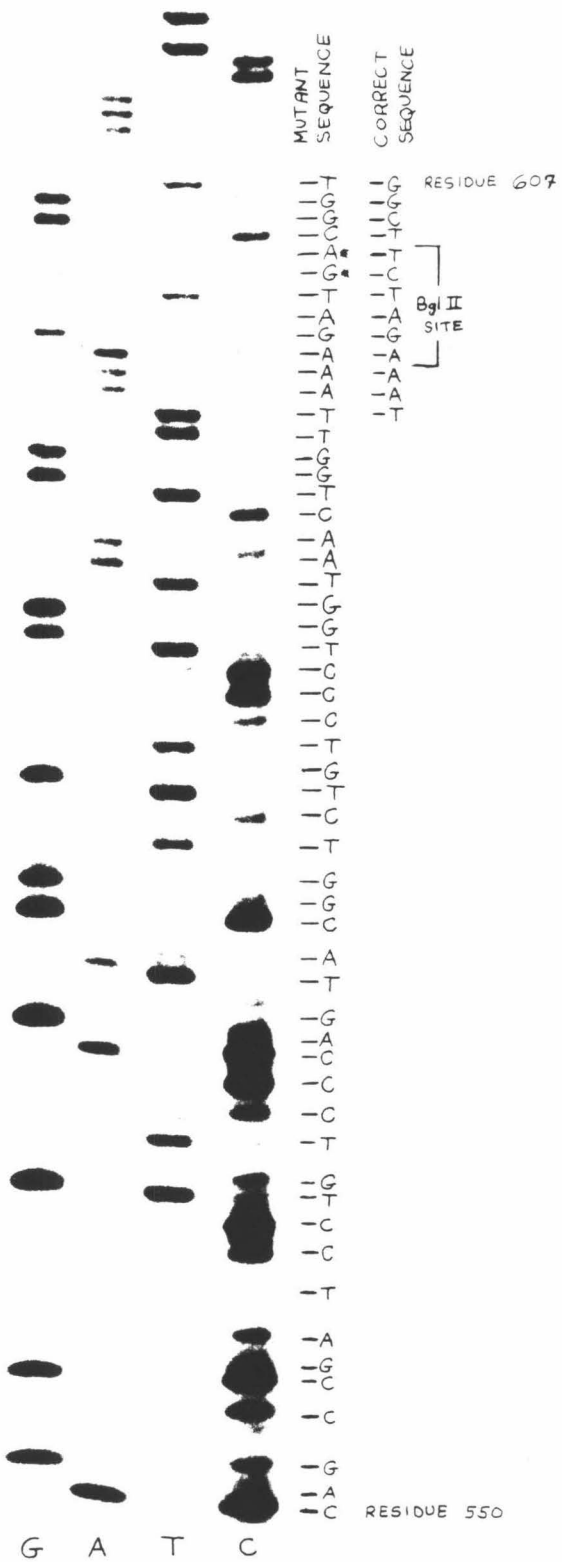

MUTANT SEQUENCE
CORRECT SEQUENCE
RESIDUE 607
Bgl II SITE
RESIDUE 550
G A T C

FIGURE 8

DNA sequencing of the synthetic gene for α-lytic protease.

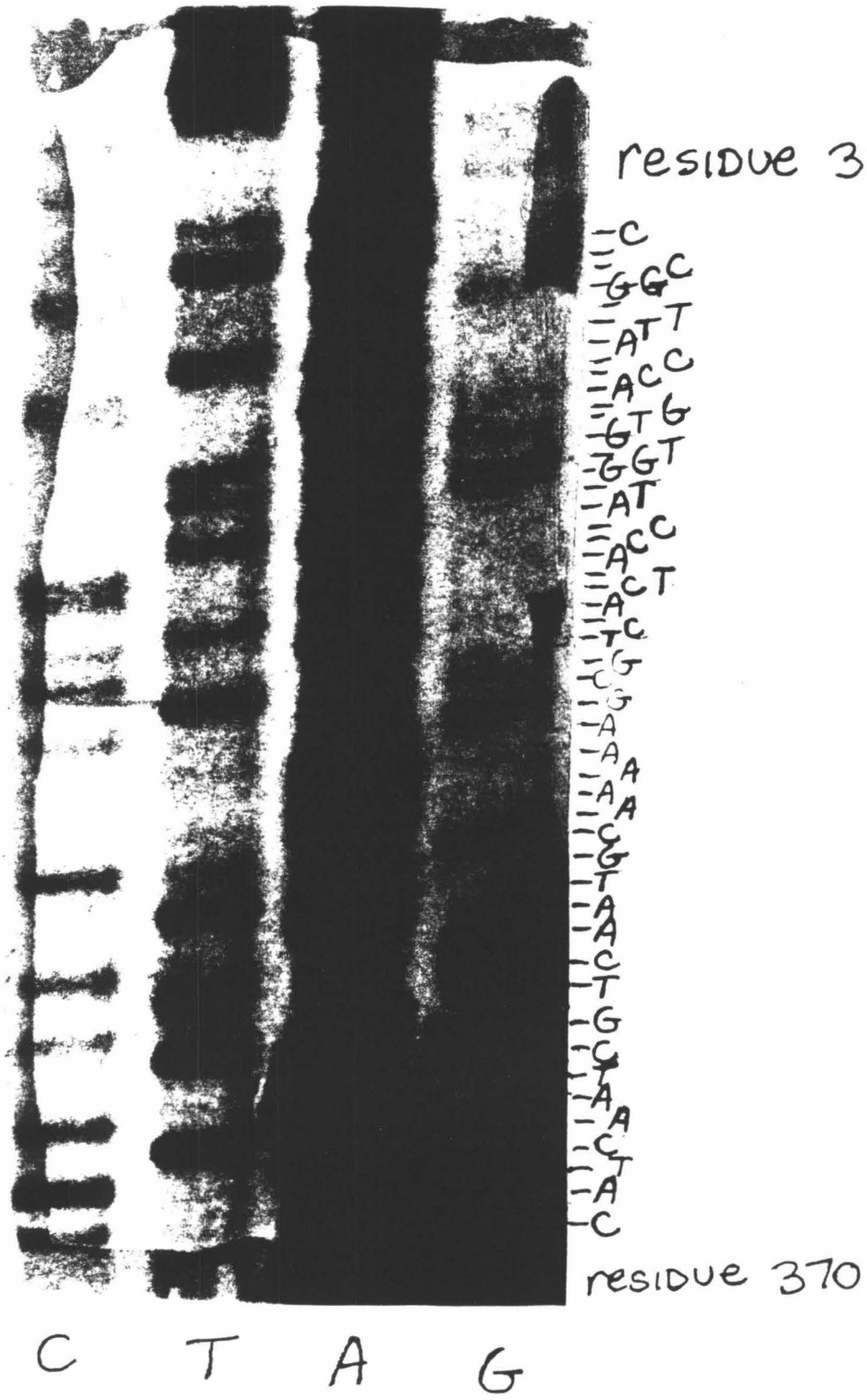

residue 320
residue 370
C T A G

FIGURE 9

DNA sequencing of the synthetic gene for ∝-lytic protease. According to the procedure used, the C and T lanes contain an extra phosphate group which results in those lanes running faster down the sequencing gel. Therefore, to correctly read the sequence, the G and A lanes need to be visually pulled down about 1.0 cm. The sequence written is in the correct order.

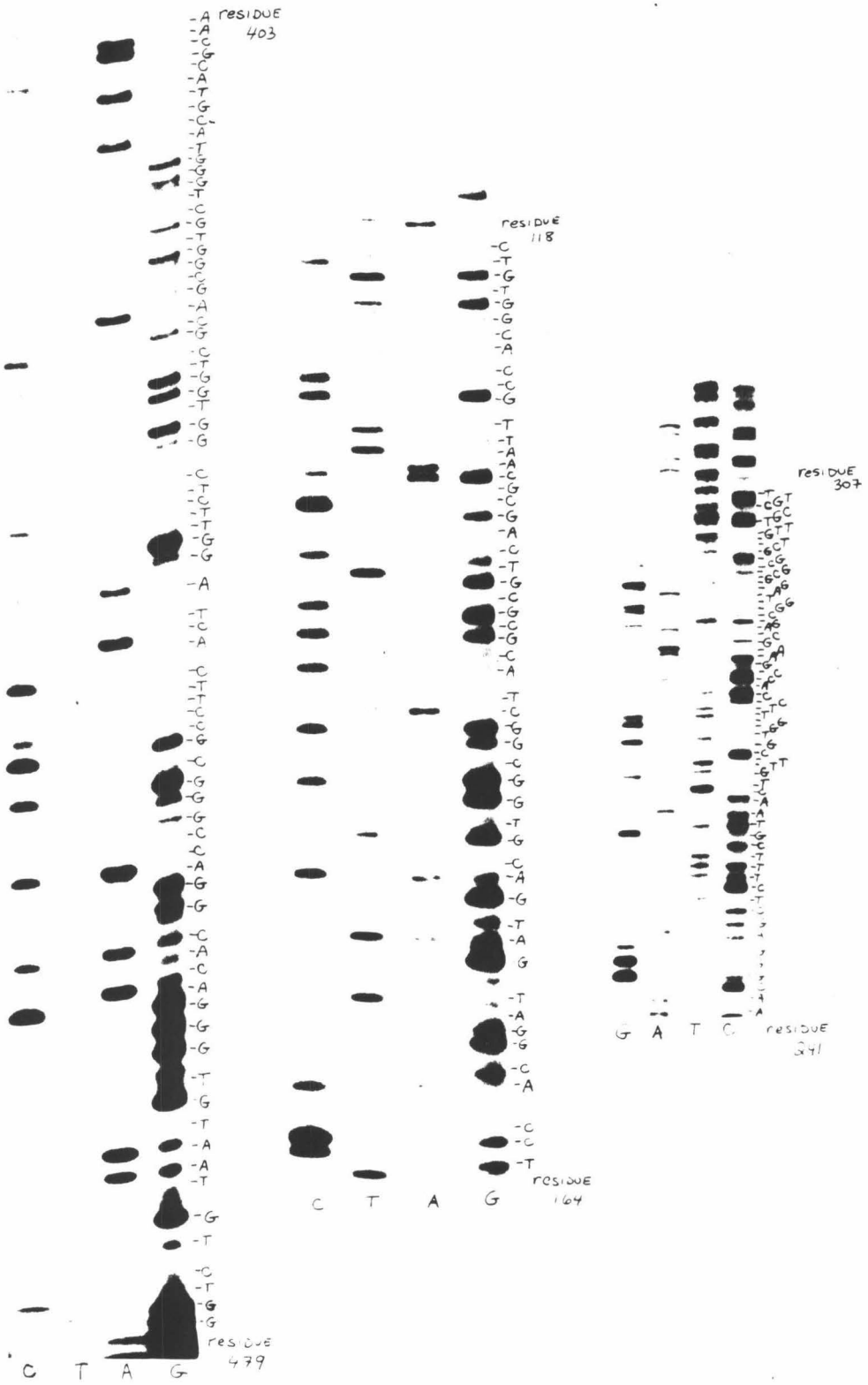

residue 403
residue 479
C T A G
residue 118
residue 164
C T A G
residue 307
residue 341
G A T C

FIGURE 10

DNA sequencing of the synthetic gene for α-lytic protease. According to the procedure used, the G and A lanes contain an extra phosphate group which results in those lanes running faster down the sequencing gel. Therefore, to correctly read the sequence, the T and C lanes need to be visually pulled down about 0.5 cm. The sequence written is in the correct order.

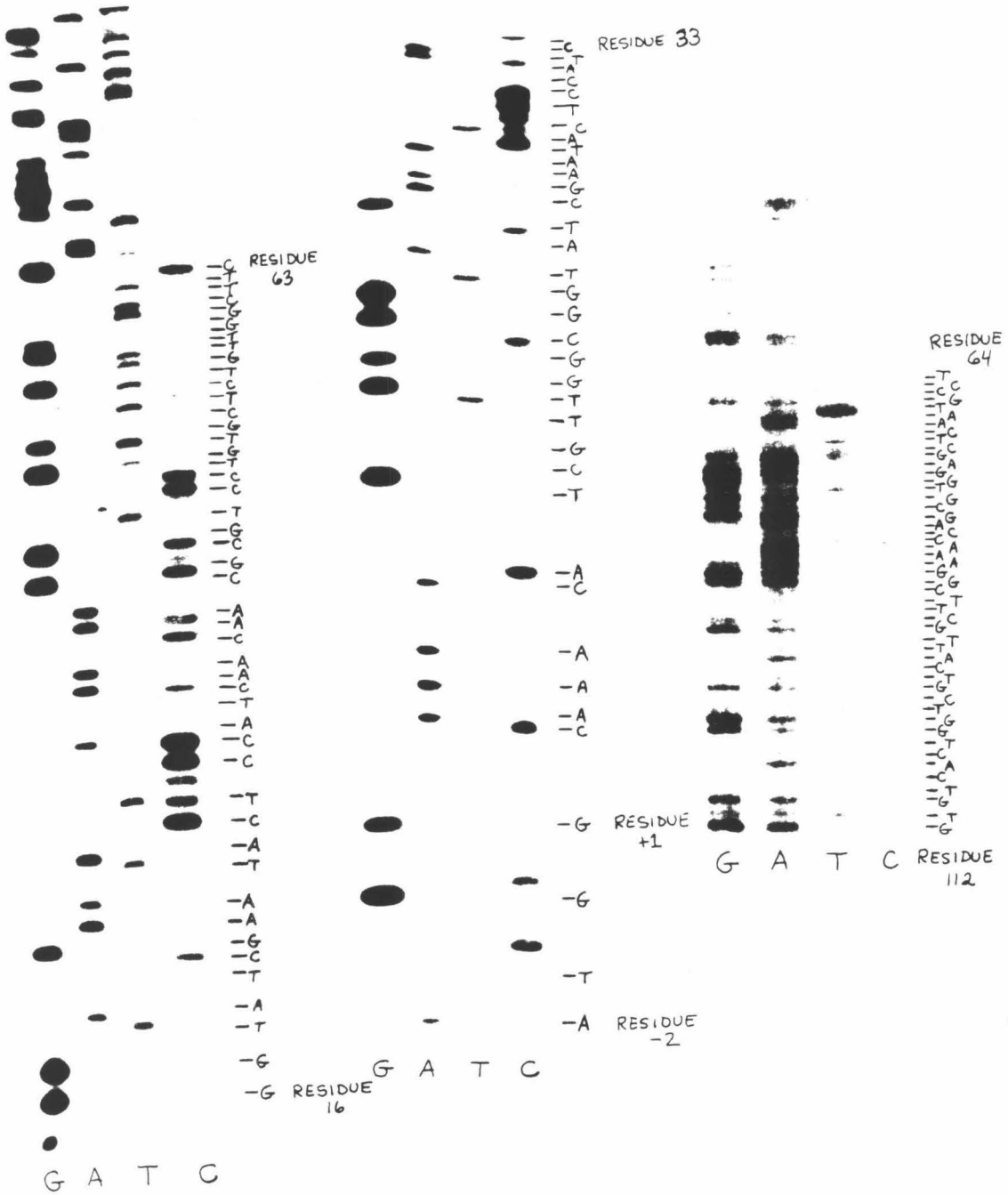

RESIDUE 63
RESIDUE 33
RESIDUE 64
RESIDUE +1
RESIDUE -2
RESIDUE 16
RESIDUE 112
G A T C
G A T C
G A T C

FIGURE 11

DNA sequencing of the synthetic gene for α-lytic protease. The G and A lanes for the sequence for residues 165 to 200 and 480 to 502 need to be pulled down 0.75 cm to read the sequence correctly. This is because an extra phosphate group is present on lanes T and C which result in their running faster on the gel. The T and C lanes for the sequence of residues 549 to 516 needs to be pulled down 1.0 cm.

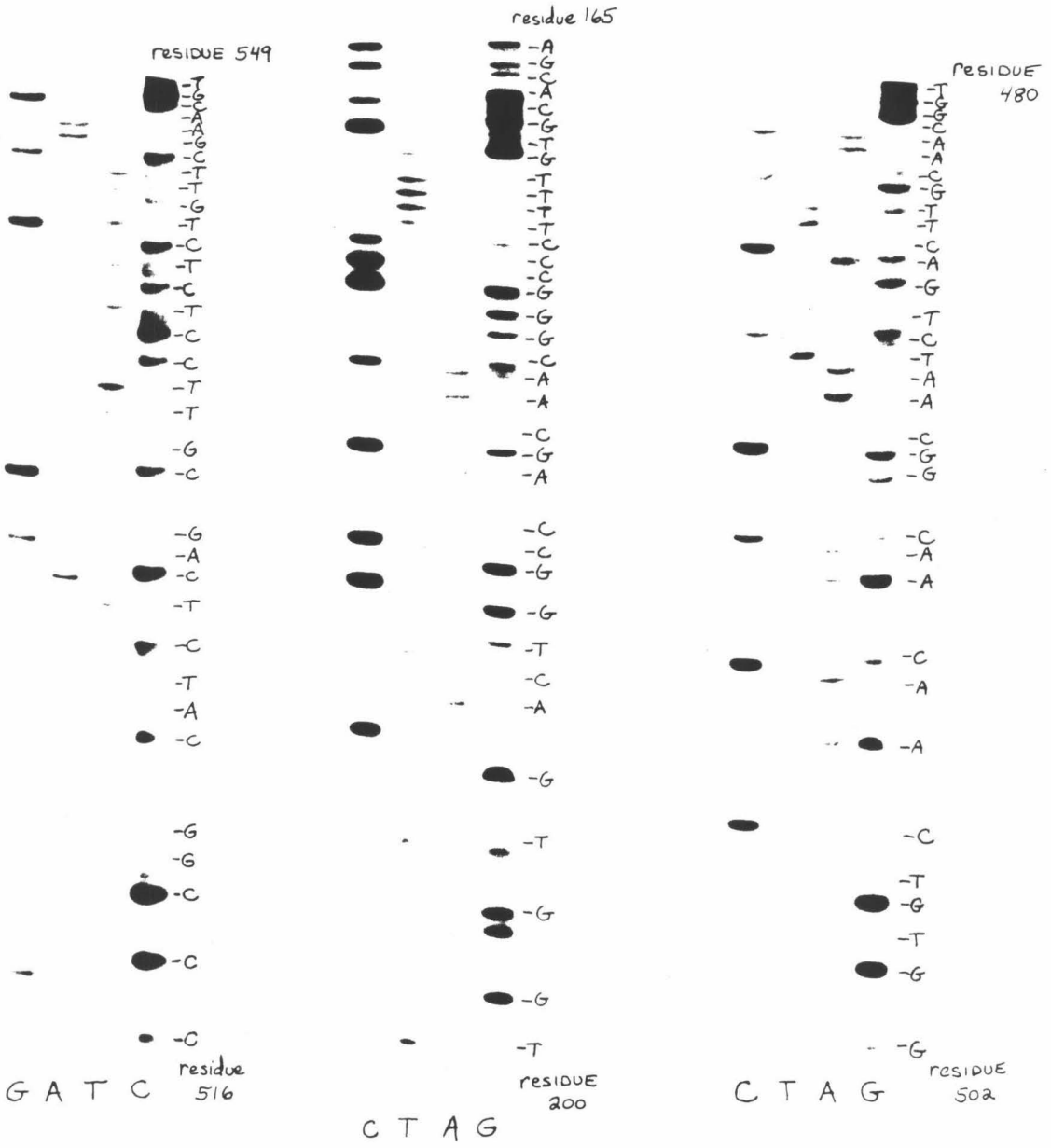

residue 549
residue 165
residue 480
G A T C
residue 516
C T A G
residue 200
C T A G
residue 502

FIGURE 12

Three fragment ligation strategy used in the correction of the mutated BglII restriction site.

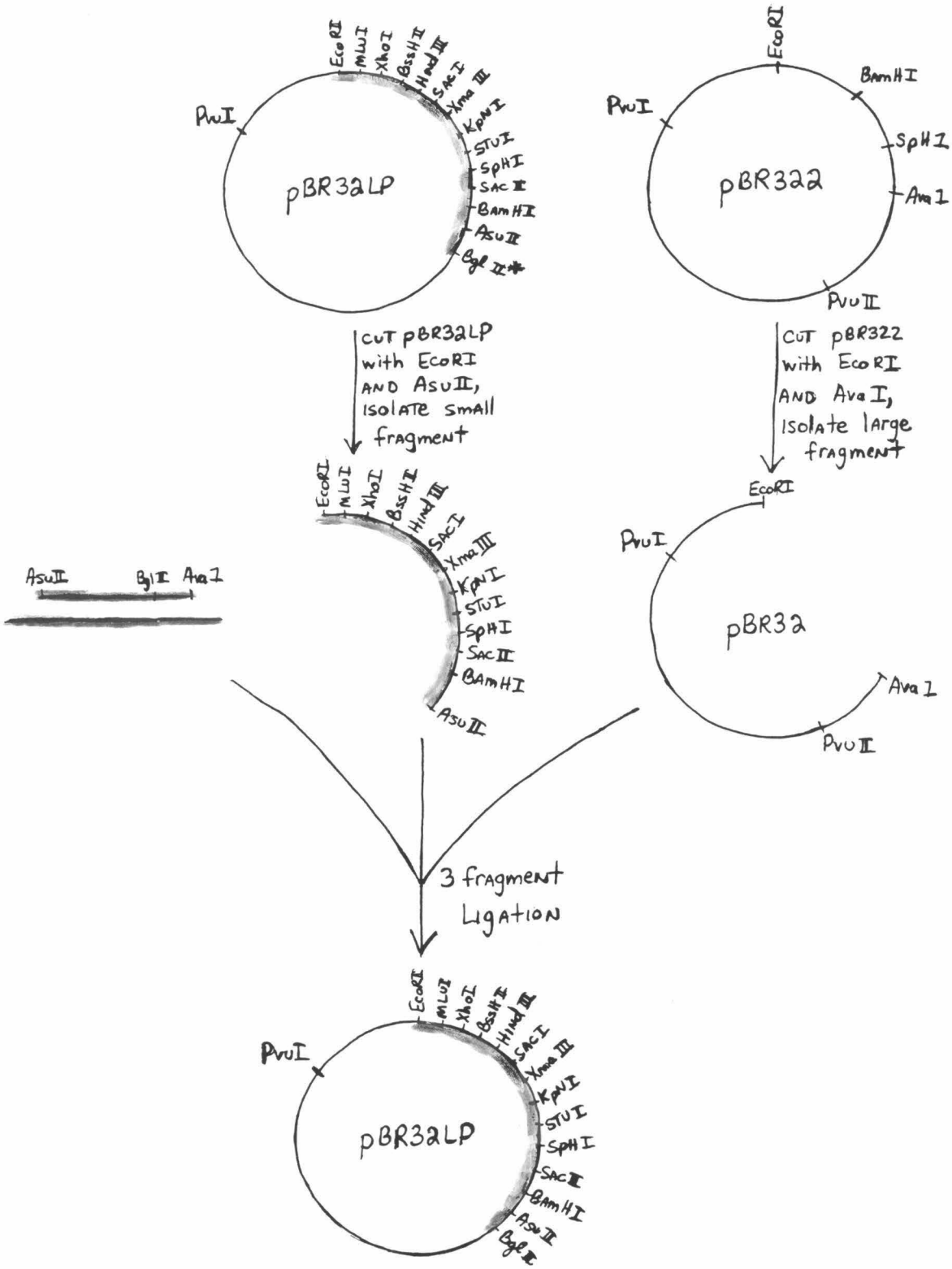

pBR32LP
EcoRI
MluI
XhoI
BssHII
HindIII
SacI
XmaIII
KpnI
StuI
SphI
SacII
BamHI
AsuII
BglII
PvuI
pBR322
EcoRI
BamHI
SphI
AvaI
PvuII
PvuI
CUT pBR32LP with EcoRI AND AsuII, ISOLATE small fragment
CUT pBR322 with EcoRI AND AvaI, ISOLATE large fragment
AsuII
BglII
AvaI
pBR32
EcoRI
PvuI
AvaI
PvuII
3 fragment LIGATION
pBR32LP
EcoRI
MluI
XhoI
BssHII
HindIII
SacI
XmaIII
KpnI
StuI
SphI
SacII
BamHI
AsuII
BglII
PvuI

FIGURE 13

Restriction map of the corrected gene for α-lytic protease. Lane 1 is the uncut plasmid pBR32LP; lane 2, pBR32LP cleaved by PvuII and EcoRI; lane 3, cleavage by PvuII and XhoI; lane 4, cleavage by PvuII and HindIII; lane 5, cleavage by PvuII and KpnI; lane 6, cleavage by PvuII and BamHI; lane 7, cleavage by PvuII and BglII and lane 8, cleavage by PvuII and AvaI. The larger fragment increases in length starting at lane 2 with EcoRI and PvuII through lane 8. This indicates that BglII is after BamHI in the sequence and so on.

LANES

Table 1. Preferred Codons in Yeast and E. coli

Preferred Codon(s) in Yeast:	Amino Acid	Preferred Codon(s) in *E. coli* and use		Use of all other triplets for that amino acid
GCU, GCC	Ala	GCC not used; no clear preference		
UCU, UCC	Ser	UCU, UCC	27	7
ACU, ACC	Thr	ACU, ACC	53	4
GUU, GUC	Val	GUU, GUA	62	8
AUU, AUC	Ile	AUC	43	4
GAC	Asp	GAC	43	11
UUC	Phe	UUC	20	3
UAC	Tyr	UAC	25	3
UGU	Cys	no clear preference		
AAC	Asn	AAC	31	1
CAC	His	CAC	12	4
GAA	Glu	GAA	40	10
GGU	Gly	GGU, GGC	80	6
CAA	Gln	CAG	28	2
AAG	Lys	AAA	17	6
CCA	Pro	CCG	34	5
UUG	Leu	CUG	56	2
AGA	Arg	CGU	33	7

Table 2. Codon Usage in the Synthetic Gene of α-Lytic Protease

AMINO ACID	CODON	#	%
A-Ala		25	12.63
	GCA	10	5.05
	GCG	9	4.55
	GCT	5	2.53
	GCC	1	.51 (Bl)
C-Cys		6	3.03
	TGT	3	1.52
	TGC	3	1.52
D-Asp		2	1.01
	GAC	2	1.01
E-Glu		4	2.02
	GAA	4	2.02
F-Phe		6	3.03
	TTC	6	3.03
G-Gly		32	16.16
	GGG	2	1.01 (Bl)
	GGT	13	6.57
	GGC	17	8.59
H-His		1	.51
	CAC	1	.51
I-Ile		8	4.04
	ATT	1	.51 (Bl)
	ATC	7	3.54
K-Lys		2	1.01
	AAA	2	1.01
L-Leu		10	5.05
	TTA	1	.51 (Bl)
	CTG	9	4.55
M-Met		2	1.01
	ATG	2	1.01
N-Asn		13	6.57
	AAC	13	6.57
P-Pro		4	2.02
	CCG	4	2.02
Q-Gln		9	4.55
	CAG	9	4.55
R-Arg		12	6.06
	CGA	2	1.01 (Bl)
	CGT	9	4.55
	CGC	1	.51 (Bl)
S-Ser		19	9.60
	AGC	2	1.01 (Bl)
	TCT	7	3.54
	TCC	10	5.05
T-Thr		18	9.09
	ACT	10	5.05
	ACC	8	4.04
V-Val		19	9.60
	GTA	9	4.55
	GTT	10	5.05
W-Trp		2	1.01
	TGG	2	1.01
X-Term		0	0.00
Y-Tyr		4	2.02
	TAC	4	2.02

Table 3. Codon Analysis of Codons Used in the Synthetic Gene

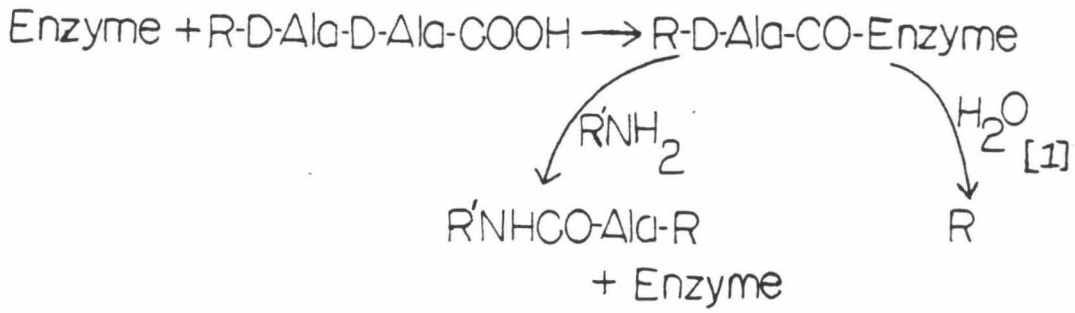
TOTAL CODONS 198

	A	G	T	C	P	Y	
A	2(1.0)	0(0.0)	0(0.0)	0(0.0)	2(1.0)	0(0.0)	A
	0(0.0)	0(0.0)	2(1.0)	0(0.0)	0(0.0)	2(1.0)	G
	0(0.0)	0(0.0)	1(.5)	10(5.1)	0(0.0)	11(5.6)	T
	13(6.6)	2(1.0)	7(3.5)	8(4.0)	15(7.6)	15(7.6)	C
	2(1.0)	0(0.0)	2(1.0)	0(0.0)	0(0.0)	0(0.0)	P
	13(6.6)	2(1.0)	8(4.0)	18(9.1)	0(0.0)	0(0.0)	Y
G	4(2.0)	0(0.0)	9(4.5)	10(5.1)	4(2.0)	19(9.6)	A
	0(0.0)	2(1.0)	0(0.0)	9(4.5)	2(1.0)	9(4.5)	G
	0(0.0)	13(6.6)	10(5.1)	5(2.5)	13(6.6)	15(7.6)	T
	2(1.0)	17(8.6)	0(0.0)	1(.5)	19(9.6)	1(.5)	C
	4(2.0)	2(1.0)	9(4.5)	19(9.6)	0(0.0)	0(0.0)	P
	2(1.0)	30(15.2)	10(5.1)	6(3.0)	0(0.0)	0(0.0)	Y
T	0(0.0)	0(0.0)	1(.5)	0(0.0)	0(0.0)	1(.5)	A
	0(0.0)	2(1.0)	0(0.0)	0(0.0)	2(1.0)	0(0.0)	G
	0(0.0)	3(1.5)	0(0.0)	7(3.5)	3(1.5)	7(3.5)	T
	4(2.0)	3(1.5)	6(3.0)	10(5.1)	7(3.5)	16(8.1)	C
	0(0.0)	2(1.0)	1(.5)	0(0.0)	0(0.0)	0(0.0)	P
	4(2.0)	6(3.0)	6(3.0)	17(8.6)	0(0.0)	0(0.0)	Y
C	0(0.0)	2(1.0)	0(0.0)	0(0.0)	2(1.0)	0(0.0)	A
	9(4.5)	0(0.0)	9(4.5)	4(2.0)	9(4.5)	13(6.6)	G
	0(0.0)	9(4.5)	0(0.0)	0(0.0)	9(4.5)	0(0.0)	T
	1(.5)	1(.5)	0(0.0)	0(0.0)	2(1.0)	0(0.0)	C
	9(4.5)	2(1.0)	9(4.5)	4(2.0)	0(0.0)	0(0.0)	P
	1(.5)	10(5.1)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	Y
P	6(3.0)	0(0.0)	9(4.5)	10(5.1)	0(0.0)	0(0.0)	A
	0(0.0)	2(1.0)	2(1.0)	9(4.5)	0(0.0)	0(0.0)	G
	0(0.0)	13(6.6)	11(5.6)	15(7.6)	0(0.0)	0(0.0)	T
	15(7.6)	19(9.6)	7(3.5)	9(4.5)	0(0.0)	0(0.0)	C
	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	P
	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	Y
Y	0(0.0)	2(1.0)	1(.5)	0(0.0)	0(0.0)	0(0.0)	A
	9(4.5)	2(1.0)	9(4.5)	4(2.0)	0(0.0)	0(0.0)	G
	0(0.0)	12(6.1)	0(0.0)	7(3.5)	0(0.0)	0(0.0)	T
	5(2.5)	4(2.0)	6(3.0)	10(5.1)	0(0.0)	0(0.0)	C
	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	P
	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	Y

CHAPTER 3

Purification and Kinetic Characterization of a T71H Mutant of RTEM-1 Beta-lactamase

The cell walls of bacteria are composed of cross-linked N-acatylmuramyl or N-acetylglucosyl polysaccharides known as peptidoglycan. The cross-linking of the peptidoglycan is regulated by D-alanyl-transpeptidases or D-alanyl-carboxypeptidases as seen in reaction 1 (1,2).



Cross-linking of the peptidoglycan is inhibited by beta-lactam antibiotics (penicillins and cephalosporins) because of the covalent binding of the antibiotic to the D-alanyl-carboxypeptidase-transpeptidase. Bacteria produce β-lactamases which catalyze the hydrolysis of the beta-lactam ring of penicillins and cephalosporins (3) and thus protect the cell wall from destruction due to the lack of cross-linking. Therefore, the usefulness of studying beta-lactamases and beta-lactam chemistry is obvious to the biochemist and clinician. Greater knowledge of beta-lactamase's catalytic mechanism and structure will help in designing new beta-lactam antibiotics to combat bacteria now resistant to common beta-lactams.

Close to 100 different beta-lactamases have been discovered and they are divided into three different classes (A,B and C) depending upon their structure and their activity towards different beta-lactams (4). Class A beta-lactamases include the enzymes from B. licheniformis 749/c, B. cereus 569/H1, S. aureus PC1 and the RTEM strain. Class A enzymes catalyze the hydrolysis of penams better than cephems and have a molecular weight around 28,000 daltons. Members of class A are found to be serine enzymes (5,6). Class B are metalloenzymes from B. cereus and P. maltophilia (7). β-lactamase IIa from class B B. cereus seems to be unselective in its specificity of β-lactams and catalyzes a wide range of substrates (8,9). β-lactamase IIa also is not affected by the common β-lactamase inhibitors. The β-lactamase from P. maltophilia is the only other enzyme known to require metal ions for catalytic activity and it seems to be very different from the B. cereus variety (10). Class C β-lactamases are also serine enzymes (11) but unlike class A β-lactamases, class C β-lactamases hydrolyze cephems better than penams, have a molecular weight of approximately 38,000 daltons and include enzymes from P. aeruginosa and E. coli K12 (12).

A relatively high resolution crystal structure of a β-lactamase has recently been available for the PC1 class A β-lactamase from S. aureus (13) The molecule is

composed of two domains with the active site located at the junction (Figure 1). The first domain is formed by an antiparallel beta sheet of five strands and three alpha helices (α1, α11, α10). The second domain is composed of eight alpha helices. Two clefts are formed at the junction of the domains. One is the active site depression and the other is located on the right side of the structure in Figure 2 (13).

Class A β-lactamases such as S. aureus PC1 have a conserved triad of amino acids, Ser-Thr-Xaa-Lys in their sequence (1). This serine residue at position 70 (according to the numbering system of Ambler) is the residue involved in the formation of an acyl enzyme intermediate (via nucleophilic attack) with the carbonyl carbon of the beta-lactam ring (14,15). The side-chain of Ser-70 lies at the bottom of the crevice with the ammonium group of the conserved Lys-73 next to it (Figure 2)(13). Lys-73 forms a salt bridge with Glu-166, another conserved residue, lying on the floor of the crevice. Val-103 and Ile-239 form a narrow alley in the depression while the ammonium group of a second conserved lysine, Lys-239, lies at the closed end of the crevice. There are other conserved residues in the active site depression such as Asn-132 and Asn-170 which make up part of the walls of the active site (13).

Although much work has been done on the β-lactamases, their catalytic mechanism is still unknown.

Evidence points to an acyl enzyme intermediate analogous to the serine proteases (16). From the crystal structure, Herzberg and Moult postulated a mechanism (13) that includes: (i) nucleophilic attack of the hydroxyl group of Ser-70. This attack leads to a tetrahedral transition state followed by formation of the acyl enzyme intermediate. The tetrahedral transition state is stabilized by hydrogen bonding of the carbonyl oxygen with the amino groups of Ser-70 and Gln-237 analogous to the "oxyanion hole" of serine proteases (16). During the nucleophilic attack, the proton on the hydroxyl group of Ser-70 would transfer to the beta-lactam nitrogen. This transfer is facilitated by Lys-73. Lys-73 orients the serine proton, and polarizes the beta-lactam nitrogen, thereby, reducing the energy barrier of transferring the proton (13) (ii) hydrolysis of the acyl enzyme intermediate occurs. It is suggested that Glu-166 optimizes the nucleophilic attack of water in the deacylation step. Moreover, the crystal structure indicates that a pocket for a water molecule exists at the bottom of the active site depression.

The β-lactamase used in the author's work is the class A RTEM-1 enzyme encoded on the plasmid pBR322. Two different RTEM enzymes have been isolated though at one time they were thought to be the same, leading to some confusion. The RTEM-1 enzyme has an isoelectric point of 5.4 and the RTEM-2 enzyme an isoelectric point of 5.6

which is accounted for by a single amino acid difference in their sequences. The RTEM-1 has a glutamine residue at position 39 while the RTEM-2 enzyme has a lysine (17).

The roll of the conserved residues in the RTEM-1 beta-lactamases has been investigated using cassette and site-specific mutagenesis (18,19,20). These studies were conducted before the crystal structure of the S. aureus PC1 class A beta-lactamase was published and so a three-dimensional active site and the orientation of the conserved residues were not known. Nevertheless, the conclusions of the mutagenesis agree well with the insights available from the three-dimensional structure. The conserved residue Thr-71 was replaced with all other nineteen amino acids. All of the mutants except T71K, T71R, T71Y, T71N and T71D demonstrate a resistant phenotype to beta-lactam antibiotics (18). The kinetic properties of the
T71S (21) , and T71I (21) and T71C (20) mutants were studied and it appears that threonine at position 71 is importantly involved in stabilition of the protein and less directly involved in catalysis.

The conserved lysine at position 73 was also replaced with all other amino acids by cassette mutagenesis (20). In this case, however, none of the mutants conferred a resistant phenotype to beta-lactam antibiotics. Kinetic analysis of the T71C and T71R did

reveal a small (0.01%) amount of activity against benzylpenicillin as compared to wild type while the K_m was not significantly changed. The pH activity profile of the T71C mutant showed that activity maximized at pH 8.3 implying that the anionic character of the Cys increased activity. Interestingly, modifying the Cys to aminoethylcysteine regained 62% of the catalytic efficiency of the wild type enzyme implying that a correctly positioned positive ion at residue 73 is needed for catalysis.

This chapter describes the purification and kinetic characterization of a T71H mutant that was created earlier by S.Schultz (18). This mutant was chosen for characterization due to the difference between a secondary alcohol and an imidizole ring. It was also interesting to explore the role of a positive charge at this site to see if it would change its pH profile or its ability to hydrolyze various substrates. The mutated gene was transferred into pJN for expression. β-lactamase production is induced by the addition of IPTG. DEAE chromatography on the extruded proteins followed by Ultragel 54 column or FPLC was used in the purification procedure. The Michaelis-Menten kinetics on the T71H mutant showed lower k_{cat} values on benzylpenicillin, cephalothin and 6-aminopenicillanic acid but no changes in K_M except for cephalothin which is approximately 4 times higher. The mutant showed little changes in its pH

profile compared to the wild-type enzyme and the mutant is more readily thermally denatured than the wild-type enzyme though the probable generation of a positive charge at that site seemed to enhance stability.

MATERIALS and METHODS

Enzymes and Chemicals

PvuI and EcoRI were purchased from Boehringer Mannheim. DNA ligase was purchased from Bethesda Research Laboratories. Antibiotics were obtained from Sigma Chemical Co. Tryptone, yeast extract and agar were from Difco. All other chemicals used were of reagent grade.

DNA

Plasmid DNA was prepared by the alkaline lysis method (23). pBR322 was purchased from Bethesda Research Laboratories. DNA fragments were purified by electrophoresis on a 1.2% agarose gel and the DNA removed from the gel by use of a geneclean kit from Bio 101 Company (La Jolla, Ca) or the use of an electroluter.

Bacterial Strains

E. coli strain LS1 (22) was used as the host of the plasmid pBR322 containing either the wild-type or mutated β-lactamase gene. *E. coli* strain D1210 was used as the host cell when transforming plasmid pJN containing the mutated β-lactamase. *E. coli* strain D1210 is a lac I^q derivative of HB101.

Subcloning into pJN

Plasmid pBR322T71H, containing the T71H mutant and previously constructed by S. Schultz (18) was digested with EcoRI and PvuI restriction endonucleases and electrophoresed on a 1.2% agarose gel. The exact digestion was also done on the plasmid pJN and electrophoresed in an identical manner. The 4800bp fragment from the pJN digestion and the 626bp fragment from the pBR322T71H digestion were isolated from the agarose gel with the use of a geneclean kit.

The two fragments wre ligated together using approximately a stoichiometric ratio of insert to vector. The ligation mixture contained approximately 0.04pmoles of each fragment in 10mM $MgCl_2$, 50mM Tris-HCl, pH 8.0, 0.5mM ATP, 5mM dithiothreitol with 20 units of DNA ligase. The ligation mixture was incubated at 16 °C overnight for approximately 18 hours. E. coli strain D1210 was made competent using standard procedures (24) and 10ul aliquot of the ligation mixture was used to transform the E. coli D1210 strain. The pJN vector carries a kanamycin resistance marker and this was used to select transformed cells carrying the ligated vector.

To insure that the mutated β-lactamase gene was present on the pJN vector, the plasmid from a kanamycin resistant colony was purified and digested with AvaI

restriction endonuclease and electrophoresed on a 1.2% agarose gel. The pJN vector containing the mutated β-lactamase gene will cut three times with AvaI but only twice if the gene is not present.

Purification of T71H

The purification procedure used was a variation of the published procedure (19).

E. coli strain D1210 containing the mutated β-lactamase gene encoded on plasmid pJNT71H was used to inoculate 50ml of L broth (10g tryptone, 5g yeast extract, 10g NaCl per liter) also containing 20mg/ml of kanamycin. The culture was grown to saturation overnight at 37 °C with continuous shaking. The saturated culture was used to inoculate 300ml of L broth also containing 20mg/ml kanamycin. The culture was grown at 37 °C with shaking until it reached log phase. All 300ml of this culture was used to inoculate 10 liters of FB media (25g tryptone, 7.5g yeast extract, 50ml of 1M Tris, pH 7.5, 5g NaCl per liter) containing 20mg/ml kanamycin. The 10 liter culture was grown in a New Brunswick Scientific Co fermentor at 30 °C with vigorous stirring and aeration for 12 to 14 hours. One drop of antifoam from Sigma Co. was added to control the foaming. After the required incubation, IPTG was added to 0.1mM. IPTG was allowed to induce for 45 minutes at 30 °C with continuous stirring and aeration. The cells were then collected in

centrifuge tubes cooled on ice. The cells were pelleted by centrifugation and the osmotic extrusion procedure was performed on the pellet as published (25). The supernatant containing the extruded proteins was concentrated to approximately 30mls using a Diaflo fitted with a PM-10 membrane. The buffer in the Diaflo was changed to 25mM triethanolamine, pH 7.25 and the 30mls of protein was applied to a 2.5 x 25cm DE-52 column. The column was run with a linear concentration gradient of 25mM to 200mM triethanolamine, pH 7.25. Approximately 4 ml fractions were collected and the mutated β-lactamase was located by assaying for its activity . Fractions containing activity were pooled and 3/4 of the sample was changed into 20mM Tris, pH 7.0 buffer and eluted on a 1 x 82cm Ultragel ACA 54 column using the same Tris buffer. Fractions of 1ml were collected and assayed for activity. A 15cm 12% polyacrylamide gel with a 2cm 3.5% stacking gel was run on samples off the Ultragel 54 column to check its purity. The other 1/4 of the sample off the DE-52 column was purified using the FPLC. The samples in 25mM triethanolamine, pH 7.25 were loaded on a Pharmacia Mono Q 5/5 column using the following discontinuouse gradient. Buffer A contained 25mM triethanolamine, pH 7.25. Buffer B contained 200mM triethanolamine, pH 7.25. At time 0, 100 % A; time 15 minutes, 85% A and 15% B; time 40 minutes, 70% A and 30% B; and at time 60

minutes, 100% A. Flow rates were run at 0.8ml/minute. Fractions were collected at the rate of 1ml/minute. Overall yield of the mutant β-lactamase was approximately 8mgs/10 liters.

Assays

Protein concentrations were measured using the published conversion factor , 29,400 $M^{-1}cm^{-1}$ at 281nm (25). The activity assay was performed at 30 C in 0.1M potassium phosphate, pH 7.0, using a Beckman DU-7 spectrophotometer with a 1cm pathlength quartz cell for assays of activity on penams. However, for cephalothin, an optical quartz block was inserted into the cell to reduce the pathlength to 1mm. Hydrolysis of the β-lactam ring can be followed for benzylpenicillin and 6-aminopenicillanic acid by measuring the decrease in absorbance at 240nm, and at 265nm for cephalothin (25). The $\Delta\varepsilon$ values used in the assay conversion are $7900M^{-1}cm^{-1}$ for cephalothin, 500 $M^{-1}cm^{-1}$ for benzylpenicillin and 500 $M^{-1}cm^{-1}$ for 6-aminopenicillanic acid (25).

Activity vs. pH

The pH of the 0.1M phosphate assay buffer was varied from pH 5 to pH 9. Aliquots of the mutant or wild-type β-lactamase were added directly and assayed. The assay is performed in exactly the same manner described

except for the change in pH of the assay buffer. The pH of the cuvette containing the assay mixture was also measured directly to insure the correct pH.

Thermal stability

Mutant β-lactamase was incubated at the indicated temperature in 0.1M phosphate buffer, pH 7.0. At various times of incubation, aliquots were removed and assayed for activity on benzylpenicillin at 30 °C.

Thermal stability vs. pH

Mutant β-lactamase was aliquoted in the indicated pH buffer (0.1M phosphate) and incubated (2-3 minutes) at various temperatures. Samples were immediately assayed for activity on benzylpenicillin at 30 °C

RESULTS

The plasmid pBR322T71H containing the T71H mutant β-lactamase was obtained from S. Schultz (18). In order to express the mutant protein in large quantities, the mutant β-lactamase gene was transferred into the plasmid pJN which contains the tac promoter (26,27) and the kanamycin resistance gene. The ligation strategy for moving the T71H mutant β-lactamase into the pJN vector is shown in Figure 3. Both pJN vector and pBR322T71H are digested with EcoRI and PvuI and the mutant β-lactamase gene takes the place of the EcoRI-PvuI fragment in pJN. The presence of the mutated beta-lactamase gene in pJN is confirmed by its kanamycin resistance and by AvaI cutting three times in the plasmid . An extra AvaI site was inserted into pBR322T71H near the active site serine 70 for cassette mutagenesis purposes (18). The host strain used to express pJNT71H was E. coli D1210 because it expresses the lac repressor constitutively. Therefore, the production of beta-lactamase is shut off until the addition of IPTG. The induction was allowed to continue for 45 minutes at 30 °C. The lower temperature was used to minimize the loss of active enzyme since other Thr-71 mutants showed thermal instability. The overall yield of mutant protein was 8mgs/10 liters which was the highest yield of any T71 mutant β-lactamase

purified in our lab thus far. Yields of other mutant β-lactamases at Thr-71 were generally around 1-3mgs/10liters (20,21). The elution profile from the DE-52 column is shown in Figure 4 and the elution profile from the Ultragel 54 column is shown in Figure 5. A 12% SDS-PAGE gel was run on fractions from the Ultragel 54 column. Later fractions from the Ultragel 54 column were shown to be homogenous as seen in Figure 6. The homogenous fractions were used for kinetic studies.

The use of the FPLC in purifying the mutant β-lactamase proved tremendously simple, quick and better than use of the Ultragel 54 column. Figure 7 shows an FPLC elution chromatogram for the T71H mutant after the DE-52 column purification step. The peak found to contain activity against benzylpenicillin is indicated by an arrow. Fractions representing this peak were collected and pooled. The pooled fractions were tested again on the FPLC and Figure 8 demonstrates the purity of the T71H mutant β-lactamase. In comparison, an FPLC was done on the fractions off the Ultragel 54 column representing supposedly pure protein. Figure 9 shows that an extra protein that co-elutes on the Ultragel column can be separated by FPLC. The difference in elution times from the FPLC between wild-type and the T71H mutant was usually 5-8 minutes with the mutant β-lactamase coming off sooner. Since the Mono Q column is composed of positively charged groups, one would expect

the mutant β-lactamase to come off sooner if the histidine residue still had some positive character at pH 7.25 at which pH the FPLC is run.

Michaelis-Menten kinetics were performed on the T71H mutant β-lactamase and the parameters are shown in Table 1. The parameters were determined for three substrates, benzylpenicillin, 6-aminopenicillanic acid and cephalothin. The values for k_{cat} and K_M were determined from Lineweaver-Burk (28) plots (done 4 times each for every substrate). The plot was constructed by following the decrease in absorbance curve and using [S], the average substrate concentration and v, the average reaction rate of the decrease in absorbance curve in accordance with published procedure (29). The K_M values were unchanged for benzylpenicillin and 6-aminopenicillanic acid but increased about 4-fold for cephalothin. The k_{cat} values decreased for all substrates. The k_{cat} decreased 7-fold for benzylpenicillin, 19-fold for 6-aminopenicillanic acid and 4-fold for cephalothin. The values of k_{cat}/K_M which is a measure of the catalytic efficiency of the enzyme were 13.5% for benzylpenicillin, 6.6% for 6-aminopenicillanic acid and 5.4% for cephalothin as compared to the k_{cat}/K_M for wild-type beta-lactamase.

The K_M and k_{cat} values were also determined at various pH values ranging from 5 to 8.5. Each of the three substrates were tested and in each case, the

T71H mutant β-lactamase mimicked the wild-type β-lactamase in its parameters (Figures 10 through 15).

The thermal stability of the mutant is greatly reduced above temperatures of 45 °C as seen in Figure 16 and Table 2 as compared to the wild-type β-lactamase. At pH 7, Figure 16 shows the rapid loss of activity of the mutant above 45 °C. However, at pH 5, the mutant does not exhibit such rapid inactivation but follows a similiar activity pattern as the wild-type enzyme (Figure 17). In contrast, at pH 8, the mutant demonstrates a much more sensitive thermal denaturation than the wild-type enzyme. At pH 8, the mutant begins to lose activity above 35 °C whereas the wild-type enzyme is stable to 45 °C (Figure 18).

DISCUSSION

The T71H mutant was constructed by S.Schultz using cassette mutagenesis which proved to be a rapid and excellent method for producing a large number of mutants at a particular site (18) They found that the histidine mutant possessed a high degree of resistance to the penam and cephem antibiotics. Through the use of western blotting, the expression and stability of the mutant enzyme at 37 °C seemed to be one of the best among the Thr-71 mutants (18). It was this reason along with the notion that a possible positive charge at the active site might provide some interesting insights into the binding of substrate and catalysis, that this mutant was chosen to be characterized.

In the T71H mutant's purification process, the mutant behaved very much like the wild-type enzyme. Except for keeping the temperature at 30 °C during growth and keeping the temperature low during the purification, no other procedure used was significantly different to isolate the mutant enzyme from bacterial cells as compared to the wild-type β-lactamase. High yields of the mutant enzyme were also obtained, about 2-3 times more than any other threonine mutant. Therefore, the histidine does not prevent or severly hamper the correct processing and secretion of the mutant β-lactamase

which can survive in the periplasm of the cell.

Using Michaelis-Menten kinetics, the K_M was found not to significantly change when the substrates benzylpenicillin and 6-aminopenicillanic acid are hydrolyzed by the mutant in comparison to the wild-type analogue. This would indicate that the enzyme-substrate complex for both benzylpenicillin and 6-aminopenicillanic acid maintained the same level of stability. However, in studying the mutant's hydrolysis of cephalothin, the K_M values increased 4-fold relative to that of the wild-type enzyme. In observing the crystal structure of the active site of β-lactamase (Figure 2), one can see the explanation for the increased K_M. The active site of β-lactamase can be described as a cave-like depression and threonine 71 in the wild-type enzyme is positioned in the back of the pocket. Therefore, the T71H mutant could alter significantly the binding pocket by making the pocket smaller, hence, creating greater steric hindrance with the acetate group of cephalothin (Figure 19). Greater steric hindrance in the binding pocket would reduce the stability of the caphalothin-enzyme complex which results in an increased K_M. An increase in K_M using cephalothin as the substrate was also observed in a T71C mutant (20) but not in a T71S mutant (30).

The T71H mutant shows considerable catalytic activity on both penams and cephems as seen in Table 1.

The mutant continues to show better efficiency on penams than on cephems as in the wild-type enzyme. The k_{cat} values of the T71H mutant also are analogous to the other threonine mutants characterized (22,23) Therefore, one can infer that the role of threonine 71 in the wild-type enzyme is not directly important in catalysis. Another interesting observation is that the mutant has a stronger effect on the K_{cat}/K_M value against cephalothin than on the K_{cat}/K_M value against penicillin G. This effect is even more pronounced in the mutant in comparing the K_{cat}/K_M value against 6-aminopenicillanic acid with the K_{cat}/K_M value against penicillin G. This effect is probably due to cephalothin and 6-aminopenicillanic acid being poorer substrates than penicillin G for β-lactamase. The binding orientation of cephalothin and 6-aminopenicillanic acid is not optimal for catalysis in the class A β-lactamases and, therefore, it would be even less so for a mutant with a large histidine residue at position 71.

The generation of a positive charge at position 71 at pH 5 did not effect the K_M or k_{cat} values of the mutant in comparison to the wild-type. For all three substrates tested, the k_m and k_{cat} values mimicked the wild-type and were relatively constant over the pH range of 5 to 8. If one assumes that the positive charge at position 71 influenced catalysis or binding, one would surely see changes of the K_{cat} or K_m over this pH range.

Since this is not the case, it seems conclusive that threonine at position 71 is not involved significantly in binding or catalysis. This conclusion is in agreement with other threonine mutants of β-lactamase (19,20,21).

The T71H mutant is more readily thermally denatured than the wild-type enzyme. As seen in Table 2, the mutant incubated at pH 7 rapidly loses activity at temperatures above 40 °C while the wild-type remains stable at temperatures up to 50 °C. An interesting observation is that the thermal stability of the mutant enzyme is pH dependent. The generation of a positive charge at pH 5 thermally stabilized the mutant as compared to the thermal stability at pH 8. As seen in Figure 14, the mutant incubated at pH 5 is comparable in thermal stability as the wild-type enzyme. However, at pH 7 (Figure 17), the mutant rapidly loses activity after 45 °C while the wild-type is stable until 55 °C. At pH 8 (Figure 18), the mutant rapidly loses activity after 35 °C while the wild-type remains stable until 45 °C. The mutant does not have greater sensitivity to alkaline pH as compared to the wild-type enzyme and , therefore, the decrease in activity is not due to pH denaturation. It seems that the charge on the histidine residue plays a role in the mutant's thermal sensitivity. An explanation for this charge dependency is not clear. Though the role of threonine 71 in β-lactamase is now conclusively shown to be of structural

integrity and stability and less directly involved in catalysis, a positive charge at position 71 could, perhaps increase stability by formation of a salt-bridge.

SUMMARY

The T71H mutant possesses some interesting properties. The mutant retains significant levels of activity against penams and cephems as compared to the wild-type enzyme. The k_{cat}/K_M is 13.5% on benzylpenicillin, 6.6% on 6-aminopenicillanic acid and 5.4% on cephalothin as compared to the wild-type enzyme. The K_M values for benzylpenicillin and 6-aminopenicillanic acid remains unchanged in the mutant but is much higher on cephalothin. The k_{cat} and K_M is not influenced by the positive charge on histidine at pH 5, indicating that threonine at position 71 is not involved directly in binding or catalysis. However, the positive charge generated on histidine influences the thermal stability of the mutant, which is more thermally stable at lower pH values. In conclusion, the role of threonine at position 71 in β-lactamase is one of stabilty and structural integrity and less of binding or catalysis.

REFERENCES

1. D.J. Waxman and J.L. Strominger (1983) Ann. Rev. Biochem.,52, 825-869.
2. J.A. Kelley, J.R. Knox, P.C. Moews, G.J. Hite, J.B. Bartolone, H. Zhao, B. Joris, J.-M. Frere and J.-M.Ghuysen (1985) J.Biol. Chem., 260, 6449-6458.
3. N. Datta and P. Kontomichalou (1965) Nature (London),208, 239-241.
4. M.H. Richmond and R.B. Sykes (1973) Adv. Microb. Physiol,9,31-88.
5. V. Knott-Hunziker, S.G. Waley, B.S.Orlek and S.G. Sammes (1979) FEBS Lett.,99 , 59-61.
6. V. Knott-Hunziker, K. Redhead, S. Petursson and S.G. Waley (1980) FEBS Lett.,121 , 8-10.
7. Y. Saino, F. Kobayashi, M. Inoue and S. Mitsuhashi (1982) Antimicrob. Agents Chemother.,22, 564-570.
8. R.B. Davies and E.P. Abraham (1974) Biochem J.,143 ,129-135.
9. E.P. Abraham and S.G. Waley (1979) in "The Beta-Lactamases" (Hamilton-Miller and J.J. Smith.,eds.) pp311-338, Academic Press, London.
10. R. Bicknell and S.G. Waley (1985) Biochemistry, 24, 6876-6887.
12. A. Coulson (1985) Biotechnol. Genet. Engng. Rev.,

3, 219-253.

12. B. Jaurin and T. Grundstrom (1981) Proc. Natl. Acad. Sci. USA, 78, 4897-4901.
13. O. Herzberg and J. Moult (1987) Science, 236 , 694-701.
14. S.J. Cartwright and A.L. Fink (1982) FEBS Lett. ,137, 186-188.
15. S.A. Cohen and R.F.Pratt (1980) Biochemistry,19, 3996-4003.
16. J. Kraut (1977) Annu. Rev. Biochem., 46, 331-358.
17. J.G. Sutcliffe (1978) Proc. Natl. Acad. Sci. USA,75, 3737-3741.
18. S.C. Schultz and J.H. Richards (1986) Proc. Natl. Acad. Sci. USA, 83, 1588-1592.
19. G.D.-McFarland, J.J.Neitzel and J.H. Richards (1986) Biochemistry, 25, 332-338.
20. S. Carroll (1987) Ph.D. Thesis, California Institute of Technology.
21. Y.-H.Chang (1986) Ph.D.Thesis, California Institute of Technology.
22. H.W. Boyer and D. Roulland-Dussoix (1969) J. Mol. Biol. ,41, 459-472.
23. D. Ish-Horowicz and J.F. Burke (1981) Nucleic Acids Res., 9, 2989-2998.
24. D. Hanahan (1983) J. Mol. Biol. , 166, 557-580.
25. J. Fisher, S. Khosla and J. Knowles (1980) Biochemistry, 19 ,2895-2901.

26. E. Amann, J. Brosius and M. Ptashne (1983) Gene ,25, 167-178.

27. H.A. Boer, L.J. Comstock and M. Vasser (1983) Proc. Natl. Acad. Sci. USA, 80, 21-25.

28. H. Lineweaver and D. Burk (1934) J. Am. Chem. Soc., 56, 658-666.

29. H.-J. Lee and I.B. Wilson (1971) Biochim. Biophys. Acta, 242, 519-522.

30. J.Neitzel (1987) Ph.D. Thesis, California Institute of Technology.

FIGURE 1

The class A β-lactamase fold. (A) Ribbon representation of the polypeptide chain of β-lactamase from S. aureus PC1. Helices are labeled and numbered sequentially through the amino acid sequence, strands are labeled . The position for the active site Ser 70 is shown. (B) Stereo representation showing the α-carbon positions in the molecule (13).

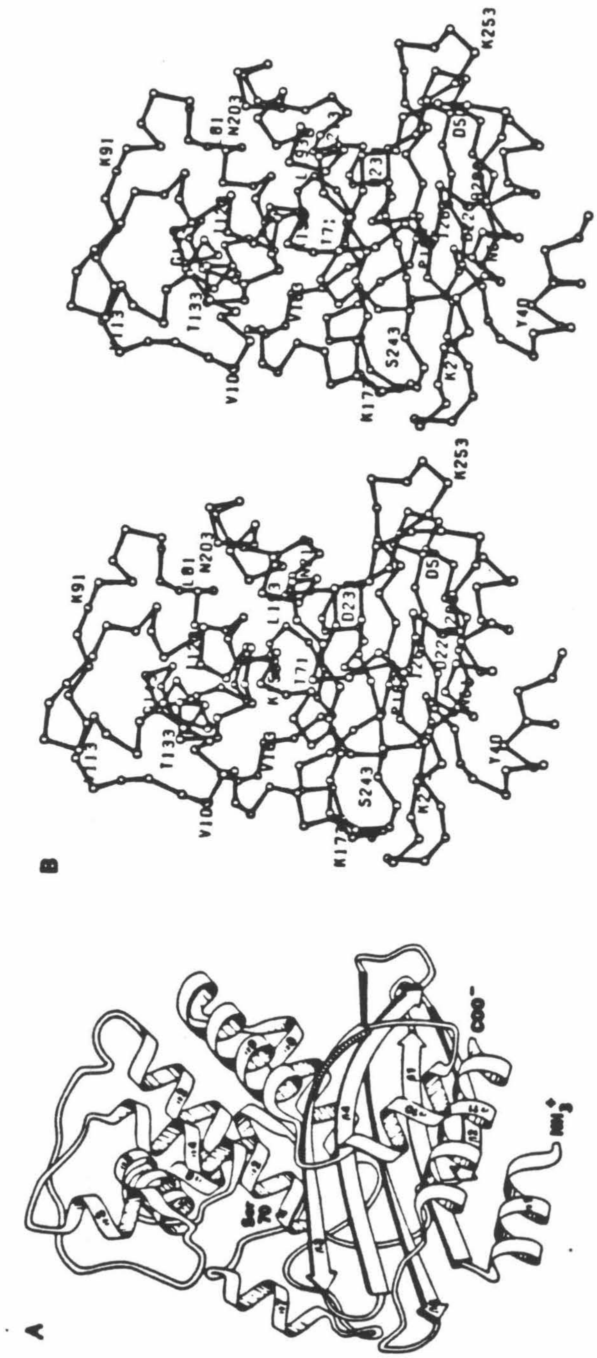

A
B
K253
K91
N203
T133
S243
D23
COO⁻

FIGURE 2

(B) Stereoscopic view of the residues that make up the active site depression. (C) Proposed position of a bound ampicillin molecule. (D) The proposed conformation of ampicillin when bound to β-lactamase (13).

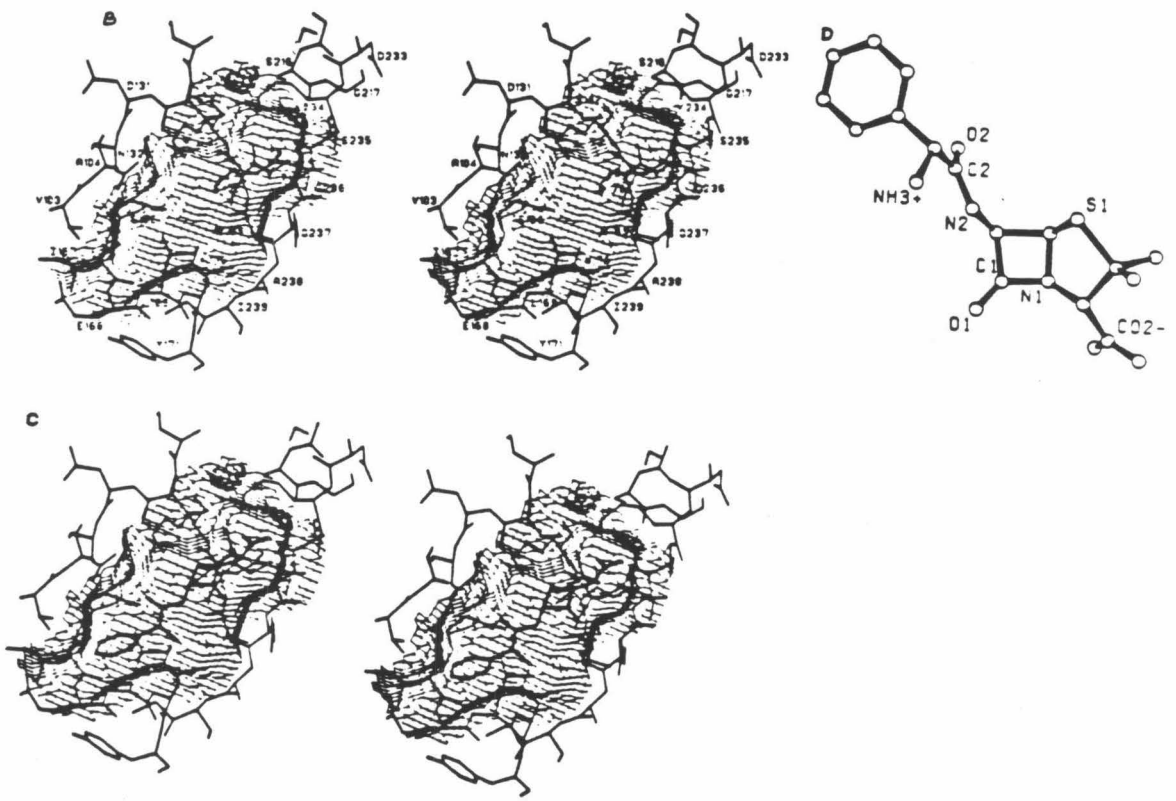

B
C
D
O2
C2
NH3+
N2
S1
C1
N1
O1
CO2-

FIGURE 3

Ligation strategy for the construction of plasmid pJNT71H.

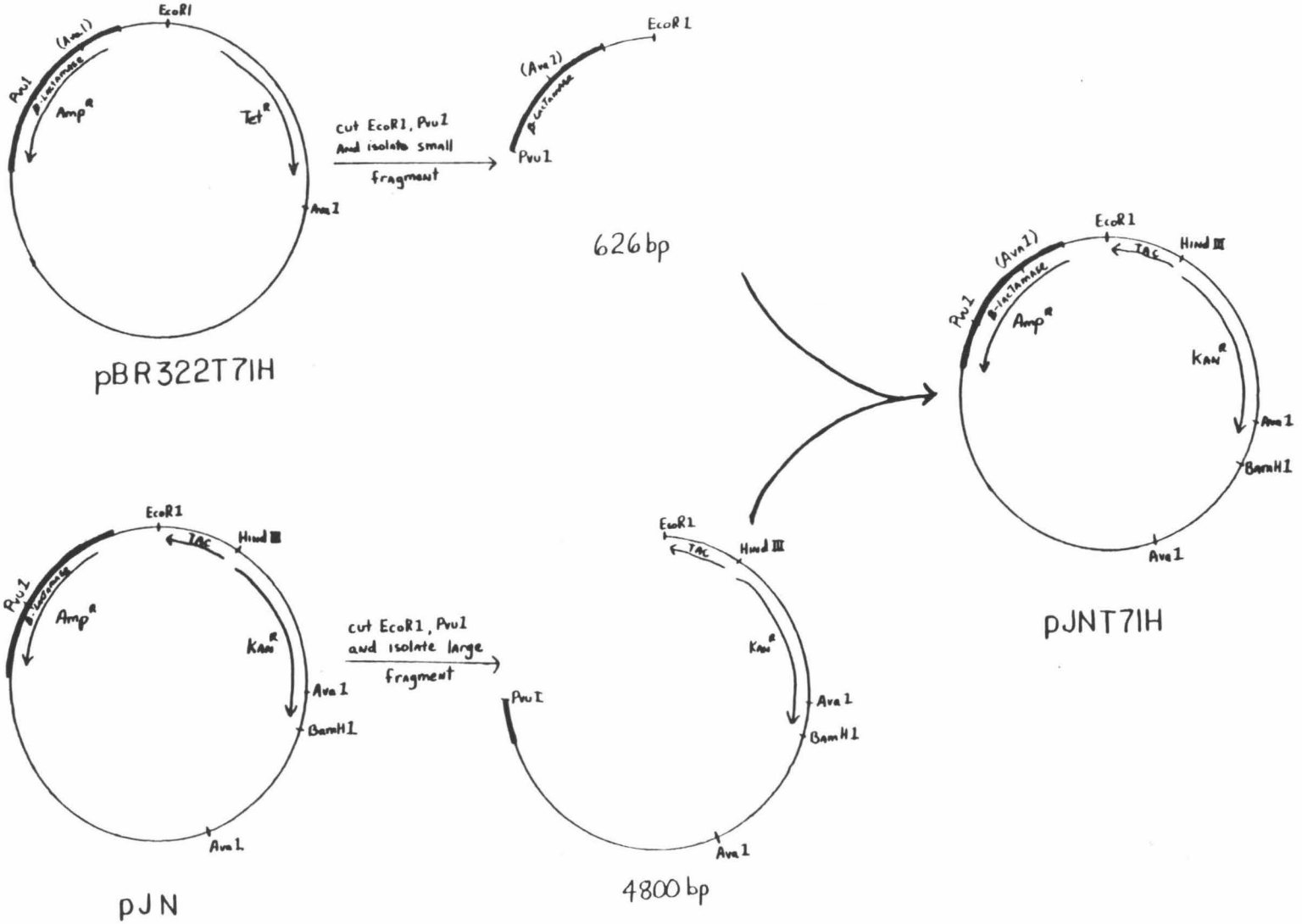

EcoR1
(Ava1)
Pvu1
β-Lactamase
Amp^R
Tet^R
Ava1
pBR322T7IH
cut EcoR1, Pvu2
And isolate small
fragment
EcoR1
(Ava1)
β-Lactamase
Pvu1
626 bp
EcoR1
Hind III
(Ava1)
TAC
Pvu1
β-Lactamase
Amp^R
Kan^R
Ava1
BamH1
Ava1
pJNT7IH
EcoR1
Hind III
Pvu1
β-Lactamase
Amp^R
TAC
Kan^R
Ava1
BamH1
Ava1
pJN
cut EcoR1, Pvu1
and isolate large
fragment
EcoR1
Hind III
TAC
Kan^R
Pvu1
Ava1
BamH1
Ava1
4800 bp

FIGURE 4

DE-52 column profile of the purification of T71H

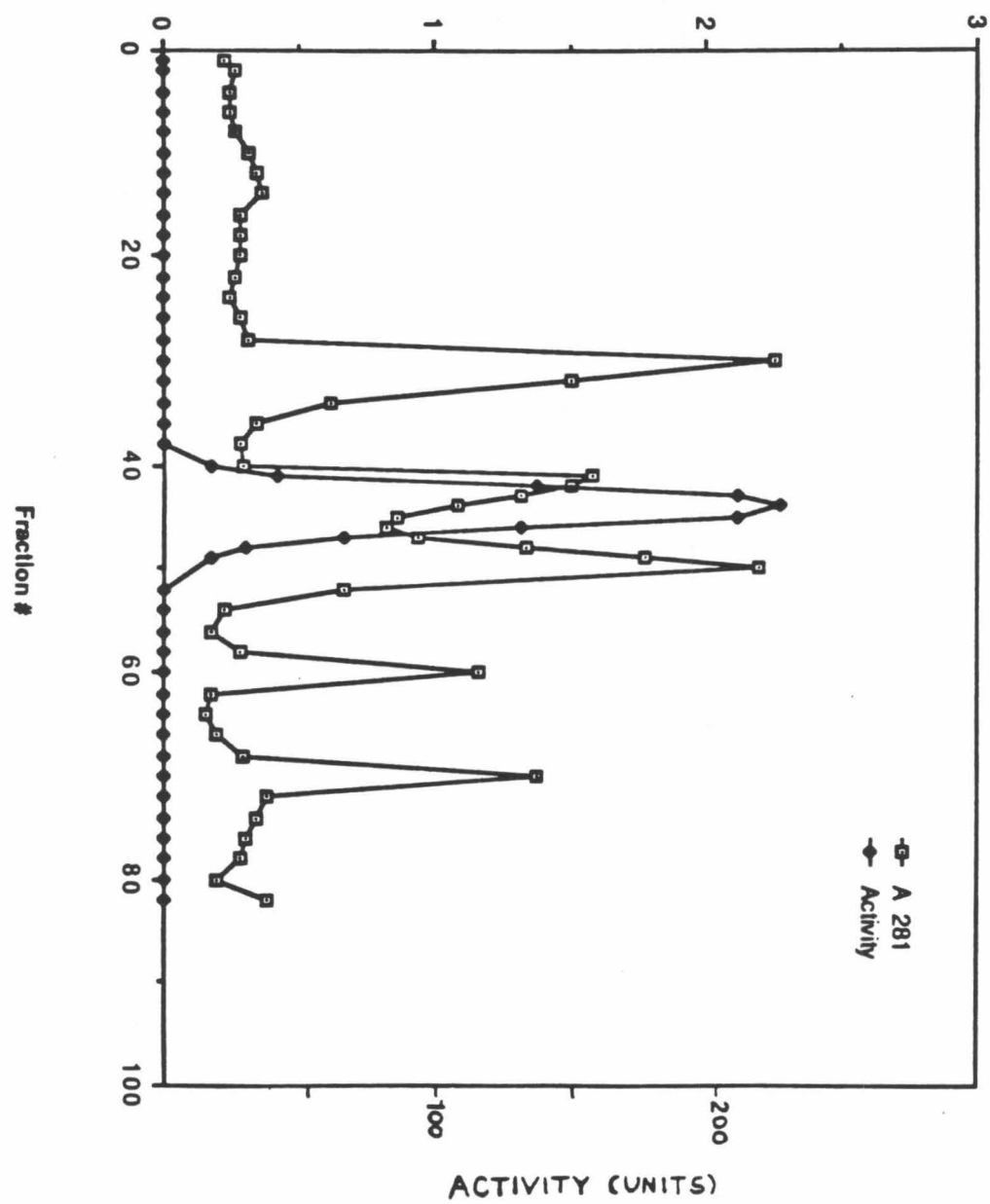

A 281
Activity
Fraction #
ACTIVITY (UNITS)
0
20
40
60
80
100
1
2
3
100
200

FIGURE 5

Ultragel 54 column profile of the purification of T71H.

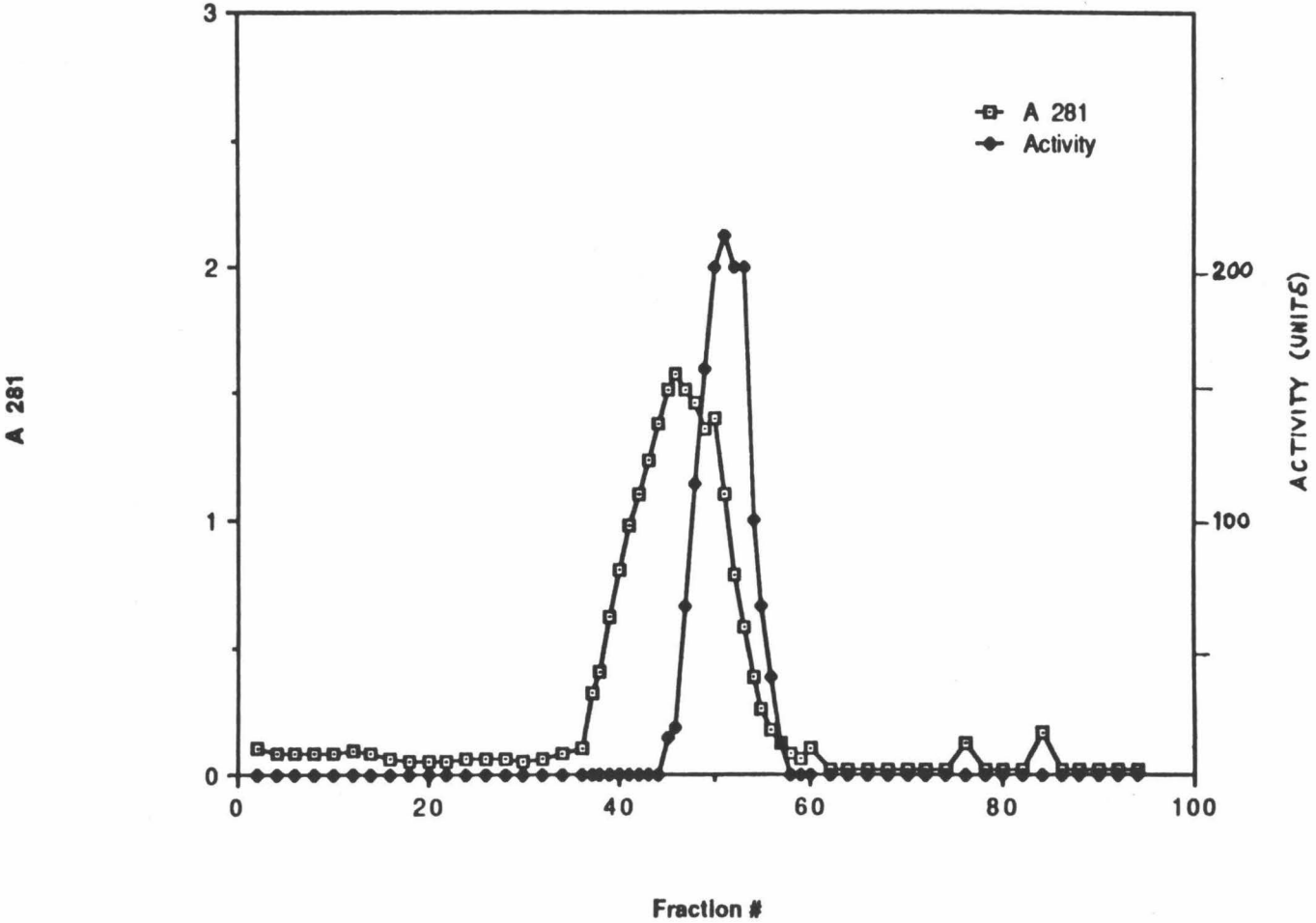

A 281
Activity
A 281
ACTIVITY (UNITS)
Fraction #
0
1
2
3
100
200
0
20
40
60
80
100

FIGURE 6

12% acrylamide gel of selected fractions from the Ultragel 54 column profile. Lane 1 is wild-type β-lactamase. Lane 2 through lane 10 are 5 ul aliquots from fractions #45, #46, #47, #48, #50, #52, #54, #56 and #57 repectively. Lane 11 is wild-type β-lactamase. Lanes 12 through 20 are 10ul aliquots of the same fractions in lanes 2 through 10.

LANES

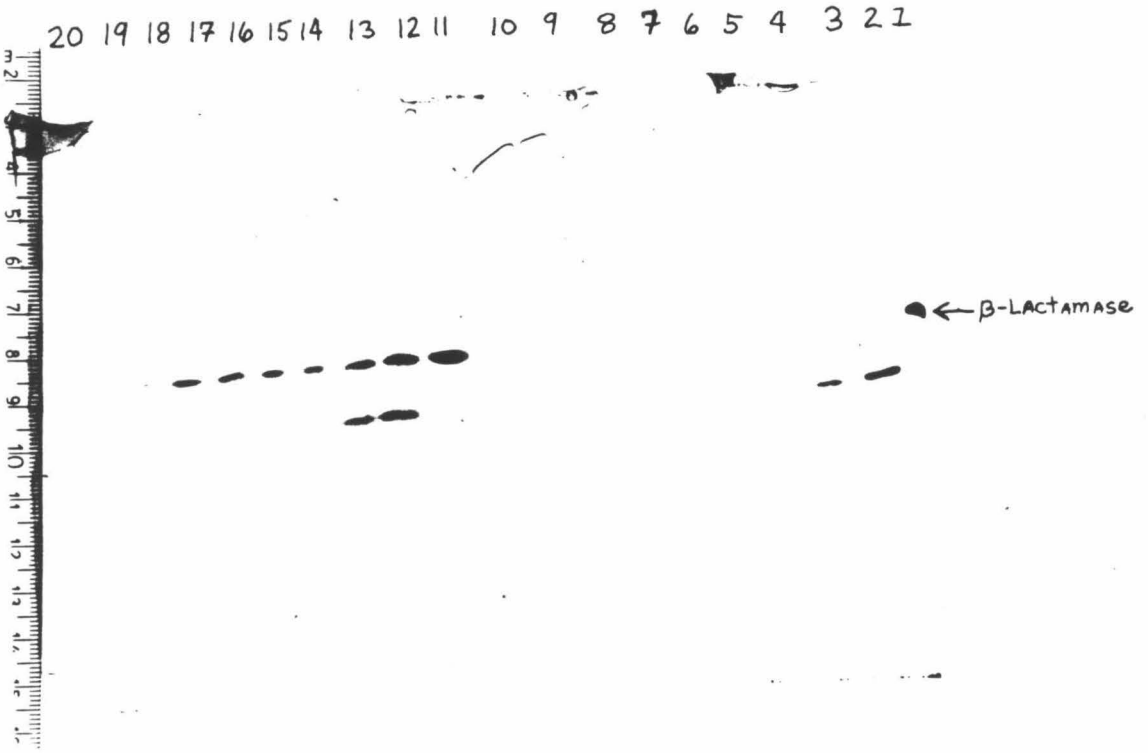


FIGURE 7

FPLC elution chromatogram for the T71H mutant after the DE-52 column purification step. The arrow indicates the peak found to contain activity.

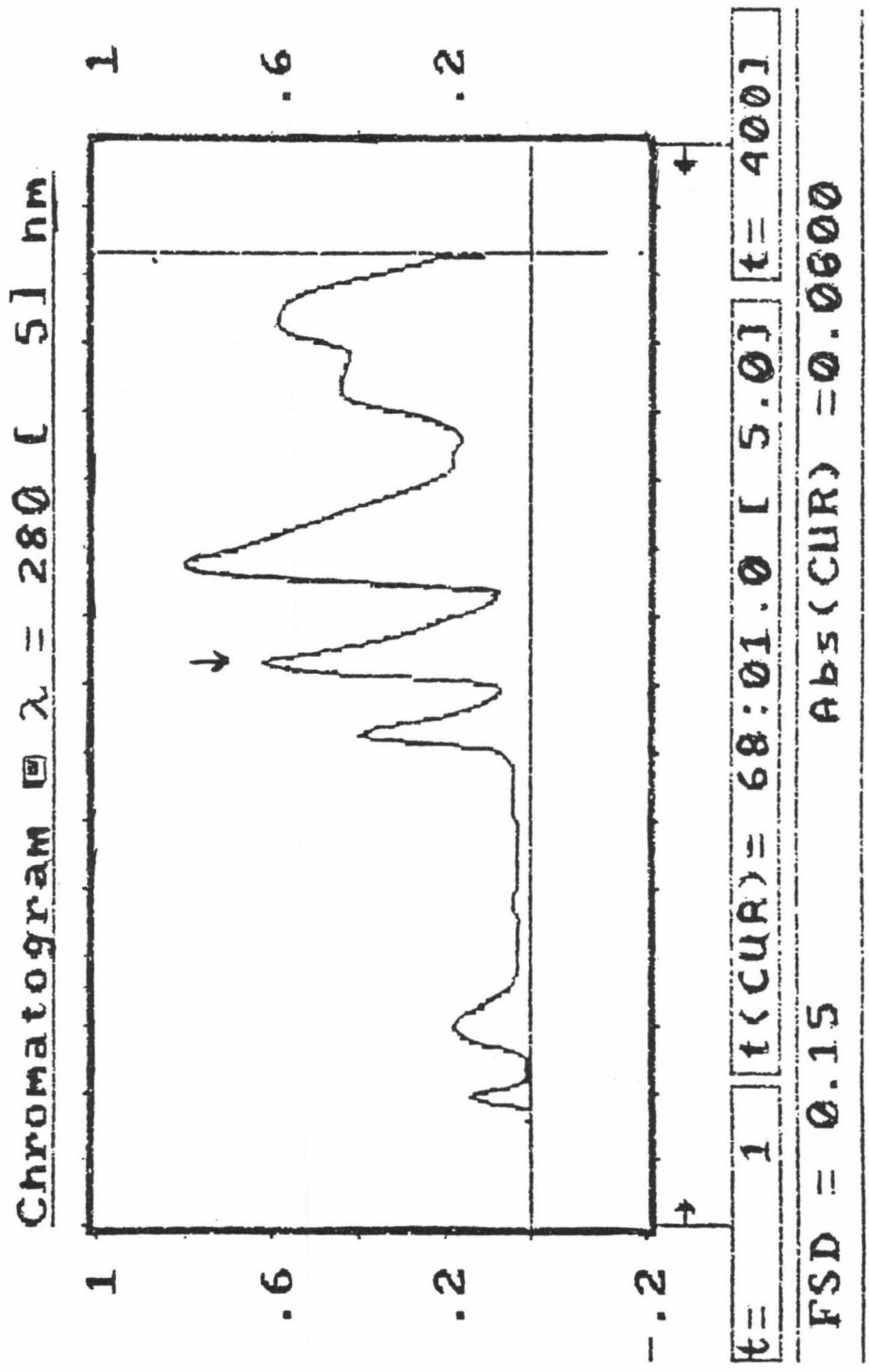

Chromatogram @ λ = 280 [5] nm
1
.6
.2
-.2
t= 1
t(CUR)= 68:01.0 [5.0]
t= 4001
FSD = 0.15
Abs(CUR) =0.0600

FIGURE 8

FPLC elution chromatogram of pooled fractions of the peak containing activity in figure 7. This is pure T71H mutant β-lactamase.

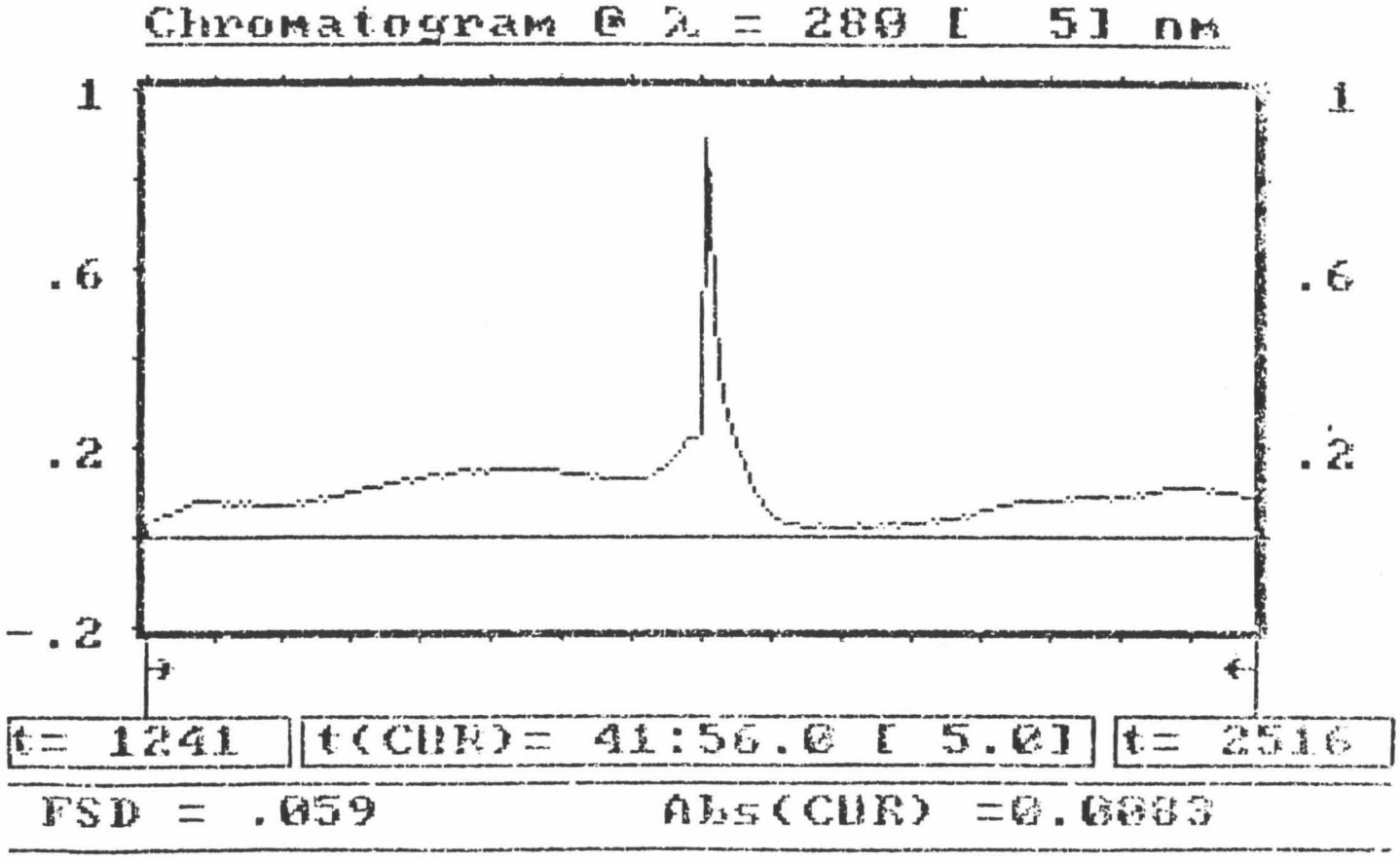

Chromatogram @ λ = 280 [5] nm
1
.6
.2
-.2
1
.6
.2
t= 1241
t(CUR)= 41:56.0 [5.0]
t= 2516
FSD = .059
Abs(CUR) =0.0083

FIGURE 9

FPLC elution chromatogram of a supposedly pure protein fraction from the Ultragel 54 column.

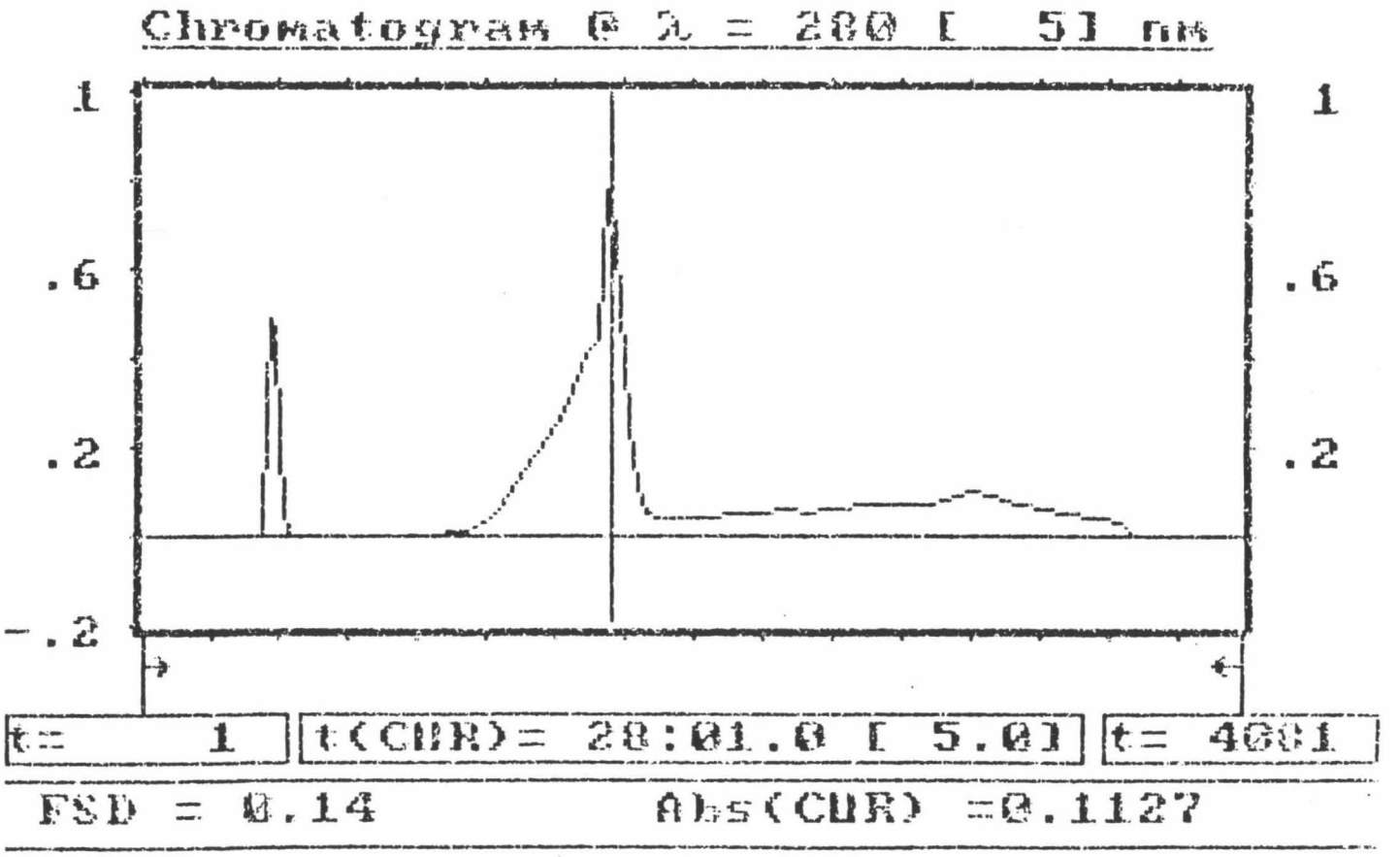

Chromatogram @ λ = 280 [51 nm
1
.6
.2
-.2
1
.6
.2
t= 1
t(CUR)= 28:01.0 [5.0]
t= 4001
FSD = 0.14
Abs(CUR) =0.1127

FIGURE 10

K_M versus pH for the T71H mutant β-lactamase. The substrate used is 6-aminopenicillanic acid.

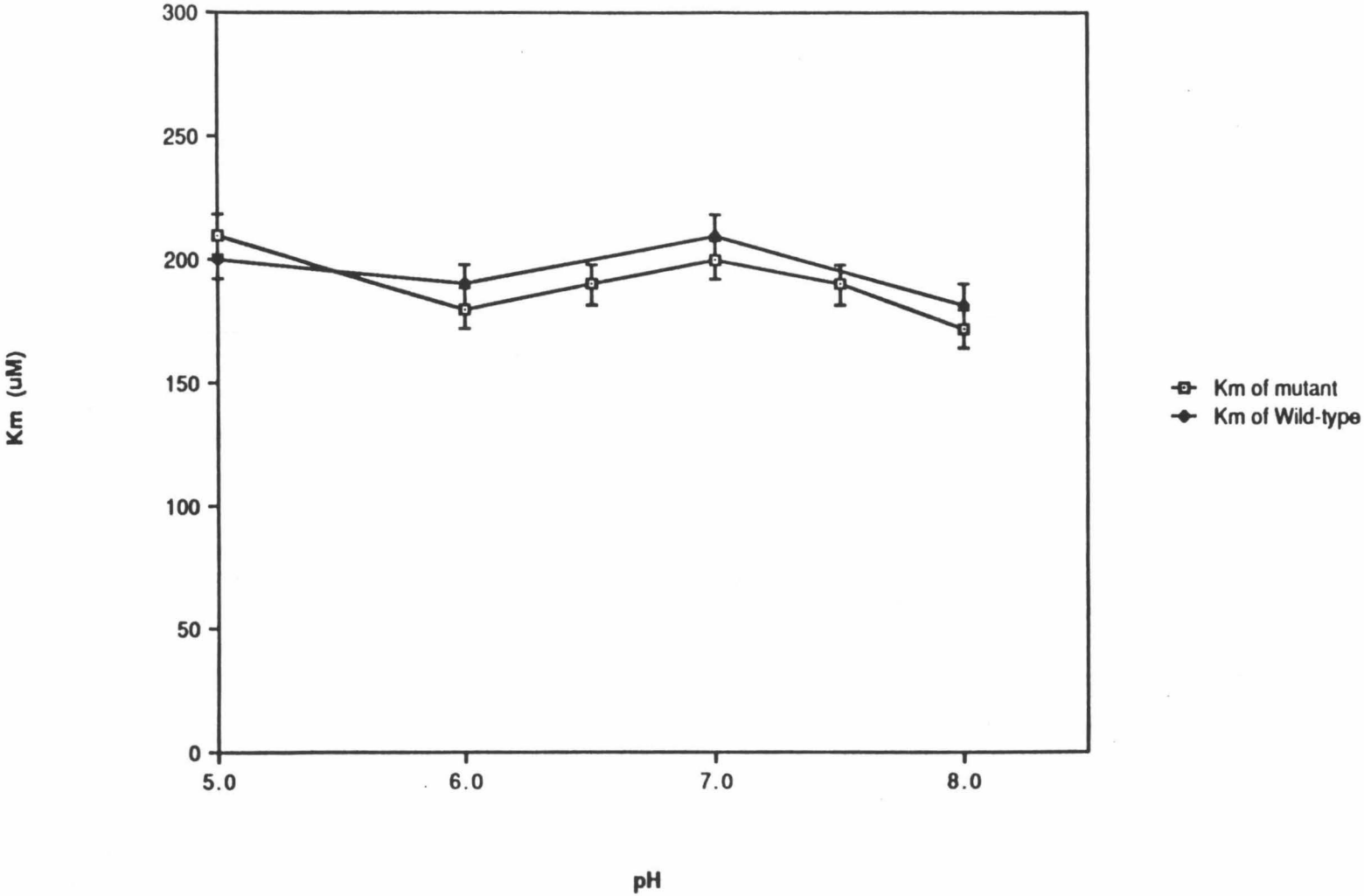

300
250
200
150
100
50
0
Km (uM)
5.0
6.0
7.0
8.0
pH
Km of mutant
Km of Wild-type

FIGURE 11

K_M versus pH for the T71H mutant β-lactamase. The subtrate used is benzylpenicillin.

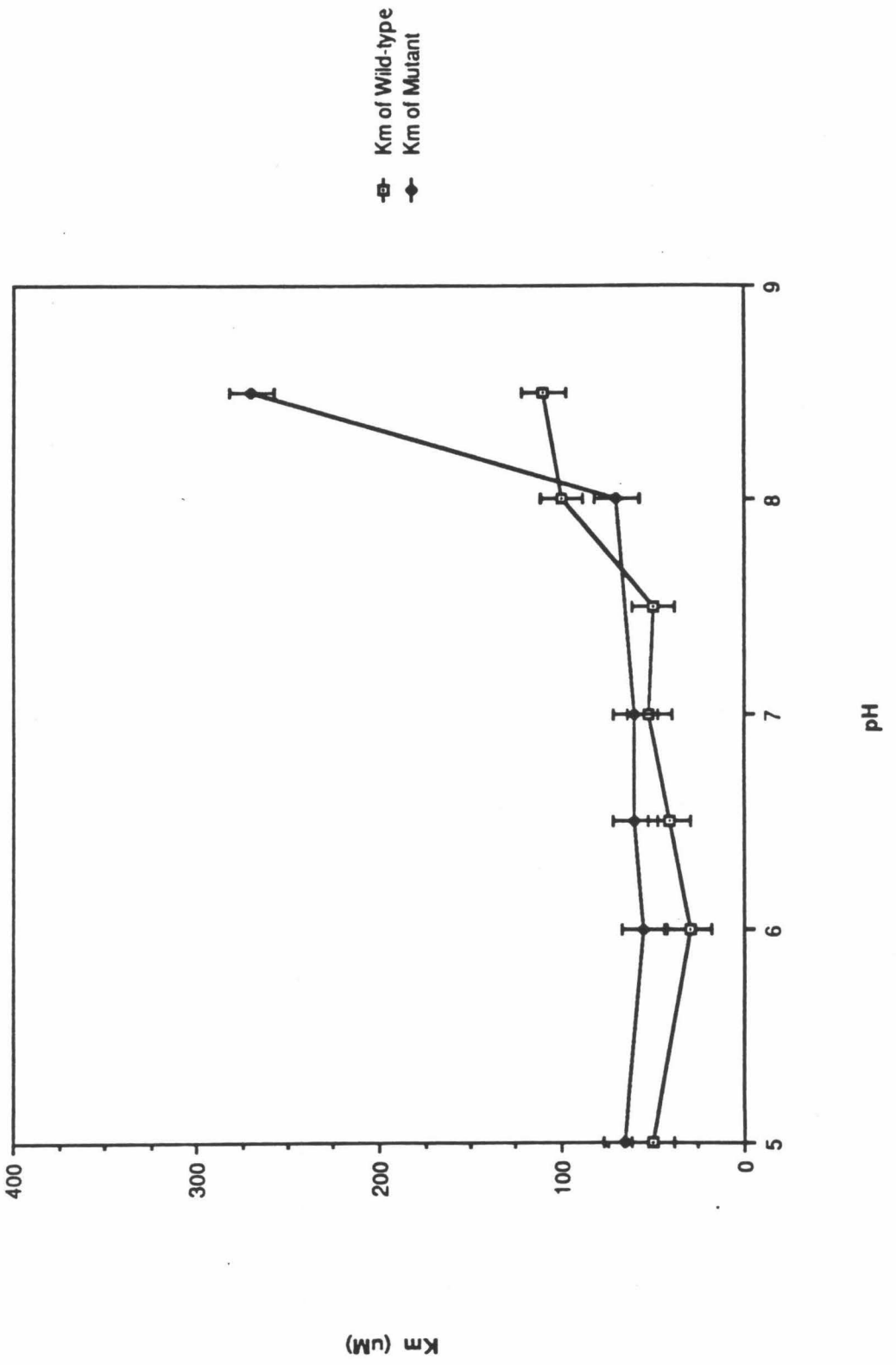

Km of Wild-type
Km of Mutant
Km (uM)
400
300
200
100
0
5
6
7
8
9
pH

FIGURE 12

K_M versus pH for the T71H mutant β-lactamase. The substrate used is cephalothin.

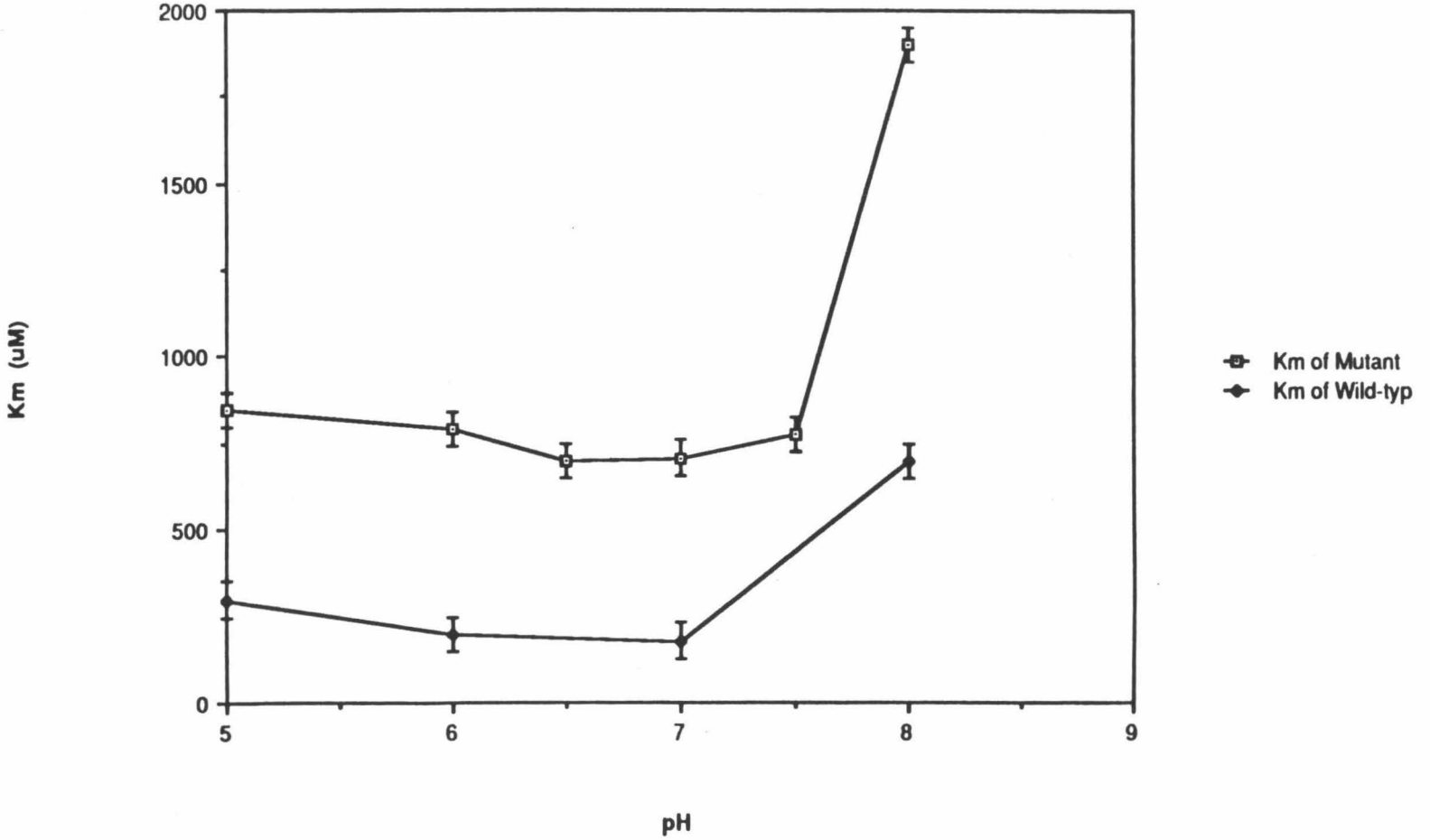

2000
1500
1000
500
0
Km (uM)
5
6
7
8
9
pH
Km of Mutant
Km of Wild-typ

FIGURE 13

K_{cat} vs. pH for the T71H mutant β-lactamase. The substrate is 6-aminopenicillanic acid.

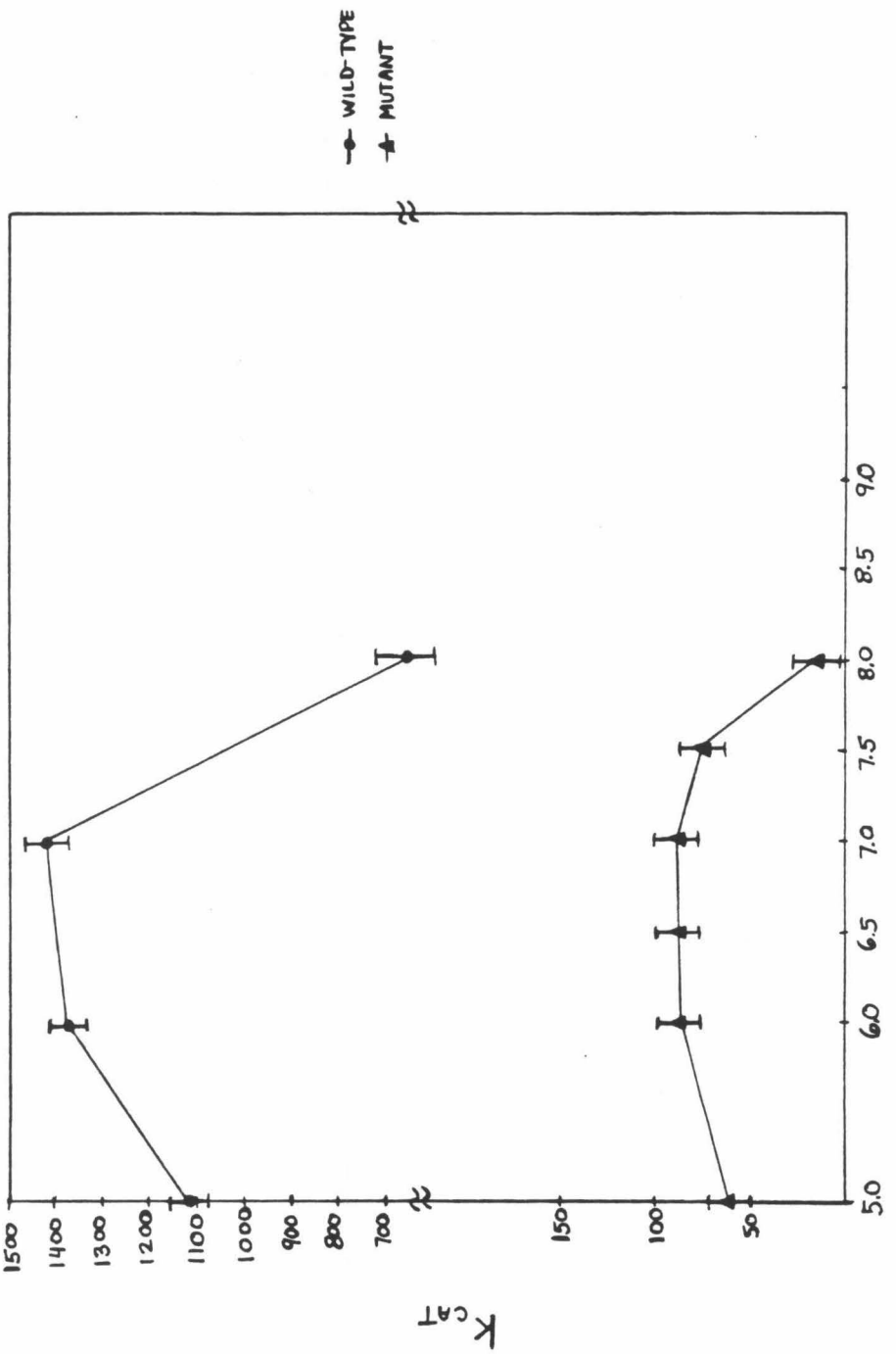

WILD-TYPE
MUTANT
1500
1400
1300
1200
1100
1000
900
800
700
150
100
50
5.0
6.0
6.5
7.0
7.5
8.0
8.5
9.0
pH
K_{cat}

FIGURE 14

K_{cat} vs. pH for the T71H mutant β-lactamase. The substrate is benzylpenicillin.

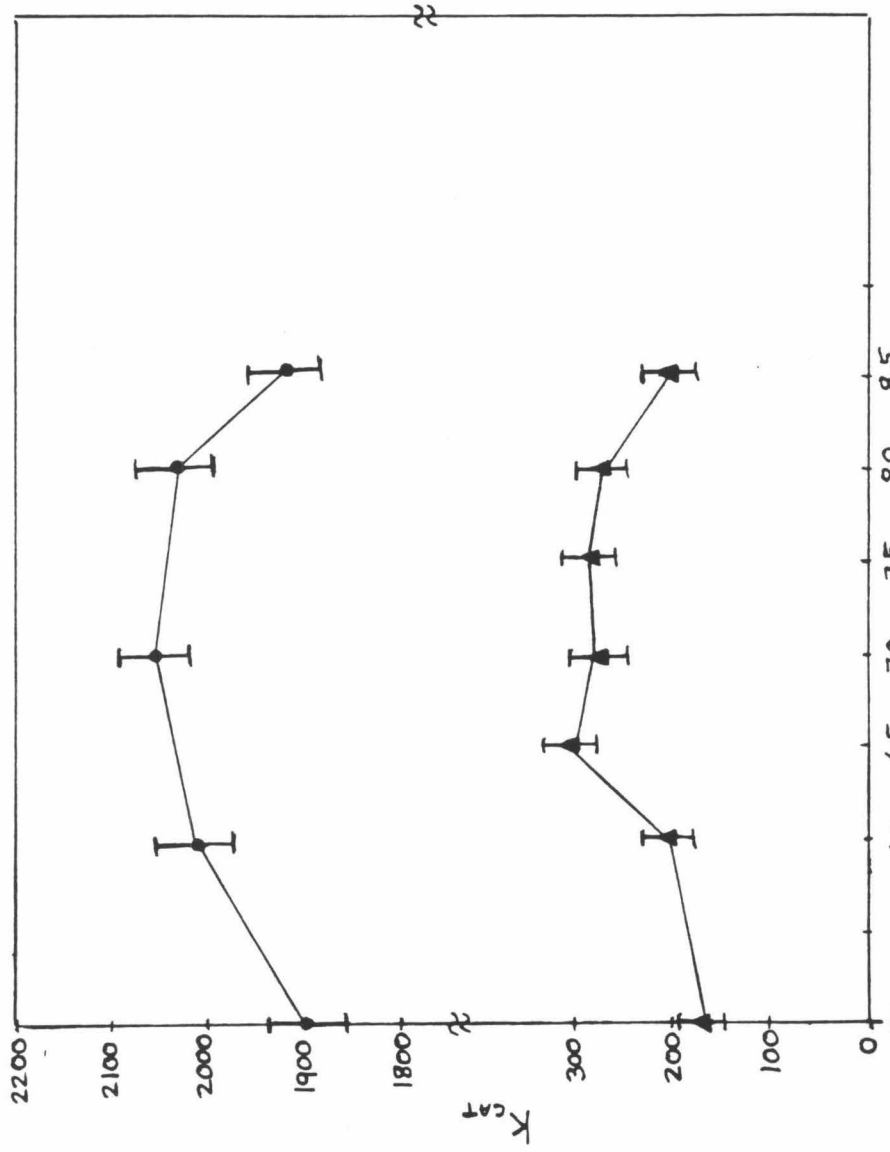

2200
2100
2000
1900
1800
300
200
100
0
5.0
6.0
6.5
7.0
7.5
8.0
8.5
pH
K_{cat}

FIGURE 15

K_{cat} vs. pH for the T71H mutant β-lactamase. The substrate is cephalothin.

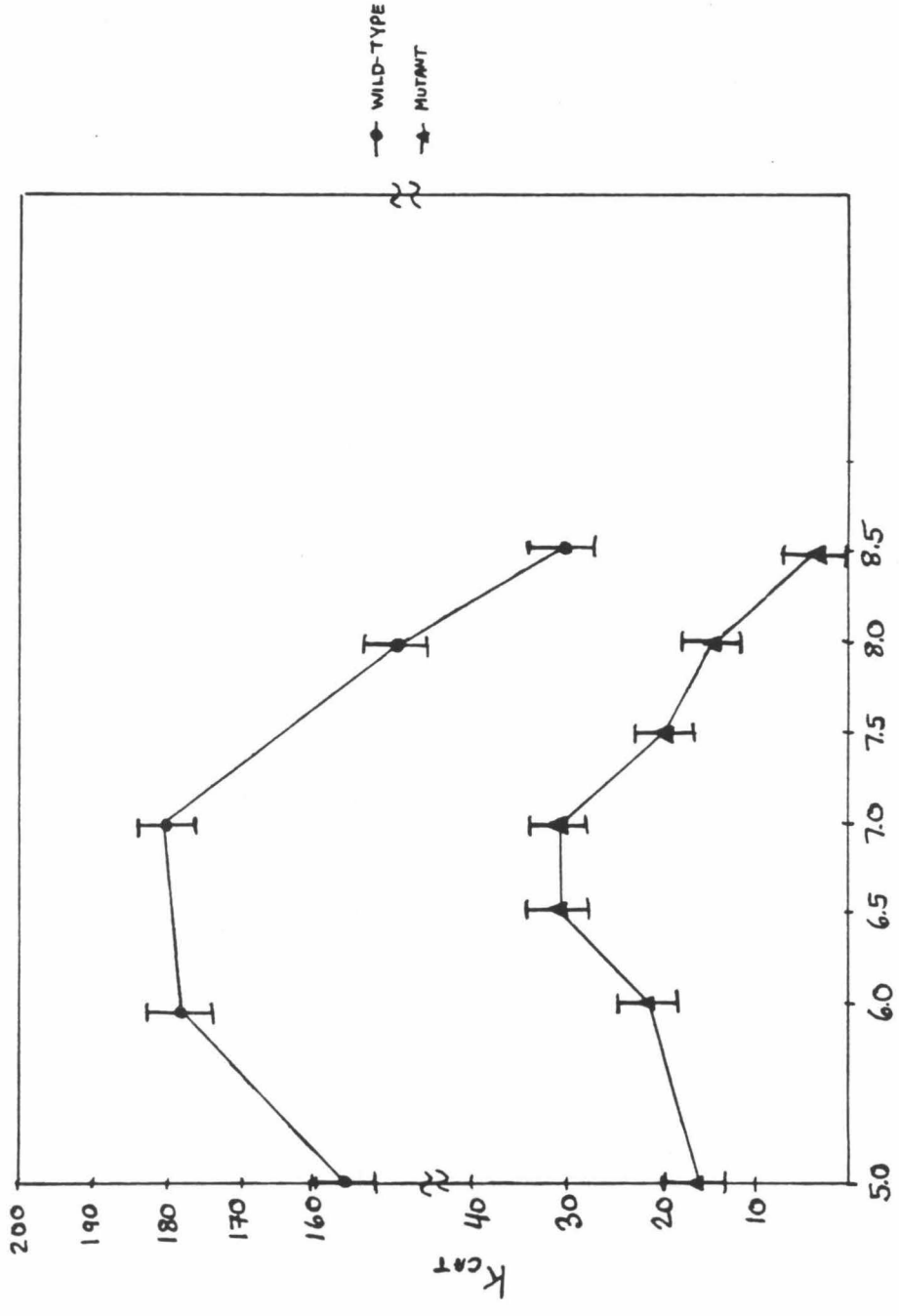

WILD-TYPE
MUTANT
200
190
180
170
160
40
30
20
10
k_{cat}
5.0
6.0
6.5
7.0
7.5
8.0
8.5
pH

FIGURE 16

Thermal stability vs. pH of the T71 mutant. Activity vs. incubation temperature at pH 7.0.

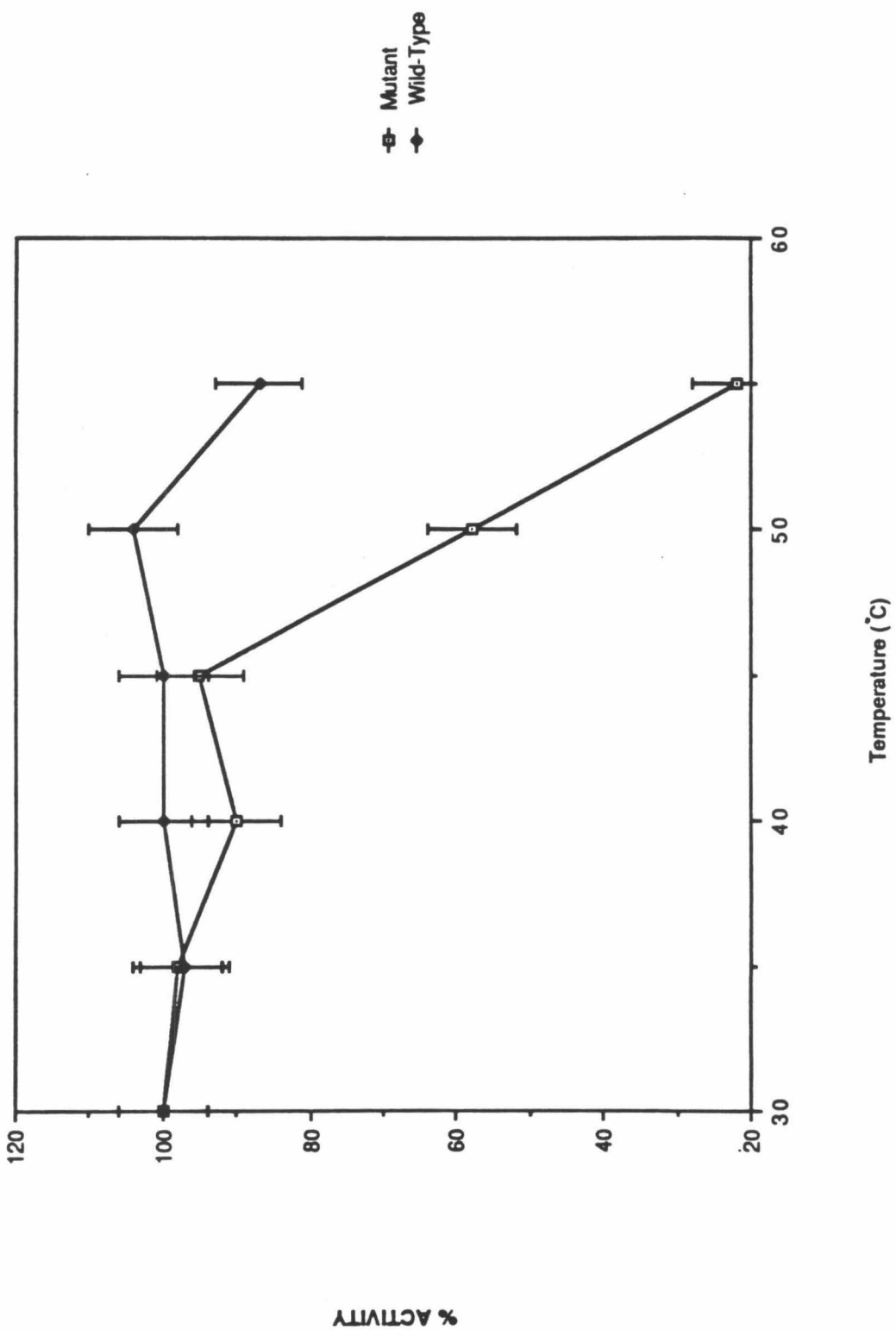

Mutant
Wild-Type
% ACTIVITY
120
100
80
60
40
20
30
40
50
60
Temperature (°C)

FIGURE 17

Thermal stability vs. pH for the T71H mutant. Activity vs. incubation temperature at pH 5.0.

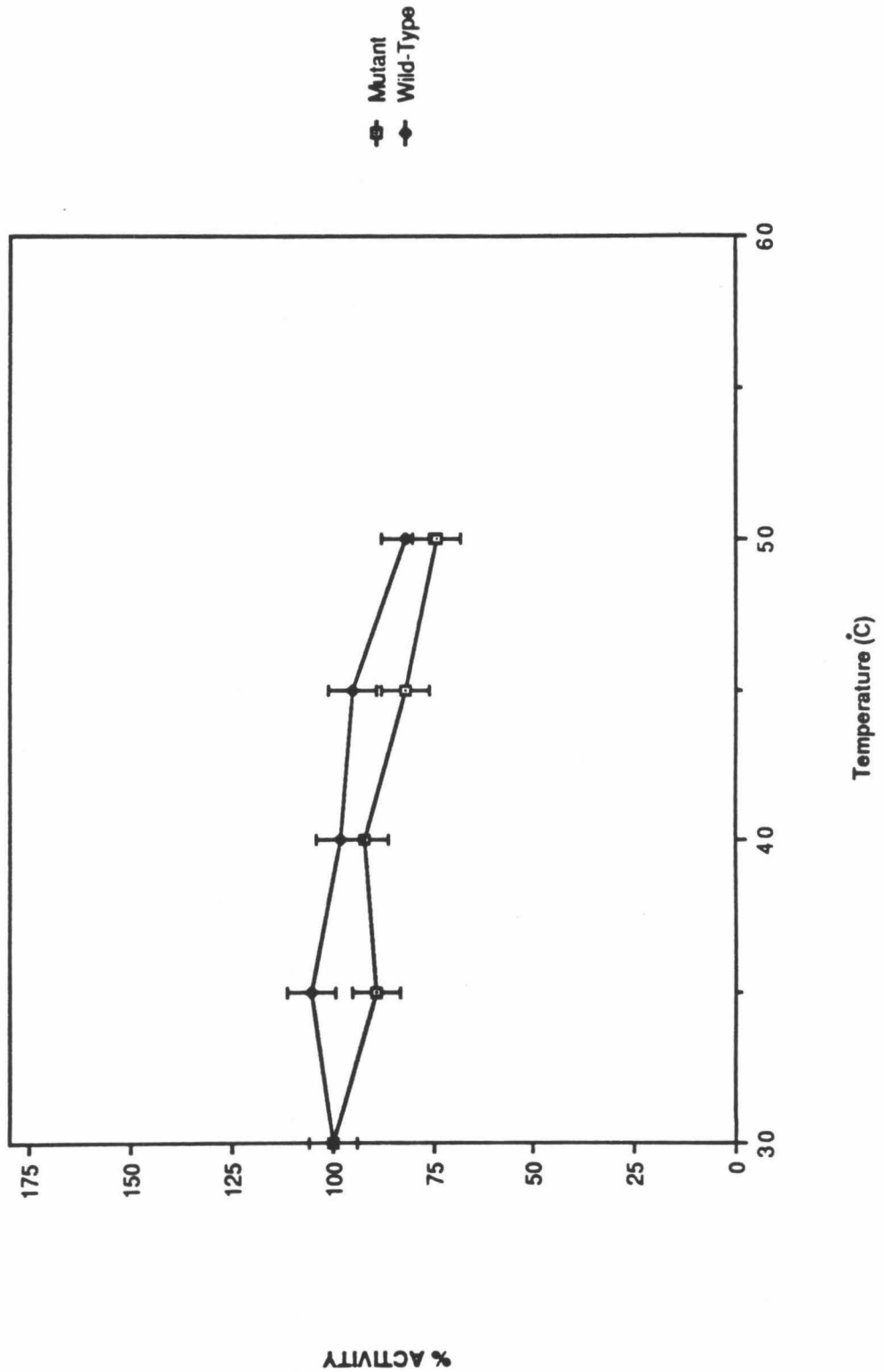

Mutant
Wild-Type
% ACTIVITY
175
150
125
100
75
50
25
0
30
40
50
60
Temperature (C)

FIGURE 18

Thermal stability vs. pH for the T71H mutant. Activity vs. incubation temperature at pH 8.0.

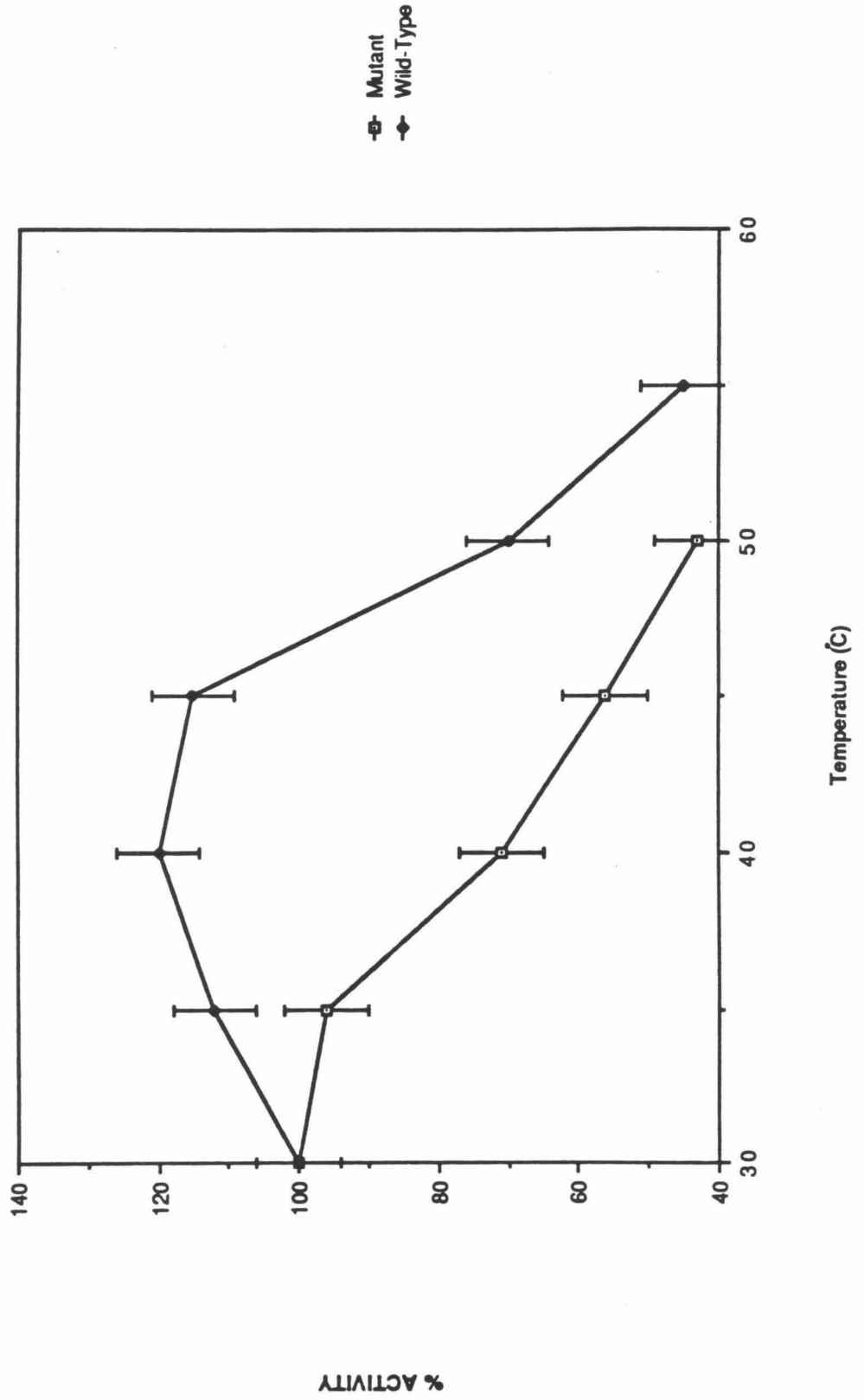

Mutant
Wild-Type
% ACTIVITY
140
120
100
80
60
40
30
40
50
60
Temperature (C)

FIGURE 19

The cephalothin molecule.

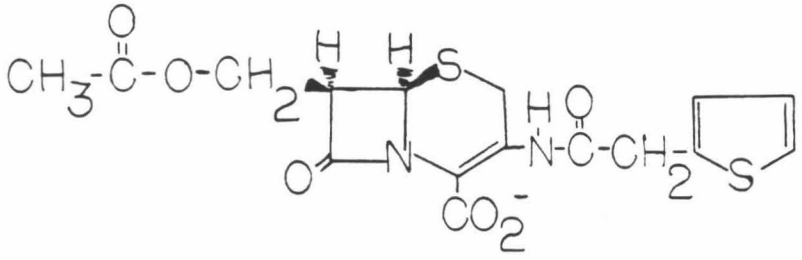


Table 1

Kinetic Parameters
T71H (wild-type)

	Kcat (sec)	Km (uM)	Kcat/Km (M sec)
Penicillin G	270+15 (2000)	20+5 (20)	1.35 x 10 (10)
Cephalothin	30+3 (120)	743+37 (190)	4.0 x 10 (6 x 10)
6-APA	77+4 (1500)	188+8 (195)	4.1 x 10 (7.6 x 10)

	Kcat/Km T71H/Wild-type
Penicillin G	0.135
Cephalothin	0.066
6-APA	0.054

	[(Kcat/Km)Cephalothin]/[(Kcat/Km)PenG]
T71H	0.30%
Wild-type	0.60%

	[(Kcat/Km)6-APA]/[(Kcat/Km)PenG]
T71H	3.00%
Wild-type	7.60%

TABLE 2

Half Lives vs. Temperature

TEMPERATURE (°C)	WILD-TYPE	T71H MUTANT
40		40 Min.
45		8 Min.
50	> 30 Min.	4Min.
55	1-3 Min.	< 1 Min.
60	1-2 Min.	

APPENDIX

Isolation of the Natural Gene for α-Lytic Protease

INTRODUCTION

Proteases are an enormously diverse group of enzymes that can be classified into four major categories: "metal proteases" (such as thermolysin)""thiol proteases" (such as papain), "acid proteases"(such as pepsin) and "serine proteases" (such as trypsin)(1).

α-lytic protease from Lysobacter enzymogenes is metal-free (2) and its cleavage pattern on insulin suggests a relationship with pancreatic elastase, which is a serine protease (3). Furthermore, α-lytic protease is rapidly and totally inactivated by DFP (Diisopropyl Phophorofluoridate). DFP inhibits serine proteases by esterfying the reactive serine residue (4).

Serine proteases can generally be subdivided into two groups (5): (i) Those that have in common the sequence Thr-Ser-Met at the active site serine residue. These serine proteases are isolated from bacteria. Subtilisin from Bacillus subtilis is an example. (ii) Those that have in common the sequence Asp-Ser-Gly at the active site serine residue. These serine proteases are isolated from higher organisms like mammals. Trypsin and elastase are good examples. α-lytic protease was the first enzyme to eliminate a previously sharp distinction between bacterial serine proteases and the

serine proteases of higher organisms. This is because of a feature of the α-lytic protease in its being a microbial enzyme but with the amino acid sequence around its active site serine and a single histidine residue (Val-Ala-Gly-His-Lys-Gly) that is conserved in pancreatic serine proteases. Therefore, it is classed as a homolog of the pancreatic serine proteases (6). In fact, its enzymatic properties are very similar to those of porcine elastase (7).

α-lytic protease has a structural homology with that of pancreatic serine proteases. It was determined that 108 α-carbon positions of α-lytic protease are topologically equivequivalent (55%) to residues of porcine elastase, even though thereis little homology in their primary sequence (<21%)(8). Topological equivalence among the bacterial enzymes is approximately 84% (8). α-lytic protease consists of a single polypeptide chain of 198 residues with a molecular weight of 19,874 (9). Table 1 shows the amino acid sequence of α-lytic protease. Table 2 compares this sequence with those of porcine elastase, bovine chymotrypsin A and B, bovine trypsin and the partial sequence of the Streptomyces griseus trypsin-like enzyme. In these sequences, deletions and insertions have been made in order to maximize homologies especially among the residues conserved, such as His-57, Ser-195, Asp-102, the disulphide bridges and the N-

terminus (9).

The activity of ∝-lytic protease depends upon an unprotonated imidazole group of its single histidine residue. This histidine is conserved in the pancreatic serine protease family (10). At 2.8 A resolution, the catalytic quartet (Ser-214,Asp-102, His-57 and Ser-195) in ∝-lytic protease seems to have the same configuration found in each member of the Gly-Asp-Ser-Gly-Gly serine protease family (11). X-ray diffraction studies of chymotrypsin reveal that Asp-102 is in a "buried" position and is hydrogen bonded to His-57 but it appears that the side-chains of His-57 and Ser-195 are not within hydrogen bonding distance of each other due to the positioning of the side -chains in the free enzyme (12). However, crystallographic evidence has indicated that the side-chain of Ser-195 rotates to a new position in formimg an enzyme-substrate complex and this rotation allows for a strong His-Ser hydrogen bond.

The molecular basis for the substrate specificity of ∝-lytic protease is understandable from the type of amino acid residues near the active site. An insertion of five residues at position 217 and the rotation of the side-chain of Met-192 account for a small specificity pocket which can only bind residues such as Ala, Ser or Val (11). Esters of L-alanine are found to be the best substrates. ∝-lytic protease also has a preference for long substrates, indicating additional secondary binding

subsites (13).

α-lytic protease has been important in the determination of the mechanism of serine proteases. The initial discovery of the single histidine in the sequence eliminated earlier proposals for the involvement of two histidine residues in catalysis (14). The presence of a single histidine makes α-lytic protease an excellent subject for NMR studies to elucidate the catalytic mechanism, since resonance peaks can be assigned to the active site histidine (15,16).

Although these studies indicate that a similiar mechanism operates in α-lytic protease as in other serine proteases, there is still controversy concerning the basic details of the mechanism (7,17). There is agreement that the role of Ser-195 is to act as a nucleophile but the disagreement comes when explaining how the serine residue in serine proteases becomes such a potent nucleophile when the $^{\gamma}O$ of Ser is a normally poor nucleophile (19). There are basically three hypotheses of the mechanism concerning the role of the Asp-His dyad in its accepting, storing and donating the proton originally on Ser-195. In Figure 1a, this shows the original "charge relay" mechanism postulated by Blow et al. (20). This mechanism argues that the charge on Asp-102 is relayed to Ser-195 via the imidazole ring of His-57. This yields an oxyanion of Ser- 195 which undergoes nucleophilic attack. In Figure 1b, this is the

modified "charge relay" or "proton shuttle" mechanism of Hunkapiller et al. (16). This mechanism argues that Ser-195 becomes polarized through the srtong hydrogen bonded system of Asp-His-Ser. Following this, proton inventory studies revealed a concerted two-proton transfer from first Ser to His, then His to Asp. In Figure 1c, Bachovchin and Roberts (21) postulated a "carboxylate assisted" mechanism. This states that Asp's role is to orient the imidazole ring of His to its proper tautomer and this His is the ultimate proton acceptor. The charges formed are stabilized through ion-pairing.

The use of site-specific mutagenesis on the gene of α-lytic protease can provide a wealth of information on the structure-function characteristics of the protein and offer valuable insights on the mechanism of serine proteases. Site-specific mutagenesis, which uses synthetic oligonucleotides to direct mutagenesis, was first developed in ϕX 174 (22,23), a single-stranded phage. Since then, site-specific mutagenesis has been used to make several mutations in ϕX 174 (24) as well as genes cloned into M13 (25).

Although site-specific mutagenesis has been widely used in the last few years to study structural changes in polynucleotides, it is only recently that the technique has been applied to the structure-function effects on enzymes. Site-specific mutagenesis may be

used to discover which residues are involved in catalysis without three-dimensional structural information on the protein (26).

In order to do site-specific mutagenesis, certain conditions must be met. The gene for the enzyme has to be cloned and expressed. The DNA sequence is also essential. The following project was started before the synthetic gene was constructed and its objective was to isolate the natural gene for α-lytic protease from Lysobacter enzymogenes. The synthesis of a long probe such as 51 base pairs from the amino acid sequence would be used to screen a library of DNA fragments from Lysobacter enzymogenes.

MATERIALS AND METHODS

Enzymes and Chemicals

Restriction enzymes were purchased from Bethesda Research Laboratories. Polynucleotide kinase was purchased from Boehringer Mannheim. [γ-^{32}P] ATP, >5000 Ci/mmol was obtained from ICN (Irvine, Ca). Tryptone and yeast extract were from Difco (Detroit, MI). DEAE membrane was purchased from Schleicher and Schuell and the screening filters were obtained from Whatman (541). All other chemicals used were of reagent grade or better.

Bacterial Strains

Lysobacter enzymogenes was started from frozen stock stored by Robert Kaiser. E. coli strain LS1 (32) was used as a host for derivatives of pBR322.

DNA Preparation

The DNA from Lysobacter enzymogenes was isolated according to published procedures by J. Marmur (27). The average length of DNA was approximately 9,000 base pairs as determined by agarose gel electrophoresis.

Plasmid DNA from pBR322 and from other cloning vectors was isolated according to the alkaline lysis method (28). DNA fragments and restriction digests were run on a 1.2% agarose gel and the DNA removed with DEAE

membrane (29,30).

The 51-base oligonucleotide probe was synthesized by using the phosphoramidite method (31) on the Applied Biosystems DNA synthesizer. The 14 and 11-base probes were synthesized manually by Steve Schultz using the same chemistry . The 14 and 11-base probes were purified by HPLC. The 51-base probe was purified by running on a preparative 20% acrylamide gel, eluting the pieces followed by ion-exchange chromatography and then desalting on a small G-25 Sephadex column.

Restriction Digests of Lysobacter DNA

To determine which enzymes give complete digestions of the Lysobacter DNA and would produce a single band upon hybridization of the probe, various restriction digests were carried out upon Lysobacter DNA.

Each reaction mixture contained 5ug of Lysobacter DNA, 1/10 volume of 10x salt buffer, 1λ of 0.1M DTT, 1ug/L of BSA and, generally, 25 units of restriction endonuclease in a volume of 100λ . The mixture was incubated for 30 minutes at 37°C,then another 25 units of enzyme was added and the reaction mixture incubated another 30 minutes. The mixture was then extracted with Tris-buffered phenol and chloroform. 50λ aliquots (2.5ug Lysobacter DNA) were run on a 1.2% agarose gel and stained with ethidium bromide for 15 minutes. The gels were photographed then treated with a gel drying

solution I (0.5N NaOH,0.1M NaCl) for 15 minutes then a gel drying solution II (0.15M NaCl,0.5M Tris-Cl,pH7.6) for another 15 minutes. The gels were dried overnight with a Biorad gel dryer. The gels are now ready for hybridization to the probe.

Hybridization to Probe

The probe was labeled by using 0.2ng per species of probe, per ml of hybridization buffer. [γ -^{32}P]ATP was added at a 2x molar excess of 5' ends. 1/10 volume of 10x kinase buffer and 1μ of 10 units/μ of T_4 polynucleotide kinase were added and the mixture incubated for 30 minutes at 37 C. 90μ of 1/10 TE (10mM Tris-Cl, 1mM EDTA, pH 8.0) was added, then the mixture was heated at 65 C for 10 minutes to inactivate the enzyme. Excess [γ -^{32}P]ATP is removed on prepacked PD10 columns (Sephadex G-25) equilibrated with 1/10 TE. 500 μ fractions are collected and the radioactive probe was found by scintillation counting. The labeled probe is then added to a hybridization mixture containing 6x SET (20x SET=3M NaCl, 0.4M Tris-Cl, pH 7.8, 20mMEDTA), 0.5% NP40, 250ug/ml yeast tRNA, 50 ug/ml pyrophosphate and 220ug/ml "cold" ATP. The gel is then sealed in a plastic bag with 10ml of the hybridization mixture at room temperature for 3 hours to overnight. The contents are removed and the gel is washed in 6x SSC (20x SSC= 3M NaCl, 0.3M trisodium citrate) (250ml) for 15 minutes at

room temperature then autoradiographed. The temperature of the washing is very important and depends upon the T_D of the probe. For the 14^{16} probe (T_D=40-46 °C), a room temperature wash, 37 °C and a 44-45 C wash are done in that order. The 11^8 probe (T_D= 34-37 C) is washed first at room temperature, then 30 °C and 34 °C in that order. For the 51-base probe, the first wash was at room temperature, then 37 °C ,55 °C and 65 °C in that order.Use of stringent hybridization criteria such as these above is done to distinguish strong binding from non-specific binding.

Cloning of the ∝-Lytic Protease Gene

Sau3a Digests of Lysobacter DNA-Preparation of Inserts

A series of different restriction digests of Sau3a were done on Lyobacter DNA to establish the proper reaction conditions for a partial digestion, thus , ensuring a random representation of fragments when the DNA is cloned. In this experiment, nine tubes containing the same amount (2.5ug) of Lysobacter DNA, 1/10 volume of 10x buffer and 0.5 ug/L BSA were digested with varying amounts of Sau3a. This was done by starting with tube 1 and 7 units of Sau3a and doing a serial dilution through eight tubes. Tube 9 contained no enzyme. All of the tubes were incubated for 45 minutes at 37 °C, then held on ice. The mixture was extracted with Tris-buffered phenol then run on a 1.2% agarose gel. The

lane was found that contained the brightest area of fragments in the 2-4kb range and the lane next to it that contained a 2-fold decrease in enzyme was chosen as being optimal conditions for digestion. This was done since B. Seed et al. (33) calculated that the brightest region of a gel corresponds to twice the extent of digestion necessary for optimal representation. Once the optimal conditions were known, a preparatory gel (7mm) was prepared with 250ug of Lysobacter DNA digested with Sau3a and the 2-4kb region was collected using DEAE membrane filters. The fragments were eluted from the membrane by adding a high salt NET (1M NaCl, 0.1mM EDTA, 20mM Tris, pH 8.0) buffer and incubating at 55-60°C for 10-45 minutes. This washing is repeated to release typically 50-90% of the bound DNA. The fragments are extracted with butanol that has been water-saturated to remove ethidium bromide then the DNA is precipitated by adding 2.5 volumes of ethanol and re-precipitated by adding 0.3M sodium acetate and 2.5 volumes of ethanol to reove any NaCl residue.

Preparation of Vector

pBR322 was cut with BamHI by adding 10ug of pBR322, 30 units of BamHI, 1/10 volume of 10x buffer, 10ug BSA, 1 ʎ of 0.1M DTT and water to a final volume of 100 ʎ . The mixture was incubated for 30 minutes at 37°C. Another 30 units of BamHI was then added and the mixture incubated another 30 minutes. An aliquot was taken (2.5ug) for

phosphatase treatment. The remaining solution was extracted with Tris-buffered phenol and chloroform. The DNA is precipitated by adding 2.5 volumes of 95% ethanol and 0.3 M sodium acetate. The phosphatase treatment involves incubating the 2.5ug of pBr322 cut with BamHI with 2 ⅄ of 10% SDS, 5 ⅄ of Tris-Cl, pH 8.5, 1 ⅄ alkaline phosphatase and sufficient water to give a volume of 31.5 ⅄ for one hour at 37 °C. The mixture is then extracted with Tris-buffered phenol 4 times, then extracted with chloroform once. 1/10 volume of 3M sodium acetate is added and 2.5 volumes of 95% ethanol to precipitate the DNA.

Ligations

Through theoretical calculations based on the theory and practice of ligation (33,34) we determined that, to optimize the number of cloned circular DNA containing both vector and insert, a 3 molar excess of insert/vector was appropriate. However, a dilution of the 3 molar excess inserts/vector ratio was also needed to ensure that a vector "sticky end" would ligate to an insert "sticky end." Therefore, a pilot reaction was conducted to determine wich dilution factor will produce the highest amount of closed circular DNA. The pilot reaction consisted of a mixture containing 0.2ug of pBR322 that had been cut with BamHI and phosphatased, 1 ug BSA, 1⅄ of 0.1M DTT, 1⅄ of T_4 ligase and sufficient

water to a volume of 10λ . This reaction mixture has a concentration that is a 3 molar excess of inserts/vector. Out of this stock, the solution was diluted 2-fold and 5-fold and the contents were incubated at 15 °C overnight. The contents of each reaction were run on a 1.2% agarose gel and stained with ethidium bromide and photographed. It was determined that the 5-fold dilution of the 3 molar excess of inserts/vector produced the greatest amount of closed circular DNA containing both vector and insert.

Transformations

Competent cells were prepared and transformations were done according to Hanahan (34). Transformed cells are spread onto ampicillin/PMSF agar plates and grown overnight (approximately 600 per 15cm plate). The frequency at which inserts were present in the vector was determined by picking colonies onto both amp and tet plates. Inserts are located in the tet gene of pBR322, therefore, cells will not grow on tet plates if they contain inserts. It was determined that approximately 70% of the colonies contained inserts.

Colony Screening Using Labeled Probe

The colonies are transferred to previously sterilized Whatman 541 filters by carefully laying the filter over the colonies. The plate and filter are

marked with a pencil and the filter lifted as soon as it is completely wet (leaving it longer inhibits re-growth on the master plate). In cases where too much of the colony is being lifted by the filter, the filter can be slightly wetted by first placing it on an agar plate. The wetted filter is then used to transfer colonies. The master plates are placed in an incubator at 37 °C for one hour and then maintained at room temperature to prevent overgrowth of the colonies.

Filters bearing colonies are placed colony side down on L agar plates containing 250 ug/L chloramphenicol. Plates are incubated for 18-24 hours at 37 °C. The filters are then lifted and washed twice for five minutes in (a) 0.5M NaOH, (b) 0.5M Tris-Cl, pH 7.4, and (c) 2X SSC with shaking. The filters are then washed briefly in 95% ethanol and air dried. A batch method works well so that 10-20 filters can be washed at one time with slow shaking in approximately 250ml of each of the solutions. After drying, the filters are ready for screening.

The probe was labeled with ^{32}P as discussed before. Filters are first pre-hybridized in a solution (10ml/pouch and 2 filters/pouch) containing 6x SET, 0.5% NP 40, 100 ug/ml denatured, sonicated salmon sperm DNA, 50ug/ml of pyrophosphate and 220ug/ml of "cold" ATP. Salmon sperm DNA, pyrophosphate and "cold" ATP are added to reduce background on filters. The filters are sealed

in a plastic bag with the pre-hybridization buffer for 2 to 4 hours at 65°C. The liquid contents are emptied and the hybridization mixture containing the labeled probe is added and hybridized at room temperature from 3 hours to overnight. The filters are then washed at 4°C in 6x SSC and autoradiographed. The filters are then washed at higher temperature to ensure passing through the temperature that best distinguishes the gene-containing colony from all others. If there are any colonies that seem to be good "candidates" for containing the gene, the colonies are first screened for tet sensitivity then a quick plasmid preparation is done and the plasmids run on a 1.2% agarose gel, dried, then hybridized again to the labeled probe.

RESULTS

Design and Synthesis of Probes

A mixed 14-based oligonucleotide and a different mixed 11-based oligonucleotide were synthesized manually by Steve Schultz by the phosphoramidite method (31). The sequences of both of these probes are located near the middle of the ∝-lytic protease gene and are made to represent all possible combinations of codons predicted from the amino acid sequence (Figure 2). The DNA sequence chosen for the probes was determined by looking for the stretch of DNA encoded with the least degeneracy. Therefore, there are 16 different 14-base oligonucleotides and 8 different 11-base oligonucleotides.

After obtaining the DNA synthesizer for our laboratory and the publication of a paper by Anderson and Kingston (35) in which a 86-base probe was constructed by deducing the amino acid codons, a 51-base probe was constructed to probe for the ∝-lytic protease gene. The 51-base probe binds to the region in the ∝-lytic protease gene with the smallest degeneracy in the genetic code (Figure 2). The 51-base probe was constructed so that any degenerate codon for an amino acid would be replaced by the codon containing the most G or C bases. This is because ∝-lytic protease is known to be highly enriched in G,C bases.

Restriction Digests of Lysobacter DNA

Lysobacter DNA was digested with different restriction enzymes to find if any of the enzymes gave complete digestions and formed any bands upon hybridization to the probes. Nine different restriction enzymes were used; the digests of seven of these are shown in Figures 3 and 4. It is evident that none gave complete digestions. Hybridizations to the 14^{16} probe showed that bands are forming on these digestions (Figures 5 and 6). However, since these bands are from incomplete digestions and are too close to the undigested DNA, one cannot determine the integrity of the "sticky ends." So, another means was attempted to isolate the gene.

Sau3a Digests of Lysobacter DNA

Sau3a was one of two restriction enzymes found that gave complete digestions of Lysobacter DNA but no bands upon hybridization to the 14^{16} probe. The result of the series of Lysobacter DNA digestions with Sau3a is shown in Figure 7. The kilobase scale on the left side of the picture was made by graphing the molecular weight markers of pBR322 on semi-log paper versus their distance traveled down the gel. Based upon this scale, lane 5 was chosen as having the brightest area in the 2-

4kb region. Therefore, the digestive conditions of lane 6 were taken as optimal. Upon preparatory isolation of these fragments, 250ug of Lysobacter DNA yielded approximately 21ug of inserts for cloning.

Screening Sau3a Inserts for ∝-Lytic Protease Gene

The results of the ligation with Lysobacter inserts and pBR322 vector cut with BamHI and phosphotased showed that the 5-fold dilution of the 3 molar excess of inserts/vector was the best choice to transform with since this produced the most closed circular vector plus insert vehicles. After transformation and plating, approximately 15,000 colonies were screened with the 14^{16} probe. Figure 8 shows how a colony bearing the gene for ∝-lytic protease could be distinguished from the other colonies. Colonies that gave positive signals such as these were further screened by hybridizing to a gel with the isolated plasmids from these colonies. However, none of the colonies picked gave positive results.

XmaIII and EcoRI Digests of Lysobacter DNA

After the unsuccessful attempts in isolating the gene from Sau3a inserts, the enzymes XmaIII and EcoRI were tried. Figure 9 show that EcoRI gives a partial digestion while XmaIII gives a complete digestion on Lysobacter DNA. At this point in time, a DNA synthesizer

was available to our laboratory and capable of making very long oligonucleotides. Therefore, the 51-base probe was made and this probe was used to hybridize with the gel containing the XmaIII and EcoRI digests. Figure 10 shows the autoradiograph of hybridizing the 51-base probe to the gel in Figure 9. After washing at 55 °C, a single band is observed in both enzymes corresponding to a molecular weight of 2000-3000bp.

Next, inserts were isolated off preparative gels from XmaIII and EcoRI digestions. These inserts were ligated into pBR322 vectors. However, not many colonies were obtained. The colonies obtained were screened using the 51-base probe. However, none of the colonies were positive. This experiment was duplicated to screen more colonies. Still, with the low frequency of transformation, a total of only 2000 colonies were screened with no results. It was at this point in time that the 51-base probe was given to Dr. Agard's laboratory in San Francisco for his attempt in finding the gene. Also at this time, the author's efforts were starting to be devoted fully and successfully to construction of a synthetic gene for α-lytic protease.

DISCUSSION

The inability of many of the restriction enzymes to cut Lysobacter DNA was probably due to a methylation pattern in the DNA. In fact, this could be one possibility for the low frequency of transformation of the pBR vectors containing the XmaIII or EcoR1 inserts. It is known that E. coli possesses mechanisms to identify and destroy foreign DNA based upon its methylation pattern (32). Another possibility for the low transformation frequency was the agarose our laboratory was using to isolate fragments from agarose gels. It was discovered that commercially available agarose contained some contaminants that interfered with the DNA ligase used in the ligation reaction. Either of these two possibilities have been known to significantly reduce transformation effiency.

It is not clear why colonies containing the α-lytic protease gene could not be detected by screening the Sau3A library with the 14^{16} probe. It was theorectically calculated that with fragments in the 2-4kb range that approximately 5000 colonies would have to be screened to obtain a 99% chance of finding the gene. It is known that Lysobacter DNA fragments are being cloned as these colonies are tetracycline sensitive and approximately 70% of all colonies grown are tetracycline sensitive. Therefore, approximately 5000 colonies were being

screened containing the Sau3A inserts. The screening technique was also ruled out as a factor since the author screened Gloria McFarland's 2269 β-lactamase mutant and obtained positive result. Probing with the 51-base oligonucleotide was promising in finding the gene. The single band obtained in hybridizing the 51-base probe to XmaIII and EcoRI digests gave evidence that the gene was contained in a single fragment and was clonable. However, since the gene was not found, it seems likely that the difficulty was that not enough colonies were being screened. In fact, after giving the 51-base probe to Dr.Agard's laboratory, he succeeded in finding the gene between two EcoRI sites. Sequencing the natural gene was incomplete since GC compression makes reading the sequence difficult. Dr.Agard's attempt at expressing the gene have not been successful to date. It seems some folding problems are becoming evident. Recently, Drs. Wensink and Abeles at Brandeis University have also cloned the ∝-lytic protease gene and are using the author's synthetic gene as a probe to verify their results. David Epstein at Brandeis University has sequenced a portion of the natural gene that includes the sequence region for the 51-base probe. His results are seen in Figure 11. Nucleotides that match the author's 51-base probe are indicated by an asterisk. It was calculated that the 51-base probe is 88.2% homologous to the natural DNA sequence. Therefore, it is

conclusive that long base probes, though not 100% homologous to the actual DNA sequence, can be used to hybridize with high specificity to DNA sequence.

REFERENCES

1. B.S.Hartley (1960) Ann. Rev. Biochem.,29, 45
2. L.Jurasek and D.R.Whitaker (1967) Can. J. Biochem., 45, 917.
3. M.A. Naughton, F.Sanger, B.S. Hartley and D.C. Show
4. D.R.Whitaker and C. Roy (1967) Can. J. Biochem, 45,911- 916.
5. G.H. Dixon (1966) in "Essays in Biochemistry," Vol.2, edited by P.N.Campbell and G.D. Greille, Acadmeic Press (London) p.190.
6. D.R.Whitaker (1970)"Methods of Enzymology," Vol.19,p.599-613.
7. H. Kaplan,V.B.Symonds,H.Dergas and D.R.Whitaker (1970) Can. J. Biochem.,48, 649-658.
8. M. James,L.Delbaere and G.Brayer (1978) Can. J. Biochem.,56,396-402.
9. M.D.J.Olson,N.Nagabhusham,M.Dziveniel,L.B.Smillie and D.R.Whitaker (1970) Nature ,228, 438-442.
10. L.B.Smillie and D.R.Whitaker (1967) J. Amer. Chem. Soc., 89,3350.
11. G.Brayer,L.Delbaere and N. James (1979) J. Mol. Biol., 131,743-775.
12. R.Henderson (1970) J. Mol. Biol. ,54, 341-354.
13. D.R.Whitaker (1965) Can. J. Biochem. ,43, 1935-1954.

14. K.A.Walsh,D.L.Kauffman,K.S.V.Sampath Kuman and H.Neurath (1964) Proc. Natl. Acad. Sci. USA., 51, 301- 308

15. G.Robellard, and R.G.Shulman (1974) J. Mol. Biol. ,86, 519-540.

16. M.W.Hunkapillar,J.H.Smallcombe,D.R.Whitaker and J.H.Richards (1973) Biochemistry, 12, 4732-4773.

17. J.Kraut (1977) Ann. Rev. Biochem. , 46, 331-358.

18. D.M.Blow (1976) Acc. Chem. Res., 9, 145-152.

19. R.Kaiser (1983) Ph.D.Thesis, California Institute of Technology.

20. D.M.Blow,J.J.Birktoft,and B.S.Hartley (1969) Nature,221,337-339.

21. W.W.Bachovchin and J.D.Roberts (1978) J. Amer. Chem. Soc. ,100, 8041-8047.

22. A.Razin,T.Hirose, K.Itakura and A.D.Riggs (1978) Proc. Natl. Acad. Sci. USA, 75 , 4268.

23. D.Shortle and D. Nathans (1978) Proc. Natl. Acad. Sci. USA, 75, 2170.

24. O.S.Blanot, S.A.Khan and R.W.Chambers (1979) J. Biol. Chem. ,254, 12684.

25. I.Kudo,M.Leinewiber,and V.L.Rajbhandary (1981) Proc. Natl. Acad. Sci. USA., 78, 4753.

26. R.B.Wallace,P.F.Johnson,S.Tanaka, M.Schoeld, K.Itakura and J. Abelson (1980) Science, 209, 1396.

27. J.Marmur (1961) J. Mol. Biol., 3,208-218.

28. D.Ish-Horowicz and J.F. Burke (1981) Nucleic Acids

Res. 9, 2989-2998.

29. G.Winberg and M.-L.Hammarskjold (1980) **Nucleic Acids Res.** ,**8**, 253-264.

30. D.Dritzen,M.Bellard, P.Sassone-Corsiand P.Chambon (1981) **Anal. Biochem.**,**112**, 295-298.

31. S.L.Beaucage and M.H.Caruthers (1981) **Tetrahedron Lett.**,**22**, 1859-1862.

32. H.W.Boyer and D.Roulland-Dussoix (1969) **J. Mol. Biol.** ,**41**,459-472.

33. B.Seed,R.C.Parker and N. Davidson (1982) **Gene**, **19**, 201- 209.

34. D.Hanahan (1983) **J. Mol. Biol.** , **166**, 557-580.

35. S.Anderson and I.B.Kingston (1983) **Proc. Natl. Acad. Sci. USA**, **80**, 6838-6842.

FIGURE 1

Hypotheses concerning the mechanism of serine proteases (18).(a) "charge relay" mechanism by Blow et al. (20), (b) "proton shuttle" mechanism by Hunkapillar et al. (16) and (c) "carboxylate assisted" mechanism by Bachovchin and Roberts (21).

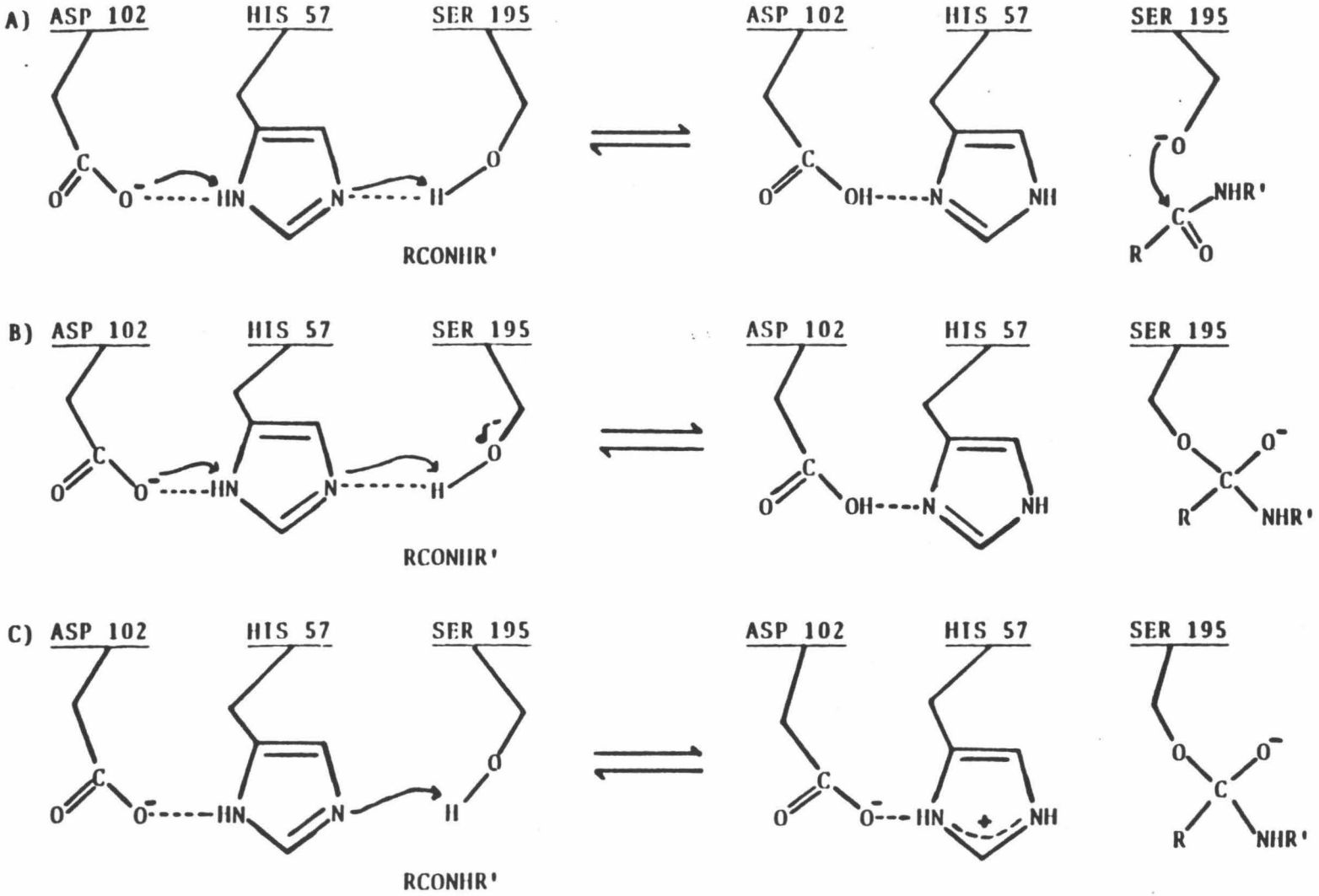

A)
ASP 102
HIS 57
SER 195
RCONHR'
B)
ASP 102
HIS 57
SER 195
RCONHR'
C)
ASP 102
HIS 57
SER 195
RCONHR'

FIGURE 2

Reverse translation of α-lytic protease. 11^{8}, 14^{16} and 51-base probes are enclosed by arrows.

AMBIGUOUS BASE CODE. N=A,T,C,G; P=A,G; Y=C,T; R=A,T; S=C,G, Z=A,T,C.

```
1     A N I V G G I E Y S I N N A S L C S V G     20
1     GCNAAYATZGTNGGNGGNATZGAPTAYRSNATZAAYAAYGCNRSNYTNTGYRSNGTNGGN     60

21    F S V T R G A T K G F V T A G H C G T V     40
61    TTYRSNGTNACNXGNGGNGCNACNAAPGGNTTYGTNACNGCNGGNGCAYTGYGGNACNGTN     120

41    N A T A R I G G A V V G T F A A R V F P     60
121   AAYGCNACNGCNXGNATZGGNGGNGCNGTNGTNGGNACNTTYGCNGCNXGNGTNTTYCCN     180

61    G N D R A W V S L T S A Q T L L P R V A     80
181   GGNAAYGAYXGNGCNTGGGTNRSNYTNACNRSNGCNCAPACNYTNYTNCCNXGNGTNGCN     240

81    N G S S F V T V R G S T E A A V G A A V     100
241   AAYGGNRSNRSNTTYGTNACNGTNXGNGGNRSNACNGAPGCNGCNGTNGGNGCNGCNGTN     300
                                    5' base probe
101   C R S G R T T G Y Q C G T I T A K N V T     120
301   TGYXGNRSNGGNXGNACNACNGGNTAYCAPTGYGGNACNATZACNGCNAAPAAYGTNACN     360
                        11'
121   A N Y A E G A V R G L T Q G N A C M G R     140
361   GCNAAYTAYGCNGARGGNGCNGTNXGNGGNYTNACNCAPGGNAAYGCNTGYATGGGNXGN     420
                                    14'

141   G D S G G S W I T S A G Q A Q G V M S G     160
421   GGNGAYRSNGGNGGNRSNTGGATZACNRSNGCNGGNCAPGCNCAPGGNGTNATGRSNGGN     480

161   G N V Q S N G N N C G I P A S Q R S S L     180
481   GGNAAYGTNCAPRSNAAYGGNAAYAAYTGYGGNATZCCNGCNRSNCAPXGNRSNRSNYTN     540

181   F E R L Q P I L S Q Y G L S L V T G     198
541   TTYGAPXGNYTNCAPCCNATZYTNRSNCAPTAYGGNYTNRSNYTNGTNACNGGN     594
```

FIGURE 3

1.2% agarose gel of Lysobacter DNA digests stained with ethidium bromide. Lane 1 is 0.42ug of pBR322; lane 2, 0.37ug of pBR322; lane 3, 0.2ug of Lysobacter DNA; lane 4, 0.50ug of Lysobacter DNA; lane 5 through 8, 2.5ug of Lysobacter DNA cut with ClaI, BamHI, EcoRI and PstI, repectively; lane 9, 0.4ug of salmon sperm DNA.

LANES

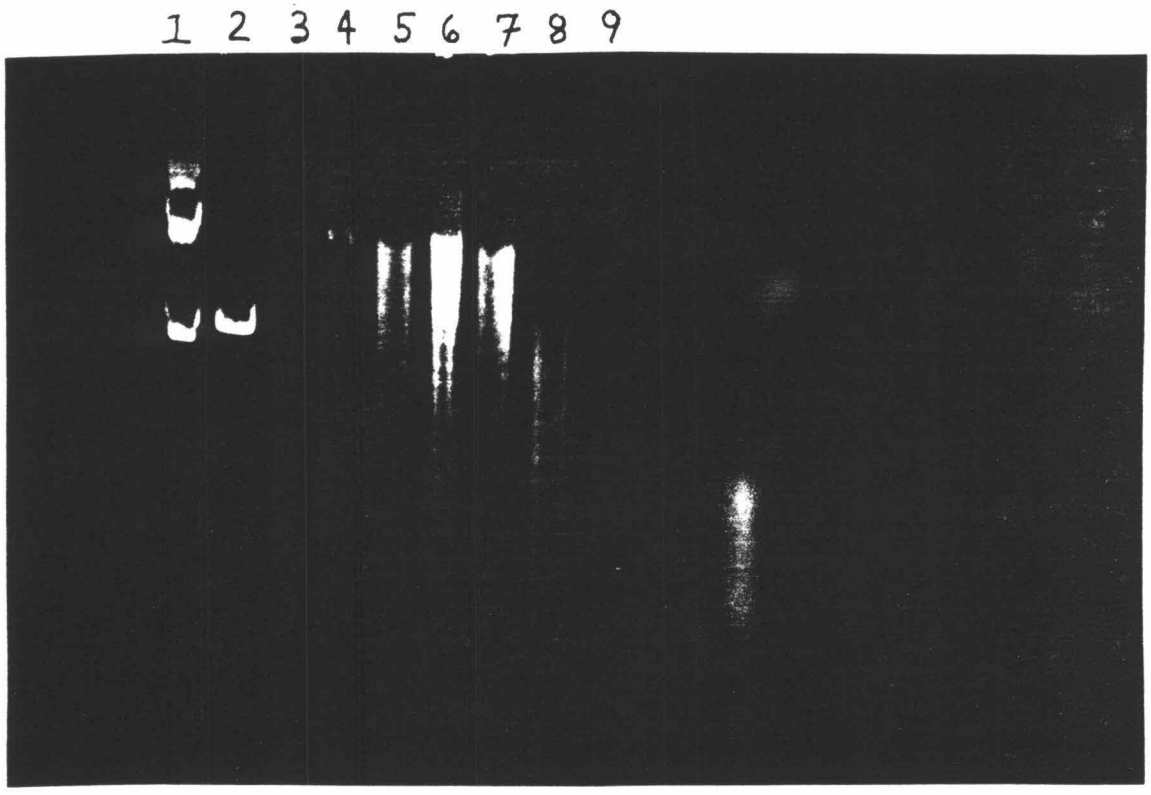


FIGURE 4

1.2% agarose gel of Lysobacter DNA digests stained with ethidium bromide. Lane 1 is 0.85ug of pBR322; lane 2, 1ug of pBR322 cut with HincII; lane 3 through 6, 15ug of Lysobacter DNA cut with BamHI, SphI, SalI and NruI, respectively; lane 7, 5ug of Lysobacter DNA.

LANES

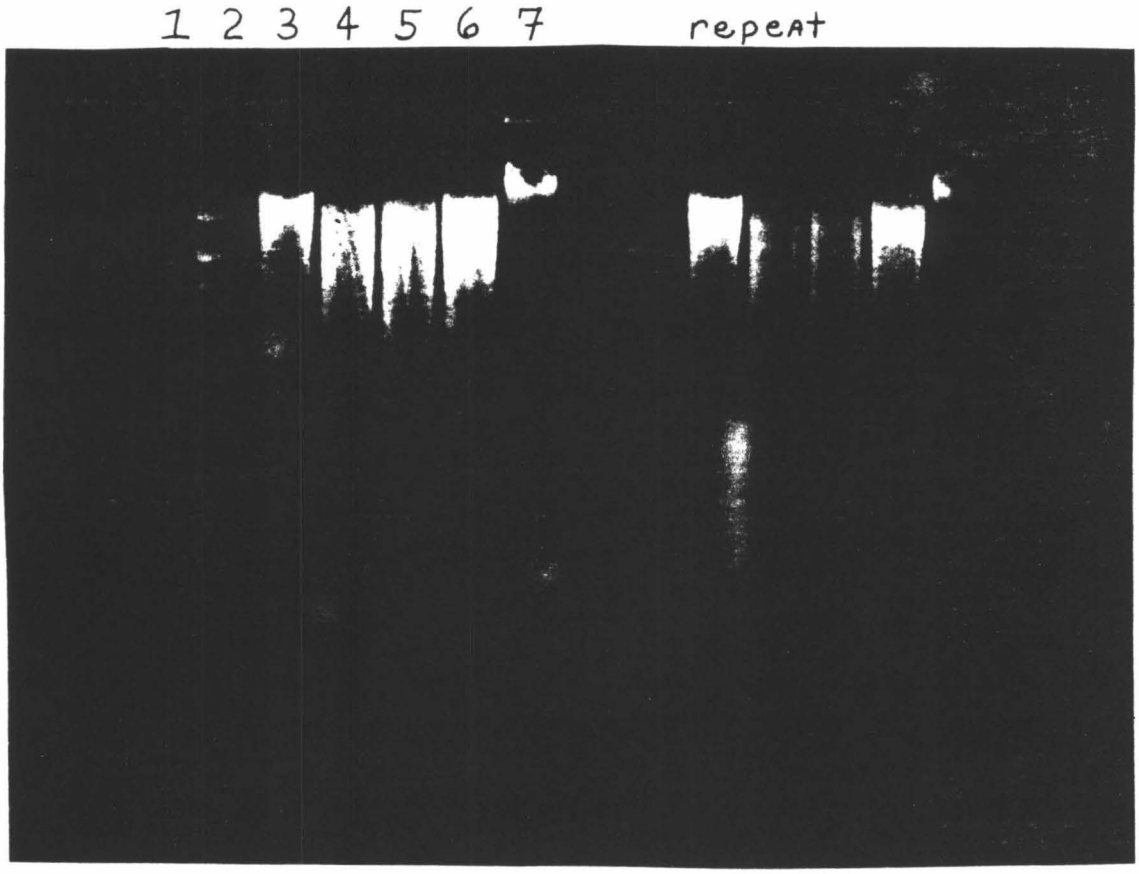


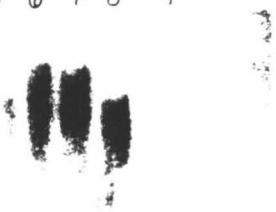
FIGURE 5

Hybridization of ^{32}-P labeled oligonucleotide (14^{16}) probe with Lysobacter DNA digests of Figure 3. Lane 5 is ClaI cut Lysobacter DNA; lane 6 through 8, Lysobacter DNA cut with BamHI, EcoRI and PstI, respectively. (a) Hybridization followed by a room temperature wash in 6X SSC. (b) Same gel after a 40-42 C wash in 6X SSC.

A)

LANES

3 4 5 6 7 8 9



B)

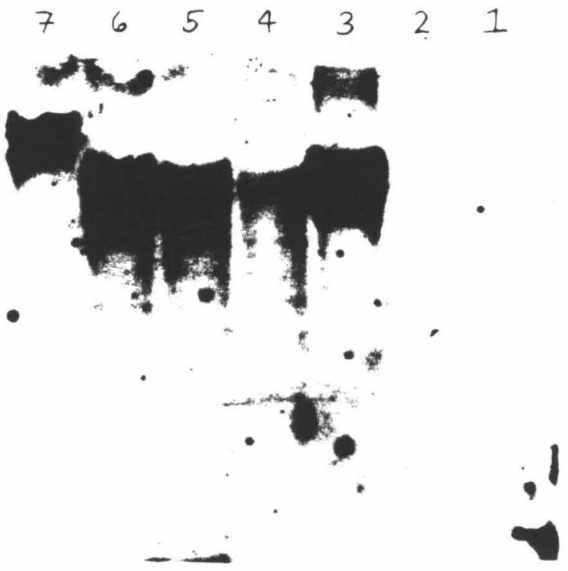
LANES

4 5 6 7 8

FIGURE 6

Hybridization of ^{32}-P labeled oligonucleotide (14^{16}) probe with Lysobacter DNA digests of figure 4. Lane 1 is pBR322; lane 2, HincII cut pBR322; lanes 3 through 6, Lysobacter DNA cut with BamHI, SphI, SalI and NruI, respectively; lane 7, uncut Lysobacter DNA. (a) Hybridization followed by room temperature wash in 6X SSC. (B) Same gel after a 37°C wash in 6X SSC.

A)



B)

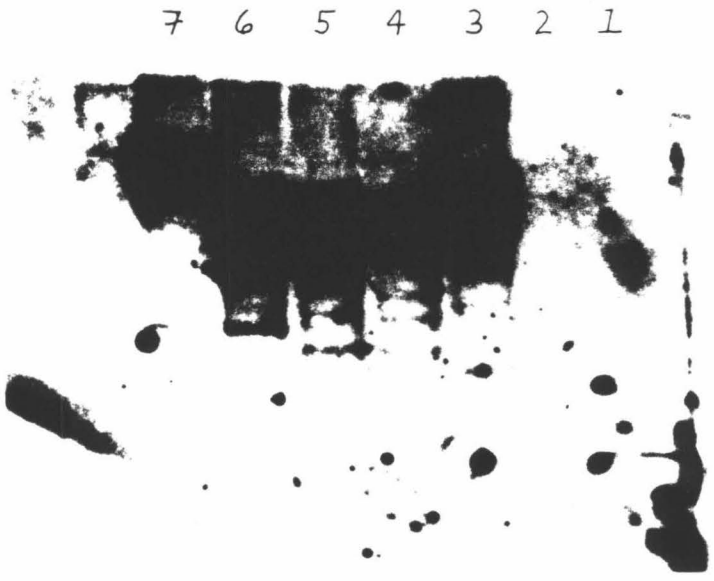


FIGURE 6

Hybridization of ^{32}-P labeled oligonucleotide (14^{16}) probe with Lysobacter DNA digests of figure 4. Lane 1 is pBR322; lane 2, HincII cut pBR322; lanes 3 through 6, Lysobacter DNA cut with BamHI, SphI, SalI and NruI, respectively; lane 7, uncut Lysobacter DNA. (a) Hybridization followed by room temperature wash in 6X SSC. (B) Same gel after a 37°C wash in 6X SSC.

LANES

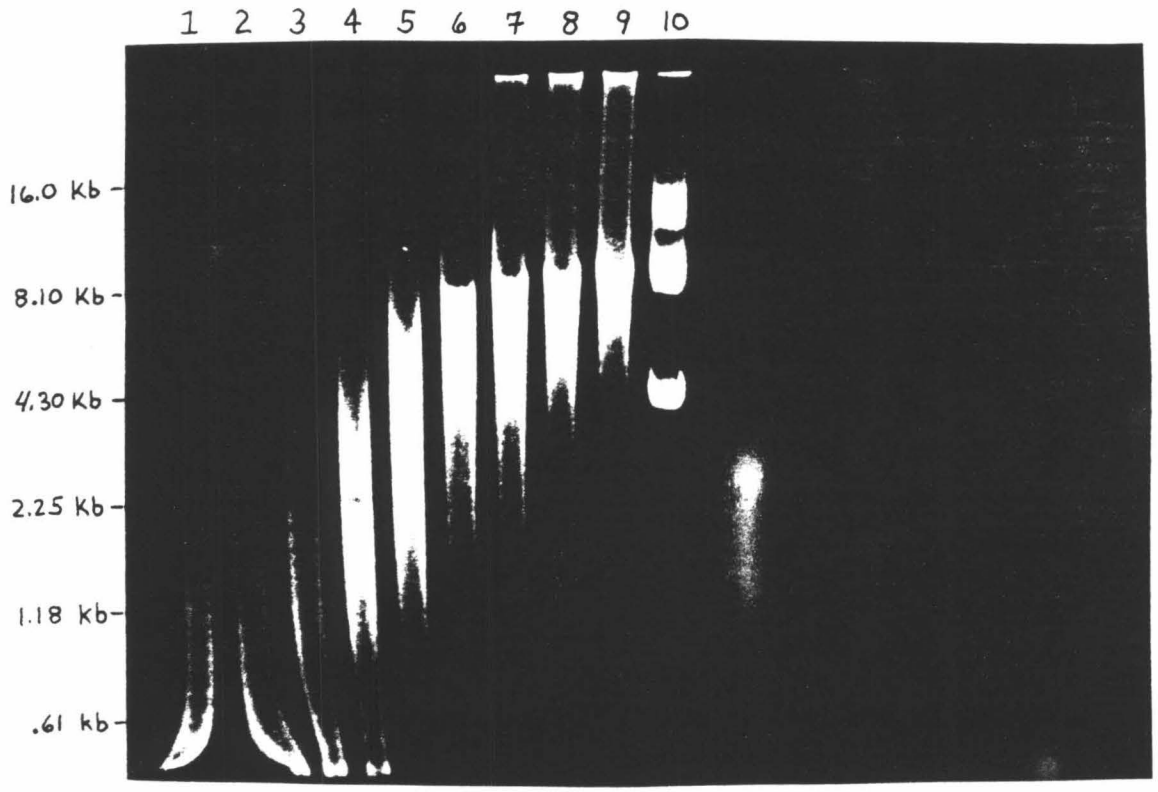


FIGURE 8

Colony screening for the α-lytic protease gene by a ^{32}P-labeled oligonucleotide (14^{16}). (a) Hybridization to a colony lifted on a Whatman 541 filter followed by washing at room temperature with 6x SSC. (b) Same filter after being washed at 37°C in 6x SSC. (c) Same filter after being washed at 45°C in 6x SSC.


A
B
C

FIGURE 9

1.2% agarose gel of Lysobacter DNA digests. Lane 1 is uncut Lysobacter DNA; lane 2, HincII cut pBR322; lane 3, Lysobacter DNA cut with EcoRI; lane 4, XmaIII digestion of Lysobacter DNA ; lane 5 and lane 6, pBR322. Lanes 14 through 18 are a repeat, respectively, of lanes 1 through 6.

LANES

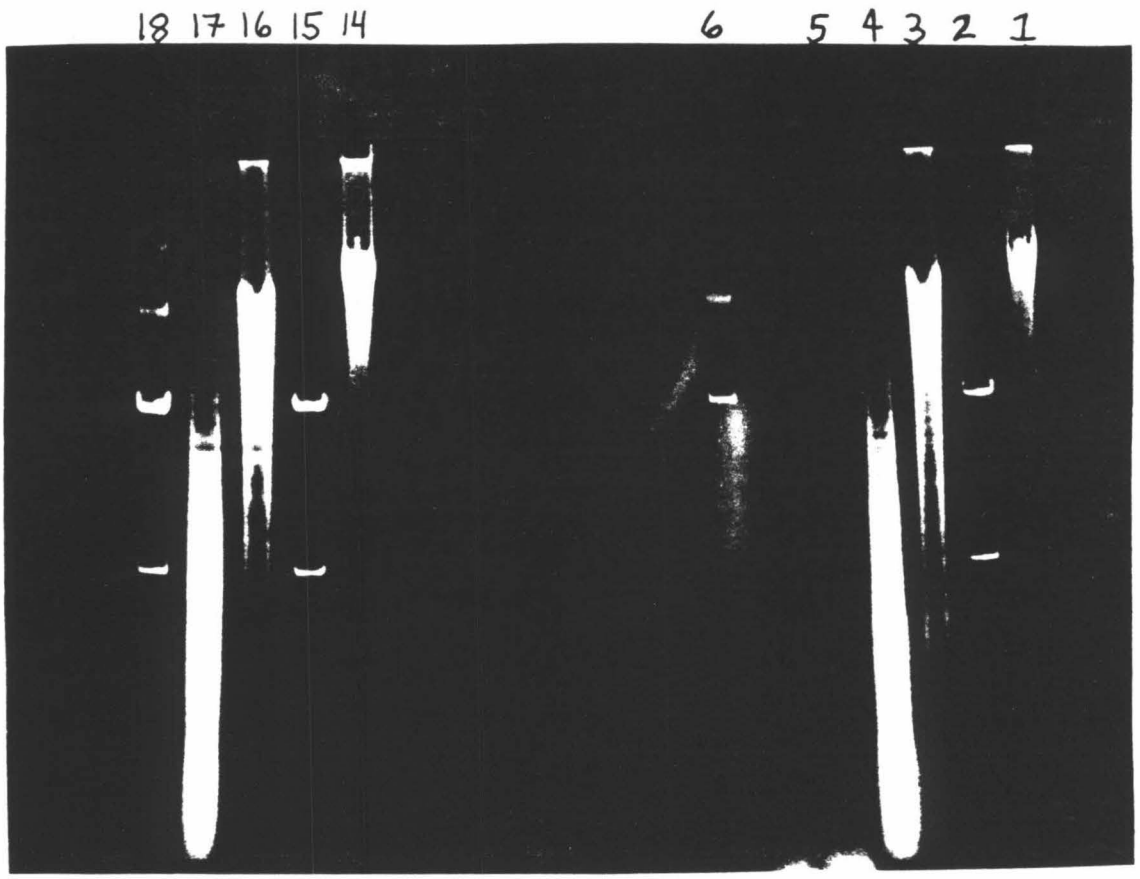


FIGURE 10

Autoradiogram of the hybridization of the 51-base probe to the agarose gel in figure 9.

LANES

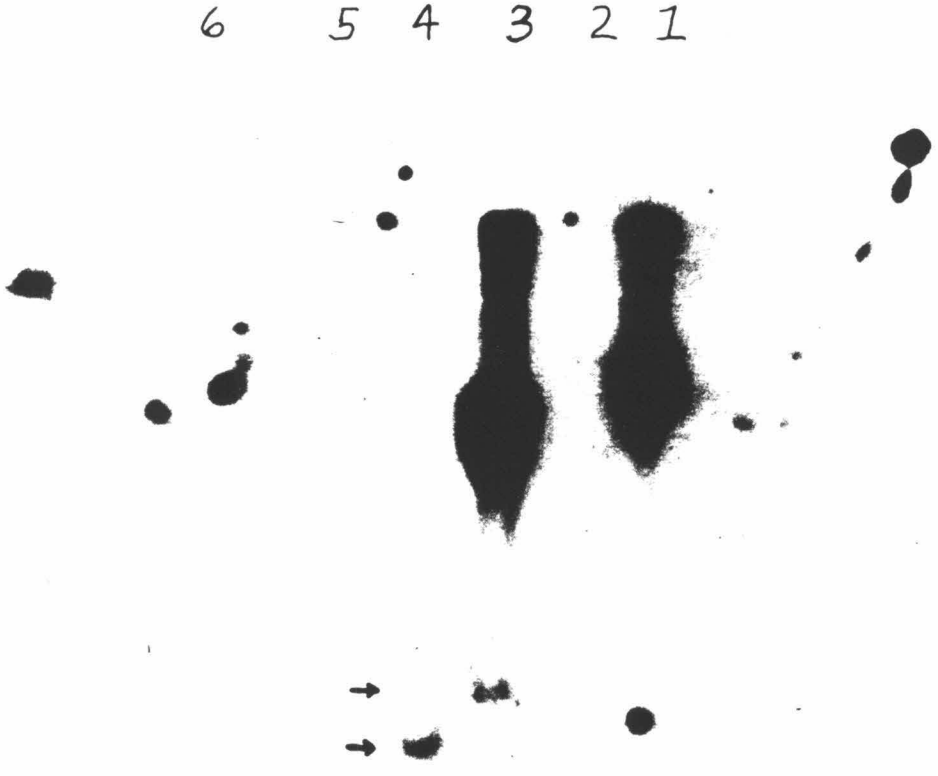


FIGURE 11

Partial amino acid and DNA sequence of the natural gene for α-lytic protease. The DNA sequence for the 51-base probe is below. Homologies of the probe DNA with the gene DNA is highlighted by an asterisk.

a.a. seq. GLY TYR GLN CYS GLY THR ILE THR ALA

DNA 5′ GGT TAC CAG TGC GGC ACC ATC ACC GCC

LYS ASN VAL THR ALA ASN TYR ALA GLU GLY

AAG AAC GTC ACC GCC AAC TAC GCC GAA GGT 3′

51-base probe 5′ TAC CCG TGC GGC ACC ATC ACC GCG AAG AAC

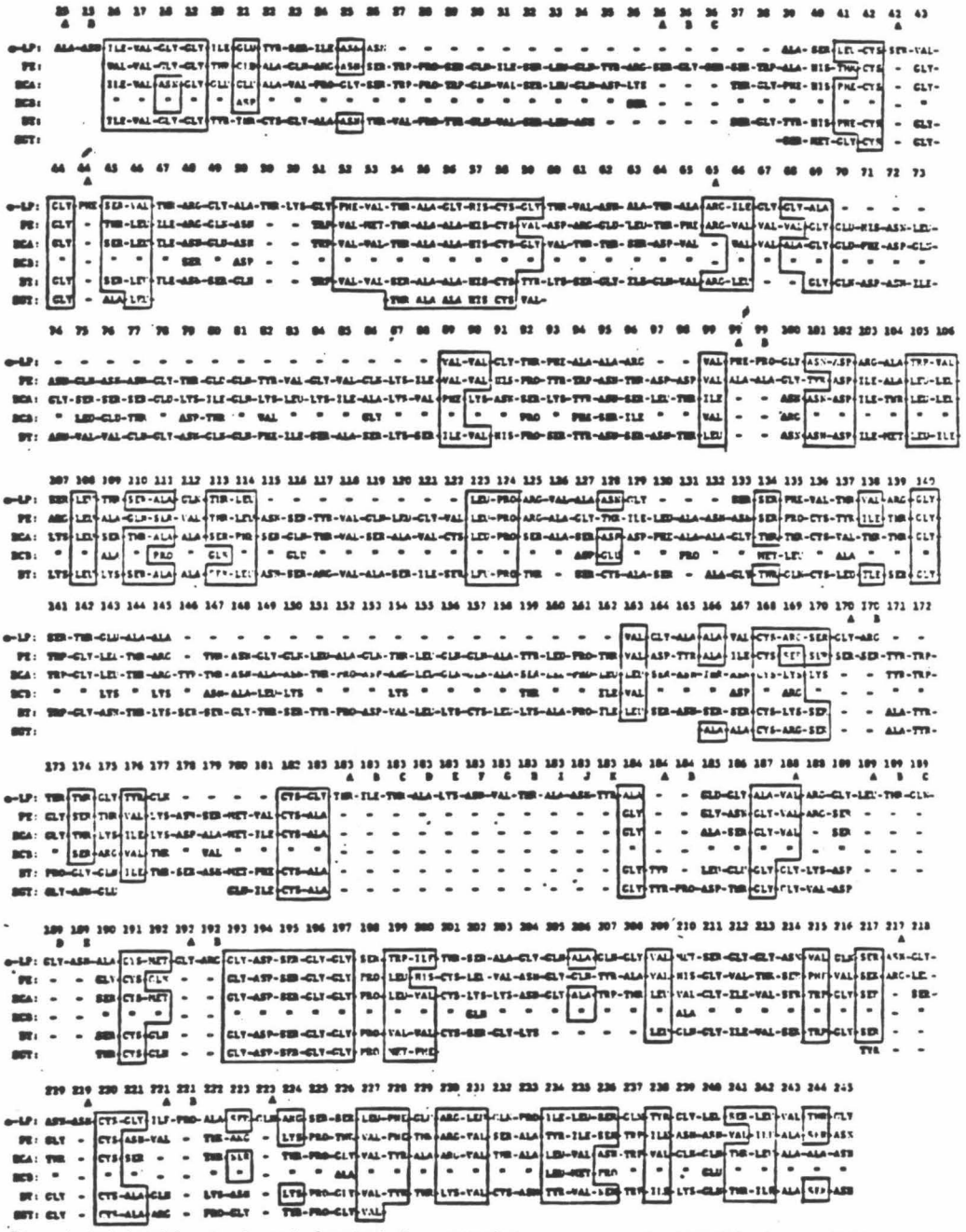
GTG ACC GCG AAC TAC GCG GAG 3′

Table 1. Amino-acid sequence of α-lytic protease (9).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	ALA	ASN	ILE	VAL	GLY	GLY	ILE	GLU	TYR	SER	ILE	ASN	ASN	ALA	SER	LEU	CYS	SER	VAL	GLY
21	PHE	SER	VAL	THR	ARG	GLY	ALA	THR	LYS	GLY	PHE	VAL	THR	ALA	GLY	HIS	CYS	GLY	THR	VAL
41	ASN	ALA	THR	ALA	ARG	ILE	GLY	GLY	ALA	VAL	VAL	GLY	THR	PHE	ALA	ALA	ARG	VAL	PHE	PRO
61	GLY	ASN	ASP	ARG	ALA	TRP	VAL	SER	LEU	THR	SER	ALA	GLN	THR	LEU	LEU	PRO	ARG	VAL	ALA
81	ASN	GLY	SER	SER	PHE	VAL	THR	VAL	ARG	GLY	SER	THR	GLU	ALA	ALA	VAL	GLY	ALA	ALA	VAL
101	CYS	ARG	SER	GLY	ARG	THR	THR	GLY	TYR	GLN	CYS	GLY	THR	ILE	THR	ALA	LYS	ASN	VAL	THR
121	ALA	ASN	TYR	ALA	GLU	GLY	ALA	VAL	ARG	GLY	LEU	THR	GLN	GLY	ASN	ALA	CYS	MET	GLY	ARG
141	GLY	ASP	SER	GLY	GLY	SER	TRP	ILE	THR	SER	ALA	GLY	GLN	ALA	GLN	GLY	VAL	MET	SER	GLY
161	GLY	ASN	VAL	GLN	SER	ASN	GLY	ASN	ASN	CYS	GLY	ILE	PRO	ALA	SER	GLN	ARG	SER	SER	LEU
181	PHE	GLU	ARG	LEU	GLN	PRO	ILE	LEU	SER	GLN	TYR	GLY	LEU	SER	LEU	VAL	THR	GLY		

Disulphide bridges: 17–87, 101–111, 137–170

Table 2. Comparison of amino-acid sequences of α-lytic protease (α-LP), bovine chymotrypsins A and B (BCA and BCB), bovine trypsin (BT) and Streptomyces griseus trypsin-like enzyme (SGT) (9).



The numbering is that of chymotrypsinogen A. Insertions in the sequences of other enzymes are numbered 36A, 36B and so on. Deletions are indicated by a dash. Only those residues of chymotrypsinogen-B which differ from chymotrypsinogen-A are indicated. Solid lines enclose residues in which chemically similar residues are present in α-lytic protease and at least two other enzymes. Chemically similar is defined in this paper as: Arg = Lys; Ala = Gly; Asp = Asn = Glu = Gln; Ser = Thr; Val = Trp = Ile = Leu = Tyr; Phe = Met. Homologous regions are as follows: all: PE, BCA, BCB, BT and SGT: 62-68, 166-182 and 191-220; PE, BCA, BCB and BT: 136-201; BT only: 228-257 and 178-232; BCA and BCB only 1-121.