SITE SATURATION MUTAGENESIS OF RESIDUE 71 IN ß-LACTAMASE

Thesis by

Steve Chester Schultz

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosphy

California Institute of Technology

Pasadena, California

1986

(Submitted April 14, 1986)

ACKNOWLEDGEMENT

I would like to thank my advisor, Dr. John H. Richards for his encouragement and support. His creativity and academic insight have influenced me tremendoulsy.

I wish to thank Suzanne Horvath for her guidance, help, and steadfast friendship during my years at Caltech. Special thanks also go to Gloria Dalbadie-McFarland and Jim Neitzel for their many hours of patient assistance and helpful discussions as well as for many hours of enjoyment in more social pursuits.

I wish to thank the Richards' group; I am deeply grateful for contributions by each member of this uniquely talented and supportive group. Along with Gloria and Jim, special thanks go to Steve Carroll, Sheila Iverson, Bill Healy, and Tom Chang.

I am also deeply grateful to Dr. Jerry R. Mohrig of Carleton College; his expert guidance, encouragement, and enthusiasm initiated my interest in research.

I would like to express my appreciation to Wayne McFarland for his friendship, support, and encouragement as well as for many enjoyable and fruitful social endeavors; to Brenda Baldwin-Zuschke for her friendship and support, especially during our early years at Caltech; and to Rosa Oefinger and Ruth (Cathy) Duborow for their friendship, encouragement and all the fun times.

Finally, I am deeply grateful to my family for their

support and encouragement: my mother and father, Ken and Wanda Schultz; my brothers Dan, Tim, and Jay; my sister Beth; my sister-in-law Darcey; and my nephews Nicholas and Matthew. Their love made it all possible.

My work was supported by NIH Grant GM 16424, NIH-NRSA Traineeship 5 T32 GMO 7616-08, and the William Barton Jones Fellowship.

SITE SATURATION MUTAGENESIS OF RESIDUE 71 IN ß-LACTAMASE

Abstract

A method was developed to "saturate" a particular residue in a protein with all possible amino acid substitutions. The method essentially involves introducing into the gene at the codon for the residue of interest a mixture of nucleotides: A,C,G, and T at the first two positions, C and G at the third position. This mixture of oligonucleotides can be readily introduced into the gene as a cassette when restriction sites are conveniently nearby. This method was used to saturate residue 71 in the pBR322 encoded ß-lactamase. The first step was introduction of restriction sites into pBR322 that flank the codon for residue 71: an Ava I site at 3972 and a Sca I site at 3937. The DNA fragment between these residues was removed and replaced with a mixture of synthetic double-stranded oligonucleotides that included 32 codons for residue 71 (codons for all 20 amino acids and the amber codon). The method was nearly 100% efficient for introducing this set of mutations and the observed distribution of codons indicates that the various codons were introduced at equal frequencies.

Class A ß-lactamases, which include the pBR322 encoded enzyme, contain a conserved triad (Ser-Thr-x-Lys) at the catalytic site; this site includes Ser 70 whose hydroxyl

group opens the ß-lactam ring. This study focuses on the conserved Thr 71, whose role in the activity of ß-lactamase was previously unknown. Mutants with all 19 possible amino acid substitutions at residue 71 were generated as described above and characterized by the in vivo penam and cephem antibiotic resistance they provided to E. coli LS-1 cells. Surprisingly, cells producing any of 14 of the mutant enzymes display appreciable resistance to ampicillin; only cells with mutants having Tyr, Trp, Asp, Lys, or Arg at residue 71 had no observable resistance to ampicillin. However, all of the mutants are less stable to cellular proteases than the wild type enzyme. These results suggest that Thr 71 is not essential for binding or catalysis but is important for stability of the ß-lactamase protein. An apparent change in the substrate specificity of the various mutant enzymes indicates that residue 71 influences the region of the protein that accommodates the side chain attached to the ß-lactam ring of the substrate.

To study the role of the naturally occurring disulfide bond in stabilizing ß-lactamases that contain mutations at residue 71, a Cys 77 --> Ser mutation was introduced into all 19 of these mutants. The resulting doubly mutant proteins were analyzed for their ability to confer resistance to ampicillin in vivo and the relative quantities of the mutant enzymes in whole cell extracts were analyzed by antibody stains of electrophoretic blots. Although the activity and stability of the single mutant Cys 77 --> Ser is indis-

tinguishable from the wild type enzyme in physiological conditions, the doubly mutant proteins are much less stable than the analagous single mutants that contain a disulfide bond. Apparently, the presence of the disulfide bond in ß-lactamase enables the enzyme partially to overcome instabilities arising from mutations at residue 71. The results presented here also demonstrate that Thr at residue 71 performs a unique and essential role in the stability of ß-lactamase.

TABLE OF CONTENTS

LIST OF FIGURES

APPENDIX II

LIST OF TABLES

NOMENCLATURE

Expressions of the form Thr 71 --> Ser indicate that a mutation that replaces threonine at residue 71 with serine has been introduced.

INTRODUCTION

INTRODUCTION

Proteins can perform many diverse functions because of their ability to assume unique structures that position various chemical groups into precisely defined three-dimensional arrangements such that these groups can interact with other molecules in very specific ways. However, precise mechanisms by which proteins can fold to generate unique structures remain unclear as do the precise interactions by which the resulting arrangement of chemical groups can give rise to a particular function. Discovering the underlying principles of the relationships between structure and function in proteins is currently among the most basic and interesting areas of protein biochemistry.

Biophysical techniques such as protein crystallography, nuclear magnetic resonance spectroscopy, circular dichroism, and florescence spectroscopy provide many insights into structure-function relationships in proteins. By using such techniques to examine thoroughly particular structures and/or functions, we can learn what conformations a protein assumes, which functional groups interact with other groups, and what chemical pathways a reaction utilizes. Although such studies can elucidate many aspects of the relationships between structure and function, several features cannot be studied by investigating only existing structures.

Methods for producing specific changes in protein structure would complement such biophysical studies by allowing one to investigate how changes in structure affect function.

Until recently, obtaining stuctural variants of proteins was difficult; one could isolate naturally occurring variants from different organisms or individuals (as for cytochrome c and hemoglobin) or produce structural variants by chemical modification or random mutagenesis. Obtaining specified structural variants was often impossible using these methods. Presently, however, by utilizing modern techniques for site-specific in vitro mutagenesis, one can produce any of 19 possible structural variants at any residue in a protein. This ability to produce any desired mutation allows a rational study of the relationships between structure and function in proteins. Just as organic chemists could rationally study the nature of carbocations by synthesis and characterization of many different molecules and, likewise, study factors influencing specificity in electrophilic aromatic substitution by synthesis and characterization of numerous aromatic structures, site-specific mutagenesis allows protein biochemists to produce and characterize large numbers of structural variants to learn what basic principles govern structure and function in proteins.

Since the introduction of site-specific mutagenesis as an approach for studying structure-function relationships in 1982 (1,2), many such studies have been reported. This approach has been used to study the roles of specific residues in catalysis for tyrosyl-tRNA synthetase (2,3,4), alkaline phosphatase (5), lac carrier protein (6), cytochrome c (7), aspartate transcarbamoylase (8), triose phosphate iso-

merase (9), and ß-lactamase (1,10); to alter substrate specificity and affinity for trypsin (11) and tyrosyl-tRNA synthetase (12); to change the pH dependence of activity for subtilisin (13); to study protein folding for cytochrome c (14) and dihydrofolate reductase (15); to study the stabilization of proteins by disulfide bonds for lysozyme (16,17), interleukin-2 (18,19), dihydrofolate reductase (15), and ß-lactamase (20); to alter protein-protein interactions for tyrosyl-tRNA synthetase (21); and to study protein-DNA interactions for bacteriophage 434 repressor (22) and bacteriophage P22 mnt repressor (23).

However, a large number of possible amino acid substitutions exist. For a protein with 200 amino acid residues, 3,800 possible single amino acid subtitutions exist; 14,436,200 possible combinations of two amino acid substitutions exist. Consequently, rational studies of structural variants of proteins require high levels of theory to predict which structures should be studied. However, prejudices necessarily resulting from the need to identify interesting mutations prior to their study may exclude certain mutations from studies that might generate new insights into structure-function relationships in proteins. Moreover, since most of the principles governing structure and function are uncertain or unknown, the effects of specific amino acid substitutions usually are difficult to predict. Due to these limitations, studies of mutant proteins are restricted to well characterized systems for which insights are available from studies

such as x-ray crystallography, computer modeling, sequence homologies, and catalytic mechanisms; even then, surprising results often arise.

Circumventing these limitations requires efficient methods for production and characterization of large numbers of structural variants. For characterization of such structural variants, if the protein in question affects the phenotype of cells, one can easily determine the activities of large numbers of mutations by screening cells for a particular function. A phenotypic screen is particularly effective if the protein is essential for cell growth; active mutants can then be identified from among 10^6 or more inactive mutants.

For producing large numbers of structural variants, several methods have been introduced. One approach is to make site-specific mutagenesis more general by using mixtures of oligonucleotides in procedures such as oligonucleotide directed mutagenesis to produce a specified set of mutations. Several examples exist, including: production of 11 different mutations at five sites in the gene for the simian virus 40 large-T antigen using a mixture of mutagenic primers that contained 1024 different sequences (24); production of up to five different mutations at codon 12 in the c-Ha-*ras*1 gene by using mixtures of up to 8 different mutagenic primers (25); and production of five different mutations in the fibroin gene promoter by using a mixture of six different mutagenic primers (26). Although such methods can be used to

generate small sets of mutations, limitations arising from low and variable efficiencies make production of large sets of mutations difficult.

The opposite approach is to make random mutagenesis more specific. Several methods of targeting random mutagenesis to specific areas in a gene have been reported (27). More recently, long synthetic oligonucleotide segments containing low levels of random base mismatches have been used to introduce one or two base mutations at high freqencies. For example, 132 single base mutations were generated in the mouse ß-globin promoter region using the following procedure (28): a single stranded plasmid vector that contained the promoter region was mutagenized using various chemicals; an oligonucleotide primer was extended across this region of the plasmid using avian myeloblastosis virus reverse transcriptase to produce DNA segments with incorrect bases; the desired fragment was then isolated from appropriate restriction digests and inserted into a new host vector; the resulting plasmids were used to transform cells directly, or DNA fragments were enriched for single base mismatches and then reincorporated into the plasmid. In another method, oligonucleotide segments were synthesized to contain low levels of single base mismatches and ligated into the region preceding the iso-1 cytochrome c gene to produce various single and double base mutations (29). Although these methods produce large numbers of mutations, they do not explore all possible amino acid substitutions and require a very efficient selec-

table phenotype.

An efficient method for generating a desired set of many known mutations would provide an approach intermediate to the two methods discussed above and, therefore, would allow a complete set of desired structures to be investigated. To study thoroughly a residue of particular interest, one could "saturate" a specific residue with all possible amino acid substitutions. A method for site saturation, whereby one produces and characterizes structural variants with all 19 amino acid substitutions at a specific residue, would enhance our ability to study rationally the relationships between structure and function in proteins.

If a procedure exists for efficiently generating the appropriate mutants and the protein in question affects the phenotype of cells such that one can easily screen for mutants that perform a particular function, one could then extend this approach to simultaneous saturation of two or more sites to assess the effects of combinations of residues on protein function.

Site saturation can be accomplished by introducing into the gene at the codon for the residue of interest a mixture of nucleotides: A,C,G, and T at the first two positions; C and G at the third position. The resulting mixture of 32 oligonucleotides, which codes for all 20 amino acids and the amber codon, can be readily introduced into the gene as a "cassette" when restriction sites are conveniently nearby. If desired, appropriate mixtures of oligonucleotides

can be used to encode a particular subset of amino acids.

I have used site saturation to study the role of Thr 71 in the pBR322 encoded ß-lactamase (EC 3.5.2.6) (30). ß-Lactamase catalyzes the hydrolysis of the ß-lactam ring of penam and cephem antibiotics (31). This activity confers resistance to these antibiotics on cells that produce the enzyme and thereby provides a convenient screening procedure to assess the activities of mutant ß-lactamases.

The class A ß-lactamases (32) (which includes the pBR322 encoded enzyme) contain a conserved triad (SerThr-x-Lys) (32) which includes Ser 70, whose hydroxyl group opens the ß-lactam ring (33,34). This study focuses on the conserved Thr 71, whose role in the activity of ß-lactamase was previously unknown.

References

1. Dalbadie-McFarland, G., Cohen, L.W., Riggs, A.D., Morin, C., Itakura, K. & Richards, J.H. (1982) Proc. Natl. Acad. Sci. USA 79, 6409-6413.
2. Winter, G., Fersht, A.R., Wilkinson, A.J., Zoller, M. & Smith, M. (1982) Nature (London) 299, 756-758.
3. Wilkinson, A.J., Fersht, A.R., Blow, D.M., Carter, P. & Winter, G. (1984) Nature (London) 307, 187-188.
4. Leatherbarrow, R.J., Fersht, A.R. & Winter, G. (1985) Proc. Natl.Acad. Sci. USA 82, 7840-7844.
5. Soumitra, S.G., Bock, S.C., Rokita, S.E., Kaiser, E.T. (1986) Science 231, 145-148.
6. Trumble, W.R., Viitanen, P.V., Sarkar, H.K., Poonian, M.S. & Kaback, R.H. (1984) Biochem. Biophys. Res. Comm. 119, 860-867.
7. Pielak, G.J., Mauk, A.G. & Smith, M. (1985) Nature (London) 313, 152-154.
8. Robey, E.A. & Schachman, H.K. (1984) J. Biol. Chem. 259, 11180-11183.
9. Straus, D., Raines, R., Kawashima, E., Knowles, J.R. & Gilbert, W. (1985) Proc. Natl. Acad. Sci. 82, 2272-2276.
10. Dalbadie-McFarland, G., Neitzel, J. & Richards, J.H. (1986) Biochemistry 25, 332-338.
11. Craik, C.S., Largman, C., Fletcher, T., Roczniak, S., Barr, P.J., Fletterick, R. & Rutter, W.J. (1985) Science 228, 291-297.

12. Wilkinson, A.J., Fersht, A.R., Blow, D.M., Carter, P. & Winter, G. (1984) Nature (London) 307, 187-188.

13. Thomas, P.G., Russell, A.J. & Fersht, A.R. (1985) Nature (London) 318, 375-376.

14. Ernst, J.F., Hampsey, D.M., Stewart, J.W., Rackovsky, S., Goldstein, D. & Sherman, F. (1985) J. Biol. Chem. 260, 13225-13236.

15. Villafrance, J.E., Howell, E.E., Voet, D.H., Strobel, M.S., Ogden, R.C., Abelson, J.N. & Kraut, J. (1983) Science 222, 782-788.

16. Perry, L.J. & Wetzel, R. (1984) Science 226, 555-557.

17. Perry, L.J. & Wetzel, R. (1986) Biochemistry 25, 733-739.

18. Wang, A., Lu, S.-D. & Mark, D.F. (1984) Science 224, 1431-1433.

19. Liang, S.-M., Thatcher, D.R., Liang, C.-M. & Allet, B. (1986) J. Biol. Chem. 261, 334-337.

20. Schultz, S.C., Dalbadie-McFarland, G., Neitzel, J.J. & Richards, J.H. (1986) manuscript in preparation.

21. Jones, D.H., McMillan A.J. & Fersht, A.R. (1985) Biochemistry 24, 5852-5857.

22. Wharton, R.P. & Ptashne, M. (1985) Nature (London) 316, 601-605.

23. Youderian, P., Vershon, A., Bouvier, S., Sauer, R.T. & Susskind, M.M. (1983) Cell 35, 777-783.

24. Kalderon, D., Richardson, W.P., Markham, A.F. & Smith, A.E. (1984) Nature (London) 311, 33-38.

25. Seeburg, P.H., Colby, W.W., Capon, D.J., Goeddel, D.V. &

Levinson, A.D. (1984) Nature (London) 312, 71-75.

26. Hirose, S., Takeuchi, K., Hori, H., Hirose, T., Inayama, S. & Suzuki, Y. (1984) Proc. Natl. Acad. Sci. USA 81,1394-1397.
27. Dalbadie-McFarland, G. & Richards, J.H. (1983) Annu. Rep. Med. Chem. 18, 237-245.
28. Myers, R.M., Lerman, L.S. & Maniatis, T. (1985) Science 229, 242-247.
29. McNeil, J.B. & Smith, M. (1985) Mol. Cell. Biol. 5, 3545-3551.
30. Schultz, S.C. & Richards, J.H. (1986) Proc. Natl. Acad. Sci. USA 83, 1588-1592.
31. Datta, N. & Kontomichalou, P. (1965) Nature (London) 208, 239-241.
32. Ambler, R.P. (1980) Phil. Trans. R. Soc. London Ser. B 289, 321-331.
33. Fisher, J., Belasco, J.G., Khosla, S. & Knowles, J.R. (1980) Biochemistry 19, 2895-2901.
34. Knott-Hunziker, V., Waley, S.G., Orlek, B.S. & Sammes, P.G. (1979) FEBS Lett. 99, 59-61.

CHAPTER I

A GENERAL METHOD FOR "SATURATION" OF A RESIDUE IN A PROTEIN WITH ALL 19 POSSIBLE AMINO ACID SUBSTITUTIONS

Introduction

For site saturation studies of the relationship between structure and function in proteins, one must first generate mutants with all 19 amino acid substitutions at a specific residue in a protein by introducing into the gene at the codon for the residue of interest a mixture of nucleotides that code for all 19 amino acids (or a subset of these 19 amino acid substitutions if desired). The mutagenic technique used to introduce the necessary mixture of nucleotides must fulfill two requirements. First, the technique must generate the specified mutations at nearly 100% efficiency; the presence of non-mutated or non-specified sequences will greatly increase the number of clones that one must analyze. Second, the mutagenic technique must produce the various individual mutations at nearly equal frequencies; non-equivalent frequencies will result in underrepresented codons and, therefore, amino acid substitutions that may go untested in the experiments. If one could efficiently generate all 19 amino acid substitutions at equal frequencies, site saturation could be used routinely to study the roles performed by particular residues in the structure and function of proteins.

This approach to studies of structure-function relationships in proteins becomes particularily appealing when the protein in question affects the phenotype of cells such that one can easily screen for mutants that perform a particular function. With an efficient technique for introducing the

desired set of mutations, site saturation could then be extended to simultaneous saturation of two or more residues to assess the effects of combinations of residues on protein function.

In an approach commonly used for cloning and subcloning genes, mixtures of many fragments generated by digesting genomic or plasmid DNA with restriction enzymes are ligated into specific restriction sites of an appropriate DNA vector. To separate individual sequences from the large mixture of vector constructions that result from such a ligation, the sequences are incorporated into cells by transformation procedures such that individual cells receive only one vector construction. By spreading an appropriate concentration of the transformation mixture onto an agar plate, single cells will give rise to isolated colonies, each of which will contain a unique vector construction.

This approach could be used to efficiently generate specified mutations if restriction sites exist in convenient locations in the structural gene for a protein of interest. The DNA segment between such sites could be removed and replaced with a synthetic double stranded oligonucleotide that introduces a desired mutation. Such studies have recently been reported (1,2). This technique had also been used to introduce sets of mutations into genes (3,4,5). Such an approach could be used for site saturation to introduce into the gene at the codon for the residue of interest a mixture of nucleotides that codes for all 20 amino acids. If

the mixture contains equimolar amounts of the various codons, this method should efficiently generate all specified mutations at equal frequencies.

Using this approach, I have saturated residue 71 in the pBR322 encoded ß-lactamase with all possible amino acid substitutions (6). The significance of residue 71 will be discussed in Chapter II.

Since convenient restriction sites did not exist in pBR322, I first introduced restriction sites that flank the codon for Thr 71 by oligonucleotide-directed mutagenesis: an Ava I site at 3972 and a Sca I site at 3937. The DNA fragment between these sites was removed and replaced with a mixture of double-stranded oligonucleotide segments. The segment was synthesized to contain a mixture of nucleotides at the codon for residue 71: A,C,G, and T at the first two positions and C and G at the third position. This combination of sequences contains 32 different codons that code for all 20 amino acids and the amber codon. The resulting mixture of plasmids was used to transform E. coli and plasmid DNA was isolated from 108 colonies. In this collection of plasmids, codons for all twenty amino acids and an amber codon were observed. Thirty of the thirty-two codons were obtained; the distribution of observed codons was similar to that predicted by a Poisson distribution based on the introduction of the various codons at equal frequencies.

Since beginning our site saturation experiments, two related studies have also been published. For the c-Ha-ras1

gene product, eighteen mutations (the Gly 12 --> Val mutant was already available) were produced by oligonucleotide-directed mutagenesis using a mixture of mutagenic primers (7). Proteins with all of the possible amino acid substitutions at residue 12 were analyzed for their ability to induce morphological transformations in Rat-1 cells. In the second study, mutant proteins with all 19 amino acid substitutions at residue 222 in subtilisin were produced by an approach somewhat similar to ours. In four separate reactions, mixtures of oligonucleotides each containing five different codons for residue 222 were ligated into the gene (8). The ratios of the various mutants obtained from each of the four pools of oligonucleotides were severely skewed. Two of the twenty codons were not obtained; these were later produced separately. Also, 30% of the characterized plasmids contained unexpected mutations. Our procedure is more reliable and general than either of these and, therefore, could be reliably used for simultaneously saturating multiple sites in a protein.

Materials & Methods

Enzymes and Chemicals

Restriction enzymes and the large (Klenow) fragment of DNA polymerase I were purchased from Boehringer Mannheim. The T4 DNA ligase and T4 polynucleotide kinase were obtained from Bethesda Research Laboratories. All antibiotics were from Sigma Chemical Co. The [α-^{32}P]dTTP, 3000 μCi/mmol (1 Ci = 37 GBq), was purchased from Amersham and the [γ-^{32}P]ATP was from ICN.

DNA

Oligonucleotides were synthesized by using the phosphoramidite chemistry (9) on the Applied Biosystems (Foster City, CA) DNA synthesizer, model 380A and purified by preparative polyacrylamide gel electrophoresis.

Bacteriophage M13 mp8 was grown in E. coli JM 103 (10) in 2 X YT medium. The DNA was prepared according to standard procedures (11). Plasmid pBR322 was grown in E. coli LS1 (12) in LB medium (13) and prepared according to standard procedures (14). Numbering of the pBR322 is the one commonly used (13).

DNA fragments were purified from 1.2% agarose gels using either DEAE-cellulose or an International Biotechnologies, Inc. (New Haven, CT) UEA electroelutor.

Reactions

Individual oligonucleotide strands were phosphorylated using standard procedures (13).

Oligonucleotide strands were annealed by mixing 0.4

pmole/μl of each strand and heating to 95°C in 10 m$\underline{M}$ $MgCl_2$ / 50 m$\underline{M}$ Tris-HCl, pH 8, followed by gradual cooling to 20°C over a 45 minute period.

For ligations, approximately 0.04 pmole of each restriction fragment and 0.4 pmole of the synthetic fragment, were mixed in 10 m$\underline{M}$ $MgCl_2$ / 50 m$\underline{M}$ Tris-HCl, pH 8 / 0.5 m$\underline{M}$ ATP / 5 m$\underline{M}$ dithiothreitol and 10 units of T4 DNA ligase and incubated at 15°C for 16 hours. The reaction mixture was extracted with phenol and the DNA was precipitated by addition of ethanol. The DNA was redissolved in 20 μl of 1 m$\underline{M}$ Tris-HCl, pH 8 / 0.1 m$\underline{M}$ EDTA; a 5 μl aliquot was used to transform competent $\underline{E. coli}$ LS1 cells. Cell transformations were done using standard procedures (15).

$\underline{Mutagenesis}$

Oligonucleotide-directed mutagenesis was carried out on M13 mp8 containing the 752-base-pair Pst I to EcoRI fragment of pBR322. The following oligonucleotides were used:

Ava I 5'-ACG TTC CTC GGG GCG

Sca I 5'-ACA TAG CAG TAC TTT

The procedures used for the mutagenesis (see Figure 1) are described elsewhere (16,17,18). To remove the preexisting Sca I site at 3846, the 61 base pair segment between HincII at 3907 and Sca I at 3846 in pBR322 was removed and replaced with the following synthetic segment using the approach presented in Figure 2.

```
5'-GAC GCC GGG CAA GAG CAA CTC GGT CTC CGC ATA CAC TAT TCT-
   CTG CGG CCC GTT CTC GTT GAG CCA GAG GCG TAT GTG ATA AGA-

   CAG AAT GAC TTG GTT GAA T
   GTC TTA CTG AAC CAA CTT A-5'
```

To introduce the 20 amino acids at residue 71, the following mixture of oligonucleotides was ligated into the plasmid as shown in Figure 3.

```
                                        [AA ]
5'-CC GAG GAA CGT TTT CCA ATG ATG AGC   [CCC]  TT*C AAA GT
        C CTT GCA AAA GGT TAC TAC TCG   [GGG]  AAG TTT CA-5'
                                        [TT ]
```

The asterisk indicates a silent mutation that removes a Dra I site at 3943 and also decreases the self-complementarity of 6 (and possibly as many as 10) bases near this end of the individual strands. After ligation and transformation as described above, 1/10 of the cells from the transformation mixture were plated on tetracycline. Of the 423 resulting colonies, 105 were selected randomly and plasmids derived from these colonies were sequenced.

DNA Sequencing

The plasmids were digested with Ava I and the resulting 2962 base pair fragments were isolated as described above. The fragment was labeled at nucleotide T-3973 using [α-^{32}P]dTTP and the Klenow fragment of DNA polymerase I. This approach is shown in Figure 4. The labeled fragments were precipitated by addition of ethanol and pelleted by centrifugation; the pellet was washed with 70% ethanol and dried at reduced pressure. The labeled fragments were sequenced using standard techniques (19). The sequencing

gels were 40 cm long and contained either 8% or 20% polyacrylamide (wt/vol) and 50% (wt/vol) urea and were run using standard procedures (13).

Results

Restriction sites that could be introduced near residue 71 without altering the amino acid sequence of ß-lactamase were identified by reverse translating the amino acid sequence in this region of the protein and searching for restriction sites in the resulting degenerate sequence (see Figure 5). The Ava I and Sca I sites were closest to residue 71 and occurred in only one other place in pBR322. Introduction of both sites require only single base changes.

Introduction of the Ava I and Sca I sites and removal of the Sca I site at 3846 were verified by restriction mapping of the plasmids. The absence of other changes in the ß-lactamase gene was verified by sequencing one of the mutants from the Pst I site through the residues of the structural gene encoding the protein N terminus; this was the portion of the gene subcloned into M13 mp8 for the oligonucleotide-directed mutagenesis.

To determine whether the oligonucleotides were properly annealed by the procedure described in the Materials & Methods section, the annealed fragments were run on a nondenaturing polyacrylamide gel (i.e., a polyacrylamide gel that contained no urea). One dark band was observed in the expected position; no other bands were observed.

To optimize conditions for ligating the synthetic fragment into the plasmid, several concentrations of synthetic fragment were used in normal ligation reactions and the resulting segments were analyzed by electrophoresis in

agarose gels. Concentrations of 4 pmole/50 μl substantially inhibited the reaction and 40 pmole/50 μl resulted in complete inhibition; concentrations of 0.08 pmole/50 μl or less gave little or no ligation; 0.4 pmole/50 μl produced the maximum amount of ligation. Subsequently, all reactions were done using 0.4 pmole of the synthetic segment and 0.04 pmoles of plasmid fragments (less by a factor of 10 than the amount of synthetic segment) for a 50 μl reaction. I also tested reactions in which the 5'-end of the shorter oligonucleotide strand (the blunt end of the double stranded fragment) was either phosphorylated or non-phosphorylated; after ligation, no detectable difference was observed on agarose gels. Moreover, similar numbers of cells were obtained after transformations using DNA from either of these two reactions.

After ligation of the mixture of oligonucleotides into the plasmid, transformation of competent E. coli LS1 cells gave 423 colonies that were resistant to tetracycline. Transformation of cells with plasmid from a parallel ligation in which the synthetic segment was omitted produced only four colonies indicating that only plasmids containing the synthetic insert were able, after transformation, to confer tetracycline resistance. Plasmid DNA derived from 105 of the colonies (plasmid from 3 of the 108 original colonies did not label properly) was sequenced from Ava I (3972) through Sca I (3937). Table I lists the codon sequences for each of these 105 plasmids. Table II lists the codon frequencies in this collection of mutations.

Discussion

The three fragment ligation shown in Figure 3 was designed to avoid two problems. First, an Ava I site occurs at nucleotide 1424 in pBR322 and, consequently, digestion of the plasmid in which the new Ava I site is present would result in cleavage at both the 1424 and 3972 Ava I sites; ligation of these fragments would result in undesirable plasmid constructions. Second, the method presented in Figure 3 will eliminate any background from wild type colonies that could arise due to incomplete digestion at the Ava I or Sca I sites. Purifying the fragments shown in this figure completely eliminates the presence of the 35-base-pair segment between the Ava I and Sca I sites from the ligation reaction. Whereas transformation of competent cells with DNA from the ligation reaction shown in Figure 3 resulted in 423 tetracycline resistant colonies, a parallel ligation in which the synthetic segment was omitted resulted in only 4 colonies indicating that only plasmids that had received the synthetic insert were able, after transformation, to confer tetracycline resistance. These results indicate that this method is highly efficient for introducing the synthetic fragments without a significant background of wild type plasmids.

The Ava I site introduced at 3972 is a nonpalindromic recognition sequence. The use of nonpalindromic sequences is desirable since the four bases that protrude on the 5' end are not complementary to themselves; therefore, polymers of the synthetic sequence cannot form at this end of the segment

and these synthetic segments can ligate efficiently only to the fragment from pBR322 as desired. The blunt end arising from cleavage by Sca I is not optimal for ligating into the plasmid since blunt end ligations are not as efficient as ligations with complementary protruding ends. However, using the next potential restriction site in pBr322 near the region of interest would require an insert that is 25 bases longer than the segment required when using the Sca I site. Fortunately, the results confirm that the efficienies of blunt end ligations are sufficient for our purposes.

Importantly, for site saturation, one must generate all specified codons at nearly equal frequencies. If codons for certain amino acids are present at low frequencies, the corresponding mutants will be underrepresented, difficult to find, and may go untested in the experiment. Nearly equimolar mixtures of oligonucleotides can be obtained by careful synthesis (20,21) and then used in cassette mutagenesis to produce the various mutations at equal frequencies.

In a Poisson distribution, the probability of finding a particular codon at least once among 101 colonies is 96% (4 of the 105 clones that were sequenced contained single base insertions at the codon for residue 71 as discussed below). As shown in Table II, in the experiments described here 94% (30/32) of the codons were observed one or more times. The absence of two codons, therefore, was probably the consequence of a small sample size. The probability of finding a codon between 2 and 4 times in a sample of 101 is predicted

to be 61%; in this experiment, 56% of the codons were observed between 2 and 4 times.

Table III lists the observed frequencies of the individual bases at each position in the codon for residue 71. The C and G nucleotides occur somewhat less frequently than expected at the first position resulting in a slight skewing of the codon frequencies (see Table II). We observed a nearly equimolar mixture of bases at the other positions. Thus, the observed codon frequencies correlate well with the expected values; this demonstrates that under the conditions of the oligonucleotide synthesis, the four nucleoside phosphoramidites couple with almost equal efficiencies. Recently, the relative coupling efficiencies of phosphoramidite derivatives of deoxynucleosides have been studied in detail (21); this study demonstrates that freshly dissolved methyl-N,N-diisopropylamino phosphoramidite derivatives couple at nearly equivalent frequencies. However, two days after being dissolved, the guanosine derivative becomes significantly less reactive; this result stresses the importance of using fresh, pure reagents in the synthesis.

In sequencing plasmids from 105 individual colonies, we found none in which the 35 base pair segment contained unexpected base substitutions. However, fourteen frameshift mutants resulting from single base deletions or additions were observed. Four mutants had one extra base (TGCC, GATC, CACC, CGGC) in the region corresponding to the codon for residue 71; one mutant had a single base addition within the

synthetic fragment not at residue 71; three mutants had single base deletions at the Sca I site; six mutants had single base additions and one a single base deletion near the Ava I site. The nine additions and deletions at restriction sites may occur during processing of the synthetic fragment or during ligation; similar deletions have been observed, for example, in ligation of a synthetic fragment into the Pst I site of the gene for subtilisin (8). The other five mutations that occur within the synthetic fragment probably result from imperfections in the DNA synthesis.

Conclusion

The approach described here will generate mutations that result in all possible amino acid substitutions in a protein at a residue of particular interest. This method was used to efficiently introduce 32 different codons at residue 71 in ß-lactamase. The distribution of codons observed in 101 colonies indicated that the various codons were introduced at equal frequencies. As an approach to studies of the relationships between structure and function in proteins, this method becomes particularly powerful for proteins that affect cell phenotype such that one can easily identify and select active mutants. One could then extend these studies to the saturation of multiple sites to assess the effect of combinations of residues on protein function. Presently, no other reported methods provide such an efficient and general technique for site saturation. Also, of the published methods, ours is the only one that could be used to saturate multiple residues.

References

1. Lo, K.-M., Jones, S.S., Hackett, N.R. & Khorana, H.G. (1984) Proc. Natl. Acad. Sci. USA 81, 2285-2290.
2. Wharton, R.P., Brown, E.L. & Ptashne, M. (1984) Cell 38, 361-369.
3. Murphy, M.H. & Baralle, F.E. (1983) Nucleic Acids Res. 11, 7695-7700.
4. Baell, G., Schulz, M.-F., Selzer, G., Chollet, A. & Movva, N. (1985) Nucleic Acids Res. 13, 1923-1938.
5. Hui, A., Hayflick, J., Dinkelspiel, K. & DeBoer, H.A. (1984) EMBO J. 3, 623-629.
6. Schultz, S.C. & Richards, J.H. (1986) Proc. Natl. Acad. Sci. USA 83, 1588-1592.
7. Seeburg, P.H., Colby, W.W., Capon, D.J., Goeddedl, D.V. & Levinson, A.D. (1984) Nature (Lodon) 312, 71-75.
8. Wells, J.A., Vasser, M. & Powers, D.B. (1985) Gene 34, 315-323.
9. Beaucage, S.L. & Caruthers, M.H. (1981) Tetrahedron Lett. 22, 1859-1862.
10. Messing, J., Crea, R. & Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
11. Messing, J. (1983) Methods in Enzymol. 101, 20-89.
12. Boyer, H.W. & Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.
13. Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

14. Ish-Horowicz, D. & Burke, J.F. (1981) Nucleic Acids Res. 9, 2989-2998.

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

16. Norris, K., Norris, F., Christiansen, L. & Fiil, N. (1983) Nucleic Acids Res. 11, 5103-5112.

17. Zoller, M.J. & Smith, M. (1983) Methods in Enzymol. 100, 468-500.

18. Dalbadie-McFarland, G. (1985), Ph.D. Thesis, California Institute of Technology.

19. Maxam, A.M. & Gilbert, W. (1980) Methods in Enzymol. 65, 499-560.

20. Suzanne Horvath, personal communication.

21. Zon, G., Gallo, K.A., Samson, C.J., Shao, K.L., Summers, M.F. & Bryd, R.A. (1985) Nucleic Acids Res. 13, 8181-8196.

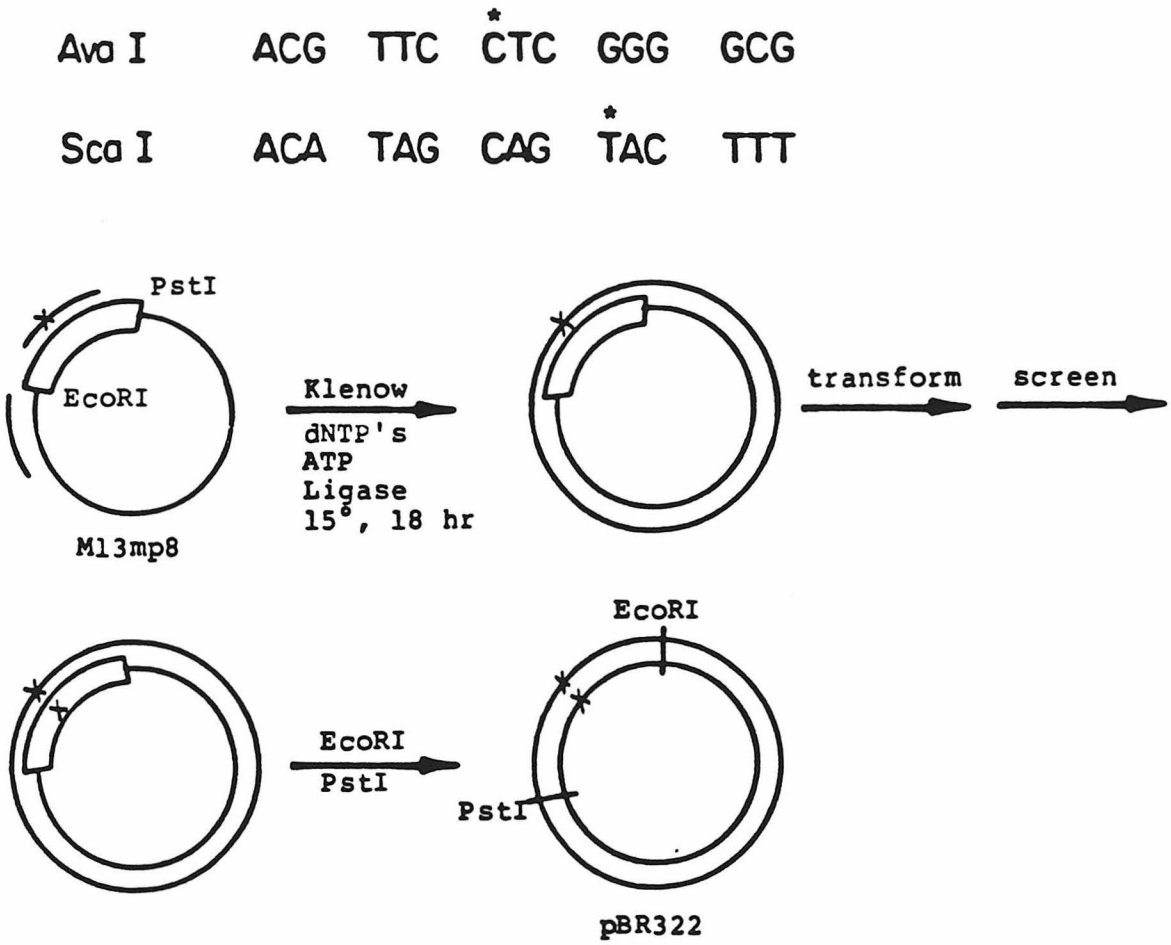


Figure I-1. Two primer method of oligonucleotide directed mutagenesis in M13 mp8 used to introduce the Ava I site at 3972 and the Sca I site at 3937.

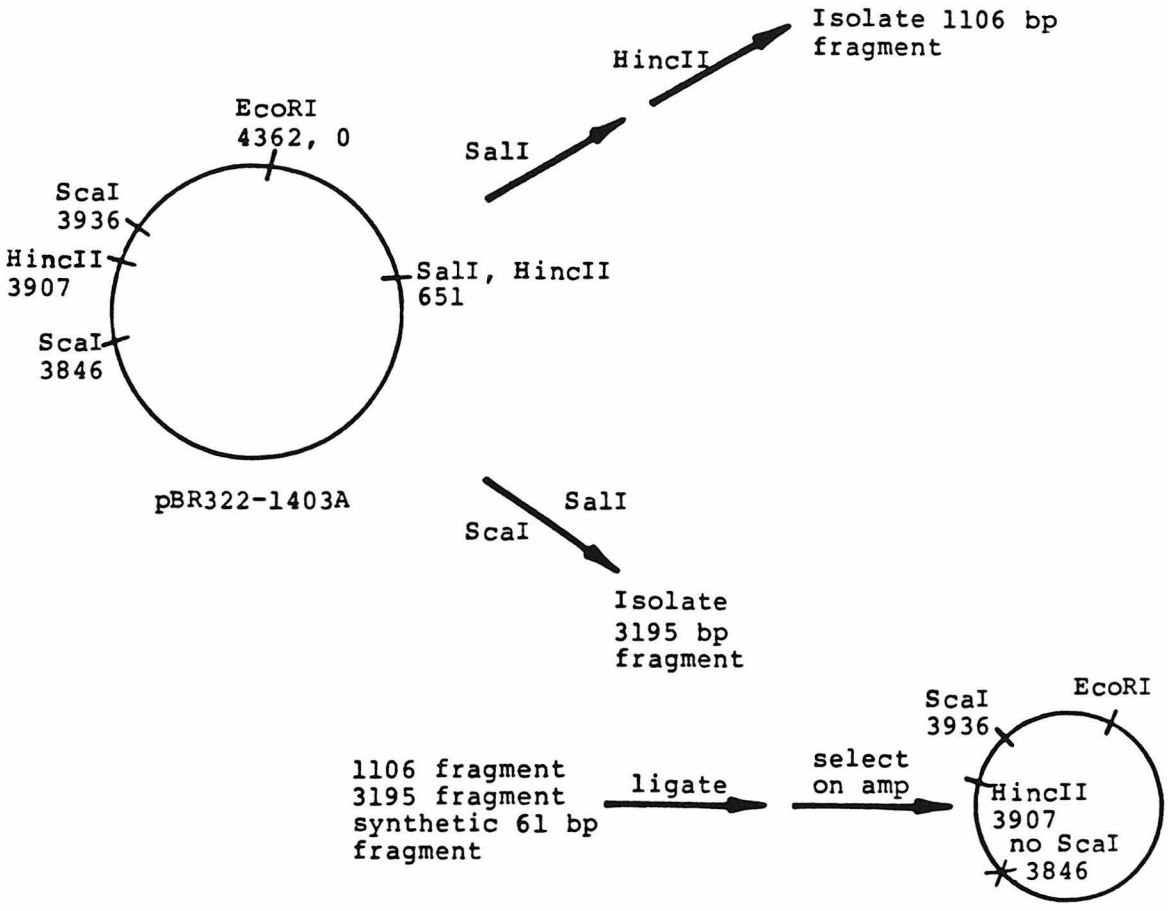


Figure I-2. Three fragment ligation used to remove the Sca I site at 3846 in pBR322.

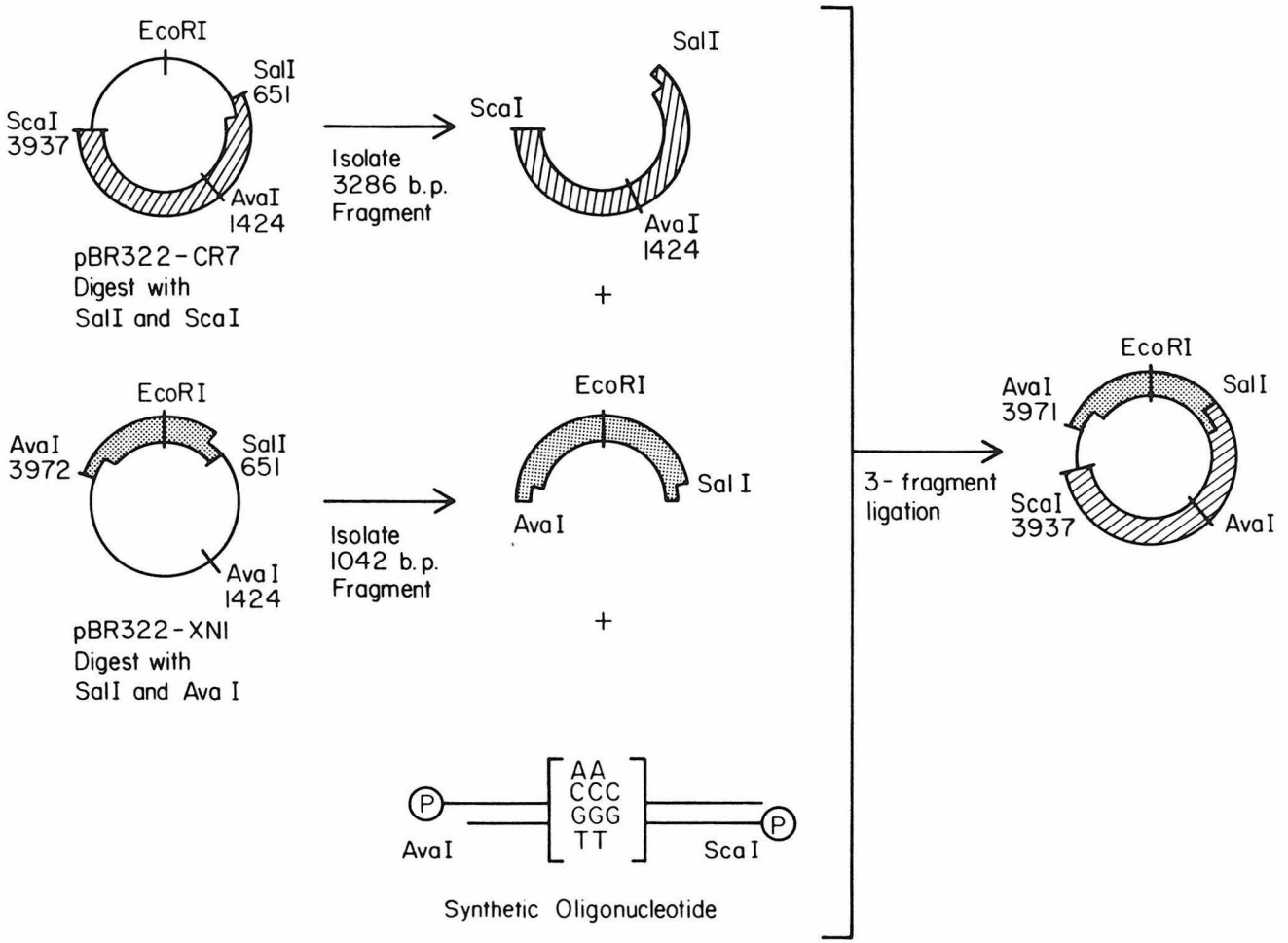


Figure I-3. Three fragment ligation used to insert the mixture of oligonucleotides containing the various codons for residue 71 in ß-lactamase.

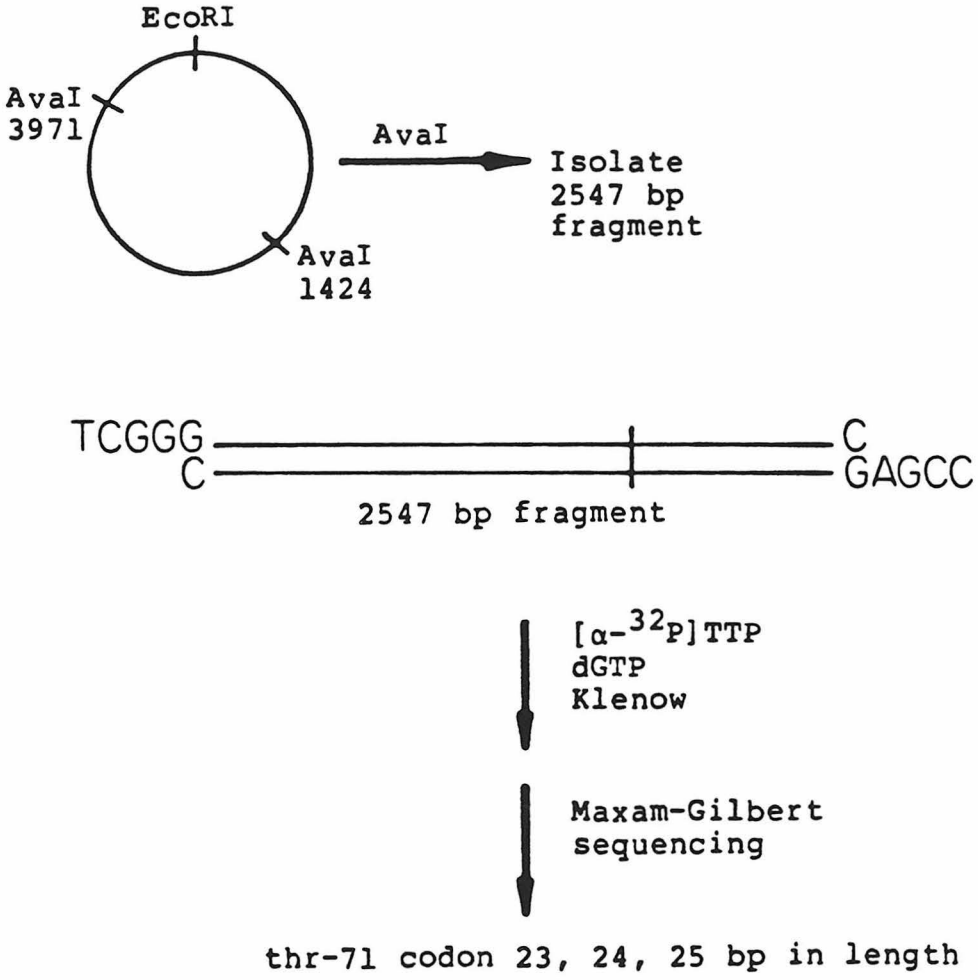
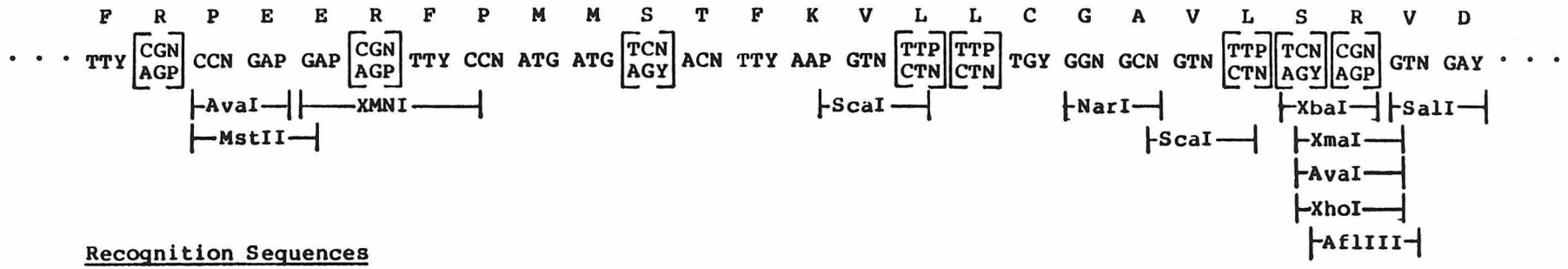


Figure I-4. Method used to incorporate ^{32}P at nucleotide T-3973 in plasmid fragment of pBR322 from site saturation.



Recognition Sequences

AvaI	C/YCGPG
ScaI	AGT/ACT
MstII (SauI)	CC/TNAGG
XMNI	GAANN/NNTTC
NarI	GG/CGCC
SalI	G/TCGAC
XbaI	T/CTAGA
XmaI	C/CCGGG
XhoI	C/TCGAG
AflIII	A/CPYGT

Figure I-5. Reverse translation of the amino acid sequence flanking Thr 71 and the potential restriction sites that exist in the resulting sequence.

Table I - Codons observed for plasmids in 108 colonies for site saturation of residue 71 in ß-lactamase.

Colony Number	Codon	Amino Acid	Colony Number	Codon	Amino Acid	Colony Number	Codon	Amino Acid
1	TTG	Leu	49	ATG	Met	97	ATC	Ile
2	GTG	Val	50	TCC	Ser	98	GCG	Ala
3	AAC i*		51	TTG	Leu	99	TAC	Tyr
4	CAC	His	52	CAG d		100	AGG	Arg
5	TGC d*		53	AGG	Arg	101	GGGC	
6	TCC	Ser	54	GAC	Asp	102	TAC	Tyr
7	GAG	Glu	55	AGG	Arg	103	ACC	Thr
8	TGCC		56	AAG	Lys	104	AAG i	
9	GGG	Gly	57	TGG	Trp	105	TCG	Ser
10	TGC i		58	TAG	stop	106	GGC	Gly
11	ACG	Thr	59	TTG	Leu	107	?	
12	TTG	Leu	60	TAC	Tyr	108	ACC	Thr
13	ATC	Ile	61	CAC	His			
14	AAG	Lys	62	CAC i				
15	CGG	Arg	63	TCC	Ser			
16	TAC	Tyr	64	ACC	Thr			
17	ATG	Met	65	TGC	Cys			
18	CCG	Pro	66	AGC	Ser			
19	GAG	Glu	67	GCC	Ala			
20	ACG	Thr	68	TTG	Leu			
21	CCG	Pro	69	?				
22	GGC	Gly	70	CAC i				
23	AAG	Lys	71	TAC	Tyr			
24	GCG	Ala	72	TAC	Tyr			
25	TGC	Cys	73	TTG	Leu			
26	CGG	Arg	74	ATG	Met			
27	TCC	Ser	75	TGG i				
28	CTC	Leu	76	AAG	Lys			
29	GATC		77	AAC	Asn			
30	CCC	Pro	78	GCC	Ala			
31	GCC	Ala	79	GCG	Ala			
32	TGG	Trp	80	AAG	Lys			
33	TGG	Trp	81	CACC				
34	ATC	Ile	82	TCG	Ser			
35	ATC	Ile	83	ATC	Ile			
36	TTC	Phe	84	TAC	Tyr			
37	TAC	Tyr	85	ATG	Met			
38	TCC	Ser	86	GCC	Ala			
39	ATC	Ile	87	?				
40	TTG	Leu	88	ATG	Met			
41	GCG	Ala	89	CTG i				
42	GAG	Glu	90	AGC	Ser			
43	ACC	Thr	91	CTG d				
44	GAC	Asp	92	TAG	stop			
45	AGC	Ser	93	ATG	Met			
46	GCC	Ala	94	TCG	Ser			
47	ATG d		95	CAG	Gln			
48	TGG	Trp	96	AGG d				

* i or d indicates an insertion or deletion within the synthetic fragment.

Table II - Frequencies of codons observed at residue 71

Codon	Number Observed	Codon	Number Observed
AAC	2	GAC	2
AAG	6	GAG	3
ACC	4	GCC	5
ACG	2	GCG	4
AGC	3	GGC	2
AGG	4	GGG	1
ATC	6	GTC	0
ATG	7	GTG	2
CAC	3	TAC	8
CAG	3	TAG	2
CCC	1	TCC	5
CCG	2	TCG	3
CGC	0	TGC	4
CGG	2	TGG	5
CTC	1	TTC	1
CTG	2	TTG	6

Insertion in degenerate codon	4
Deletion at Sca I site	3
Insertion at Ava I site	6
Deletion at Ava I site	1
Insertion in synthetic segment	1

Table III - Frequencies of bases at each position in the codon for residue 71.

First Position		Second Position		Third Position	
Base	Number Observed	Base	Number Observed	Base	Number Observed
A	34	A	29		
C	14	C	26	C	47
G	19	G	21	G	54
T	34	T	25		

CHAPTER II

CHARACTERIZATION OF MUTANT ß-LACTAMASES WITH ALL POSSIBLE AMINO ACID SUBSTITUTIONS AT RESIDUE 71.

Introduction

ß-Lactamase catalyzes hydrolysis of the ß-lactam ring of penam and cephem antibiotics (1) and thereby confers resistance to these antibiotics on cells producing this enzyme. Several distinct ß-lactamases have been isolated from a variety of bacteria; these enzymes have been classified into three groups: A, B, and C (2). The class A ß-lactamases include the S. aureus PC1, B. licheniformis 749/c, B. cereus 569/H1 as well as RTEM-1 and RTEM-2 enzymes. Many similarities exist between the enzymes composing this class of ß-lactamases, including sequence homologies, catalytic mechanism, and probable three-dimensional structural similarities (2). Class B consists of a B. cereus enzyme that requires zinc for catalysis (3). Class C enzymes are the chromasomal ß-lactamases from E. coli K12 and Ps. aeruginosa. The catalytic mechanism of class C enzymes resembles that of class A enzymes in that a catalytic serine forms an acyl-enzyme intermediate; however, unlike the class A enzymes, these ß-lactamases hydrolyze cephems more rapidly than penams (4,5) and also exhibit transferase activity toward added nucleophiles and will catalyze hydrolysis of the resulting esters (6). We have studied the RTEM-1 enzyme encoded on pBR322, which is a class A ß-lactamase.

The precise mechanism by which ß-lactamases hydrolyze ß-lactam antibiotics remains uncertain. The class A enzymes have been studied more thoroughly than class B or C enzymes; this discussion will address only class A enzymes. The ratio

of k_{cat} to K_m for good substrates such as benzylpenicillin approaches 10^8; the rate of catalysis for these good substrates appears to be diffusion controlled (7). The first step in hydrolysis apparently involves formation of an acyl-enzyme intermediate with an active site Ser (8,9) (Ser 70 in the consensus numbering for class A ß-lactamases (10)) which can be observed directly for very poor substrates such as cefoxitin (11) or trapped at low pH or subzero temperatures for good substrates such as dansyl-penicillin (12). For cephalosporins, reclosure of the ß-lactam ring and subsequent release of intact antibiotic from the enzyme occurs at approximately one third the rate of acylation (13) indicating that the conformation of the cephalosporinoyl group in the acyl-enzyme complex resembles that of free substrate; therefore, strain may exist in the acyl-enzyme bond. The rate of deacylation of the enzyme to produce the corresponding acid of the antibiotic is unaffected by the addition of external nucleophiles (11) suggesting the involvement of an as yet unidentified internal nucleophile. Although results such as these are adding quickly to our understanding of the catalytic mechanisms of ß-lactamase, the precise mechanism, as well as the specific amino acid residues (in addition to Ser 70) involved in catalaysis, remain unclear.

The class A ß-lactamases contain a conserved triad (2) (Ser-Thr-x-Lys, residues 70 to 73 in the consensus numbering of these enzymes). The hydroxyl group of Ser 70 opens the ß-lactam ring to form the acyl enzyme intermediate

(8,11). Replacement of this serine by Cys results in a mutant enzyme with reduced activity (14) whereas replacement by Thr results in an inactive enzyme (15); apparently, this position requires a primary nucleophile. The Lys at residue 73 is conserved in all class A (2) and class C ß-lactamases (16) as well as D-Ala-D-Ala carboxypeptidases (17), which are the target enzymes for penam and cephem antibiotics and are likely to be related to the ß-lactamases structurally (18), mechanistically, and possibly evolutionarily (19). My studies have focused on the conserved Thr 71, whose role in the acivity of ß-lactamases was unknown.

Using site saturation, I have generated all of the possible 19 amino acid substitutions at residue 71 (20) as described in Chapter I. Analysis of these mutant enzymes included determination of their ability to confer <u>in vivo</u> resistance to penam and cephem antibiotics and determination of the relative quantities of enzyme in whole cell extracts of <u>E. coli</u> by antibody stains of electrophoretic blots. Surprisingly, cells containing any of 14 of the mutant ß-lactamases displayed appreciable resistance to ampicillin, only cells containing mutants with Tyr, Trp, Asp, Lys, or Arg at residue 71 had no observable resistance to ampicillin. However, all of the 19 mutant proteins show significantly increased sensitivity to proteolysis, especially at 37°C. These results suggest that this residue is not directly involved in binding or catalysis but is important for stability of the protein.

Materials & Methods

Cells

E. coli LS-1 cells (21) containing wild type or mutant ß-lactamase genes on pBR322 as described in Chapter I were grown in LB media (22) (L-broth and L-agar) containg 15 mg/l of tetracycline or the stated concentration of penam or cephem antibiotics.

Phenotype screens

Initially, resistance to penam or cephem antibiotics was determined by using sterile toothpicks to pick individual colonies and then to innoculate specific positions on an agar plate containing an appropriate concentration of antibiotic.

More accurate values for the maximal level of antibiotic resistance provided to cells for the various mutants were determined by spreading cells onto L-agar plates containing a continuous concentration gradient of antibiotic. Antibiotic concentration gradients were generated by elevating one end of a petri dish 5 mm and pouring 15 ml of L-agar containing an appropriate concentration of antibiotic into the tilted dish. These were allowed to dry overnight. The plates were then placed on a flat, horizontal surface and 15 ml of L-agar was poured on top; the plates were used immediately after they hardened. A 50 μl aliquot of a $1:10^5$ dilution of a saturated culture (a saturated culture of *E. coli* LS1 contains approximately 2×10^9 cells per ml) was spread over half of the plate. A standard mutant was plated on the other half; standards used were as follows: for ampicillin and

benzylpenicillin at 37°C and for 6-aminopenicillanic acid at 30°C, Thr 71 --> Leu; for 6-aminopenicllanic acid at 37°C, Thr 71 --> Cys; for ampicillin and benzylpenicillin at 30°C, Thr 71 --> Glu. These standards were used because colonies containing these mutants grew up to approximately half of the maximal concentration of antibiotic in the plates under the conditions described above.

Protein gels and electrophoretic blots

Whole cell extracts were prepared for E. coli harboring plasmids encoding either mutant or wild type ß-lactamases. The cells were grown to late log-phase (O.D. at 600 nm ~ 1) and the absorbance at 600 nm was measured. The cells were pelleted by centrifugation and resuspended in loading buffer [10% (vol/vol) glycerol / 5% (vol/vol) 2-mercaptoethanol / 3% (wt/vol) sodium dodecyl sulfate/ 62.5 mM Tris-HCl, pH 6.8 / 1.0 mM EDTA] and incubated 10 minutes at 95°C. Cells from cultures with densities of 1 O.D. were resuspended in 100 µl of loading buffer; cells from cultures with slightly higher or lower densities were resuspended in proportionally greater or lesser volumes of loading buffer. 20 µl aliquots of these whole cell extracts were loaded onto a 15 cm 12% polyacrylamide gel with a 2 cm 4% stacking gel and electrophoresed at 5 mA for approximately 12 hours.

The protein was transferred from the polyacrylamide gel to nitrocellulose by electrophoresis (electrophoretic blot or western blot) according to the Bio-Rad procedure (23): the polyacrylamide gel was washed for 30 minutes in 25

mM Tris base / 192 mM glycine (pH ~ 8.3) / 20% (vol/vol) methanol; the gel was then sandwiched between blotting paper and pure nitrocellulose (Schleicher & Scheull) and electrophoresed in a Bio-Rad Trans-Blot cell at 50 V for 5 hours.

The nitrocellulose was stained according to the procedure from the Vectastain ABC kit (24) using antibody raised against ß-lactamase in rabbits. The initial injection used to raise antibody contained denatured ß-lactamase purified by SDS/polyacrylamide gel electrophoresis (25); booster shots contained protein purified as described elsewhere (25). The rabbit was bled after the booster shot and serum was collected and used without furthur purification. However, the membrane was first treated with a solution of 1% (wt/vol) BSA and 0.1% (vol/vol) normal goat serum for 1.5 hours at room temperature. For binding of the primary antibody, a 1:1000 dilution of rabbit antiserum in TPBS was incubated with the nitrocellulose membrane for 1 hour at room temperature.

Results

From the 423 colonies obtained after tranformation of cells with DNA from ligation of the synthetic insert into the plasmid as described in Chapter I, 108 colonies were picked onto L-agar plates containing tetracycline and were also tested for resistance to ampicillin (10, 50, 100, and 500 mg/l), benzylpenicillin (100 mg/l), 6-aminopenicillanic acid (10, 25, 50, and 100 mg/l), cephalothin (25, 50, and 100 mg/l), and cephalexin (25 and 50 mg/l). Ampicillin and cephalothin plates were incubated at 30°C and 37°C; all others were incubated only at 37°C. The results are presented in Table I. This phenotype screening indicated that a change in antibiotic specificity and reduced stability particularly at 37°C were characteristics of some of the mutants.

After sequencing plasmids derived from 105 of these colonies as discussed in Chapter I, the sequences were matched to the phenotypes determined above (see Table I). Mutants with Tyr, Trp, Asp, Lys, and Arg at position 71 gave no resistance to any of the five antibiotics. A mutant with Phe was resistant to 100 mg/l of ampicillin at 30°C, but sensitive even to low levels of ampicillin at 37°C. Mutants with Gly, Gln, and Glu produced resistance up to 100 mg/l of ampicillin and benzylpenicillin at 37°C. Mutants with Ala, Val, Leu, Ile, Pro, His, Cys, Ser, Thr, and Asn gave resistance to >500 mg/l of ampicillin and >100 mg/l of benzylpenicillin. Only mutants with His, Cys, Ser, and Thr were resistant to >100 mg/l of 6-aminopenicillanic acid, those with

Gly, Ala, Val, Leu, Ile, Pro, Glu, Gln, and Asn were resistant to 25 mg/l, and those with Met (although resistant to ampicillin) showed no resistance to 6-aminopenicillanic acid. Only the wild type enzyme (Thr 71) was resistant to cephalothin and cephalexin at 37°C. Mutants with Ser at residue 71 were resistant to 50 mg/l of cephalothin at 30°C, but none of the other mutants grew in the presence of this antibiotic.

More accurate values for the level of antibiotic resistance provided to cells by the mutant proteins were determined by spreading cells onto an agar plate with a continuous concentration gradient of antibiotic. The maximal concentration at which colonies were established could then be observed; these values are presented in Table II. These values were independent of dilutions of cells higher by a factor of 2 or lower by a factor of 5 than the dilutions used in determining the phenotypic characteristics.

Whole cell extracts from approximately 1 X 10^8 *E. coli* cells harboring the various wild type and mutant plasmids incubated either at 30°C or 37°C were electrophoresed in denaturing sodium dodecyl sulfate / polyacrylamide gels and then transferred to nitrocellulose. Figure 1 shows antibody stains of these blots. ß-Lactamase or its mutants are present in all lanes, but the quantity varies substantially for different mutants; this variation is greatly reduced at 30°C. We believe that the protein responsible for the band slightly above ß-lactamase is pre-ß-lactamase because it

represents a protein of the appropriate size and because this band is absent for the Thr 71 --> amber mutant. The blots suggest that the mutant proteins are processed normally, but that mature mutant proteins are more susceptible than wild-type ß-lactamase to degradation, probably by periplasmic proteases. Some of the bands below those for intact mutants of ß-lactamase likely represent proteolytic fragments and appear to differ slightly between the various mutants.

Discussion

The antibiotic resistances of colonies containing the 19 mutant ß-lactamases were initially determined by picking colonies onto agar plates with various levels of antibiotics (see Table I). More accurate values were measured by plating cells containing the mutants onto agar plates having a linear concentration gradient of antibiotics (see Table II). This latter method gives consistently lower values for antibiotic resistance as in this case individual cells must establish colonies whereas when picking colonies many cells are initially present at a single site, allowing them to cooperate in inactivating sufficient antibiotic to allow growth (this is commonly called the innoculum effect). However, both sets of data are useful as colonies with mutants having low lactamase activity would be classified as inactive if only plating were used. In Table II, colonies that grew when picked, but not when plated, are designated as having trace activity.

The antibiotic resistance of colonies containing mutant plasmids reflects many factors: plasmid copy number, rate of transcription, stability and rate of translation of the mutant mRNAs, stability of pre-ß-lactamase in the cytoplasm, rate of processing and secretion, stability of the protein in the periplasm, and, finally, the intrinsic enzymatic activity.

With regard to the steps from transcription through translation, the codon used for a given amino acid at residue

71 did not alter observed antibiotic resistance. For example, wild type enzyme with ACT (pBR322), ACC or ACG for Thr 71 all confer the same level of antibiotic resistance as do Thr 71 --> Ser mutants with TCC, AGC, TCG, or TCT for serine and Thr 71 --> Leu mutants with TTG or TTC for leucine. Likewise, for all of the active mutants listed in Table II, different codons for the same amino acid always produced identical values for the maximal level of antibiotic resistance. Though the presence of infrequently used codons in E. coli may affect translation (26), our results do not implicate this as the dominant factor in the observed phenotypic resistances for the conditions used in these experiments.

Mutations may affect processing and secretion of proteins (27) and thereby alter the relative concentrations of mutant ß-lactamases in the periplasm. Although mutants such as Thr 71 --> Ser and Ser 70 --> Thr and the double mutant Ser 70 --> Thr, Thr 71 --> Ser are processed and secreted normally in Salmonella typhimurium (25), these mutations are unlikely to affect these processes as drastically as mutations such as Thr 71 --> Trp, Tyr, Lys; even in these cases, however, abundant processed protein is apparent in cells growing at 30°C (see Figure 1).

Resistance to thermal denaturation and proteolysis can greatly influence the in vivo concentration of mutant proteins (25,28). Antibody stains of electrophoretic blots for whole cell extracts of E. coli harboring the wild type or

mutant plasmids grown at 30°C and 37°C show significant differences in the *in vivo* concentrations of the various mutants. A band representing the precursor protein shows similar intensities for the mutant and wild type enzymes; the presence of normal amounts of precursor proteins suggests that processing and secretion have not been drastically altered for the mutants. However, the amounts of protein in the periplasm vary greatly for the mutants as contrasted to wild type, particularly for cells growing at 37°C. The electrophoretic blots were carried out on samples after denaturation, boiling with 2-mercaptoethanol and elecrophoresis with SDS; consequently, the reduced amounts of mutant proteins relative to wild type ß-lactamase result from more extensive proteolysis rather than simple thermal denaturation of the mutants.

In these blots, the relative antigenicities of the various denatured ß-lactamases could significantly affect the amounts of mutant proteins apparently present. The proteins in question have been denatured, the antibody is polyclonal rabbit antiserum raised against denatured wild type ß-lactamase and the mutants in question have only single site substitutions; for these reasons, the relative antigenicities of the mutants should not differ appreciably from each other or from that of the wild type protein. The relatively constant amount of an unidentified *E. coli* protein in all lanes serves as an internal standard to ensure that nearly similar amounts of total protein were loaded onto each of the lanes.

Accordingly, the relative intensities of the bands for wild type and mutant ß-lactamases should reflect the relative amounts of these mutants actually present under the conditions specified.

A surprising number of mutants with amino acid substitutions at residue 71 display lactamase activity; indeed, of the 19 substitutions at residue 71 only 5 (Trp, Tyr, Lys, Arg, and Asp) give proteins that confer no observable resistance to ampicillin or benzylpenicillin. Enzymes with Gln, Glu, Asn, and Phe at residue 71 provide low level resistance, but enzymes with all other amino acids at this site have high activity toward benzylpenicillin and ampicillin.

Some conclusions about structure-function relationships can be drawn from these data. Mutants having amino acids with charged or very large side cains at residue 71 have at best very low activity, otherwise both polar and non-polar side chains seem compatible with appreciable activity. When 6-aminopenicillanic acid is used to select for activity, mutants with amino acids having non-polar side chains at residue 71 show decreased acivity whereas mutants with amino acids having polar side chains at residue 71 retain high levels of activity toward this antibiotic. All mutants show a dramatically decreased ability to confer resistance toward the cephem antibiotics cephalothin and cephalexin. At 30°C, wild type ß-lactamase (Thr 71) confers resistance to at least 100 mg/l of cephalothin while the mutant Thr 71 --> Ser confers resistance only to low levels of the cephem antibi-

otics; at 37°C only wild type enzyme provides observable resistance.

Because many mutants at residue 71, including those having amino acids with polar and non-polar side chains at this residue, show appreciable activity toward penams, we conclude that Thr at residue 71 is not essential for binding or catalysis, though it may perform some contributing role in these functions. However, residue 71 does play a very important role in the structural stability of the protein as observed in the increased sensitivity to proteolysis of all 19 mutants. The importance of branching at the ß-carbon of threonine with a polar hydroxyl group and non-polar methyl group is demonstrated, for example, by the increasing antibiotic resistance of cells growing at 37°C provided by the series of enzymes with Gly < Ala < Leu < Val at residue 71. The increased thermal and proteolytic sensitivity (relative to wild type) of the mutant Thr 71 --> Ser (25), lacking only the methyl group of Thr, further supports this view. Interestingly, in a recent study of proteins from mesophilic and thermophilic bacteria, Ser --> Thr substitutions were frequently observed (29) suggesting that Thr can provide greater thermal stability to proteins than Ser. The high levels of catalytic activity observed for the Thr 71 --> Cys and Thr 71 --> Ser mutants demonstrate the importance of the polar group. The reduced activity toward 6-aminopenicillanic acid of enzymes having amino acids with non-polar side chains at residue 71 suggests that the side chain of residue 71

influences the hydrophobic or hydrophilic characteristics (possibly hydration) of the pocket that accommodates the substituent attached to the ß-lactam ring of the substrate. This possibility is further emphasized by the lower activity of enzymes with Glu and Gln at residue 71 toward benzylpenicillin (neutral, non-polar substituent) than toward ampicillin (positively charged substituent).

The high level of activity observed for the Thr 71 --> Ile and Thr 71 --> His mutants was surprising; one would not consider these amino acids as suitable replacements for Thr. The Thr 71 --> Ile mutant appears to be slightly more stable than mutants with other nonpolar amino acid side chains at residue 71, perhaps emphasizing the importance of a nonpolar interaction in this region of the protein as with the methyl group of threonine. The high level of activity observed for the Thr 71 --> His mutant is difficult to interpret; possibly the hydrogen bonding abilities of the side chain of this residue resemble those of threonine. The stability of the Thr 71 --> Pro mutant as seen in Figure 2 is also noteworthy. The proline side chain may reduce the conformational flexibilty at residue 71; however, the stability of this mutant protein remains well below that of the wild type enzyme.

All mutants confer on cells at best very low resistance toward cephalothin and cephalexin; this likely results from the inherently lower activity of ß-lactamase against cephem antibiotics (30) and accords with results observed for muta-

tions in other regions of the enzyme (31). This behavior also demonstrates the importance of the conserved residues at the catalytic site in providing an enzyme with high levels of catalytic activity and stability.

Conclusion

I have used the site-saturation approach to structure-function studies of proteins to determine the role of Thr 71 in the activity of the pBR322 encoded RTEM-1 ß-lactamase. The results demonstrate that residue 71 has little or no affect on binding or catalysis but is essential for stability of the protein. One would not have predicted that Thr could perform such a unique structural role; even Ser does not adequately replace Thr at this site in the protein. The results also indicate that residue 71 may influence the region of the protein that accommodates the side chain attached to the ß-lactam ring of the substrate. Particularly surprising results were obtained for the Thr 71 --> Ile and Thr 71 --> His mutants; one would not have predicted that these residues could adequately replace Thr.

The results presented here demonstrate the potential of site saturation for determining the role of specific amino acid residues in the function of a protein. Future studies using site saturation should produce many new insights into the relationships between structure and function in proteins.

References

1. Datta, N. & Kontomichalou, P. (1965) Nature (London) 208, 239-241.
2. Ambler, R.P. (1980) Phil. Trans. R. Soc. Lond. B 289, 321-331.
3. Kuwabara, W. & Abraham, E.P., (1967) Biochem. J. 103, 27c-30c.
4. Sykes, R.B. & Matthew, M. (1976) J. Antimicrob. Chemother. 2, 115-157.
5. Jaurin, B. & Grundstrom, T. (1981) Proc. Natl. Acad. Sci. USA 78, 4897-4901.
6. Knott-Hunziker, V., Petursson, S., Waley, S.G., Jaurin, B. & Grundstrom, T. (1982) Biochem. J. 207, 315-322.
7. Hardy, L.W. & Kirsh, J.F. (1984) Biochemistry 23, 1275-1282.
8. Knott-Hunziker, V., Waley, S.G., Orlek, B.S., & Sammer, P.G. (1979) FEBS Lett. 99, 56-61.
9. Cohen, S.A. & Pratt, R.F. (1980) Biochemistry 19, 3996-4003.
10. Ambler, R.P. (1979) in Beta-Lactamases (eds. Hamilton-Miller, J.M.T. & Smith, J.H., Academic Press, London) 99-125.
11. Fisher, J., Belasco, J.G., Khosla, S. & Knowles, J.R. (1980) Biochemistry 19, 2895-2901.
12. Cartwright, S.L. & Fink, A.L. (1982) FEBS Lett. 137, 186-188.

13. Anderson, E.G. & Pratt, R.F. (1981) J. Biol. Chem. 256, 11401-11404.

14. Sigal, I.S., DeGrado, W.F., Thomas, B.J. & Petteway, S.R., Jr. (1984) J. Biol. Chem. 259, 5327-5332.

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

16. Knott-Hunziker, V., Petursson, S., Jayatilake, G.S., Waley, S.G., Jaurin, B. & Grundstrom, T. (1982) Biochem. J. 201, 621-627.

17. Nicholas, R.A., Ishino, F., Park, W., Matsuhashi, M. & Stromiger, J.L. (1985) J. Biol. Chem. 260, 6394-6397.

18. Kelly, J.A., Dideberg, O., Charlier, P., Wery, J.P., Libert, M., Moews, P.C., Knox, J.R., Duez, C., Fraipont, Cl., Joris, B., Dusart, J., Frére, J.M. & Ghuysen, J.M. (1986) Science 231, 1429-1431.

19. Waxman, D.J. & Strominger, J.L. (1983) Ann. Rev. Biochem. 52, 825-869.

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

21. Boyer, H.W. & Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.

22. Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

23. Anonymous (1984) Trans-Blot Cell Operating Instructions (Bio-Rad Laboratories, Richmond, CA).

24. Anonymous (1984) Immunodetection of Antigens on Nitrocellulose using Vectastain ABC Kit (Vector Laboratories, Burlingame, CA).

25. Dalbadie-McFarland, G., Neitzel, J. & Richards, J.H. (1985) Biochemistry 25, 332-338.

26. Ikemura, T. (1981) J. Mol. Biol. 151, 389-409.

27. Kadonaga, J.T., Gautier, A.E., Straus, D.R., Charles, A.D., Edge, M.D. & Knowles, J.R. (1984) J. Biol. Chem. 259, 2149-2154.

28. Hecht, M.H., Sturtevant, J.M. & Sauer, R.T. (1984) Proc. Natl. Acad. Sci. USA 81, 5685-5689.

29. Argos, P., Rossmann, M.G., Grau, U.M., Zuber, H., Frank, G. & Tratschin, J.D. (1979) Biochemistry 18, 5698-5703.

30. Sykes, R.B. & Smith, J.T. (1979) in Beta-Lactamases (eds. Hamilton-Miller, J.M.T. & Smith, J.T., Academic Press, London) 360-401.

31. Barany, F. (1985) Proc. Natl. Acad. Sci. USA 82, 4202-4206.

Figure II-1. Electrophoretic blots of ß-lactamases with the 19 amino acid substitutions at residue 71. The blots were treated with anti-ß-lactamase antibody and visualized using the horseradish peroxidase assay (see reference 24).

37 °C
E. coli protein
precursor β-lactamase
β-lactamase
E. coli protein
β-lactamase
E. coli LS1
amber 71
pBR 322
gly
ala
val
leu
ile
met
pro
phe
tyr
trp
β-lactamase
E. coli LS1
amber 71
pBR322
his
cys
ser
thr
asn
gln
asp
glu
lys
arg
30 °C
E. coli protein
precursor β-lactamase
β-lactamase
E. coli protein
β-lactamase
E. coli LS1
amber 71
pBR322
gly
ala
val
leu
ile
met
pro
phe
tyr
trp
β-lactamase
E. coli LS1
amber 71
pBR322
his
cis
ser
thr
asn
gln
asp
glu
lys
arg

Table I. - Maximal level of resistance determined by picking.

			Benzylpen.	Ampicillin 30°C	Ampicillin 37°C	Aminopen.	Cephalothin 30°C	Cephalothin 37°C
1.	TTG	Leu	>100	>500	100	25	0	0
2.	GTG	Val	>100	>500	>500	50	0	0
3.	AAC	i*	0	0	0	0	0	0
4.	CAC	His	>100	>500	>500	>100	0	0
5.	TGC	d*	0	0	0	0	0	0
6.	TCC	Ser	>100	>500	>500	>100	0	0
7.	GAG	Glu	0	100	50	0	0	0
8.	TGCC		0	0	0	0	0	0
9.	GGG	Gly	>100	>500	100	25	0	0
10.	TGC	i	0	0	0	0	0	0
11.	ACG	Thr	>100	>500	>500	>100	>100	>100
12.	TTG	Leu	>100	>500	>500	25	0	0
13.	ATC	Ile	>100	>500	>500	25	0	0
14.	AAG	Lys	0	0	0	0	0	0
15.	CGG	Arg	0	0	0	0	0	0
16.	TAC	Tyr	0	0	0	0	0	0
17.	ATG	Met	>100	>500	>500	50	0	0
18.	CCG	Pro	>100	>500	>500	25	0	0
19.	GAG	Glu	0	>500	100	25	0	0
20.	ACG	Thr	>100	>500	>500	>100	>100	>100
21.	CCG	Pro	>100	>500	>500	25	0	0
22.	GGC	Gly	>100	>500	100	25	0	0
23.	AAG	Lys	0	0	0	0	0	0
24.	GCG	Ala	>100	>500	>500	25	0	0
25.	TGC	Cys	>100	>500	>500	>100	0	0
26.	CGG	Arg	0	0	0	0	0	0
27.	TCC	Ser	>100	>500	>500	>100	50	0
28.	CTC	Leu	>100	>500	>500	25	0	0
29.	GATC		0	0	0	0	0	0
30.	CCC	Pro	>100	>500	>500	>100	0	0
31.	GCC	Ala	>100	>500	>500	25	0	0
32.	TGG	Trp	0	0	0	0	0	0
33.	TGG	Trp	0	0	0	0	0	0
34.	ATC	Ile	>100	>500	>500	25	0	0
35.	ATC	Ile	>100	>500	>500	25	0	0
36.	TTC	Phe	0	100	10	0	0	0
37.	TAC	Tyr	0	0	0	0	0	0
38.	TCC	Ser	>100	>500	>500	>100	25	0
39.	ATC	Ile	>100	>500	>500	>100	0	0
40.	TTG	Leu	>100	>500	>500	25	0	0
41.	GCG	Ala	>100	>500	>500	25	0	0
42.	GAG	Glu	0	>500	>500	0	0	0
43.	ACC	Thr	>100	>500	>500	>100	>100	>100
44.	GAC	Asp	0	0	0	0	0	0
45.	AGC	Ser	>100	>500	>500	>100	50	0
46.	GCC	Ala	>100	>500	>500	25	0	0
47.	ATG	d	0	0	0	0	0	0
48.	TGG	Trp	0	0	0	0	0	0
49.	ATG	Met	>100	>500	>500	0	0	0

Table I. (continued)

			Benzylpen.	Ampicillin 30°C	Ampicillin 37°C	Aminopen.	Cephalothin 30°C	Cephalothin 37°C
50.	TCC	Ser	>100	>500	>500	>100	50	0
51.	TTG	Leu	>100	>500	>500	>100	0	0
52.	CAG d		0	0	0	0	0	0
53.	AGG	Arg	0	0	0	0	0	0
54.	GAC	Asp	0	50	0	0	0	0
55.	AGG	Arg	0	0	0	0	0	0
56.	AAG	Lys	0	0	0	0	0	0
57.	TGG	Trp	0	0	0	0	0	0
58.	TAG	stop	0	0	0	0	0	0
59.	TTG	Leu	>100	>500	>500	25	0	0
60.	TAC	Tyr	0	0	0	0	0	0
61.	CAC	His	>100	>500	>500	>100	0	0
62.	CAC i		0	0	0	0	0	0
63.	TCC	Ser	>100	>500	>500	>100	50	25
64.	ACC	Thr	>100	>500	>500	>100	>100	>100
65.	TGC	Cys	>100	>500	>500	>100	0	0
66.	AGC	Ser	>100	>500	>500	>100	0	0
67.	GCC	Ala	>100	>500	>500	10	0	0
68.	TTG	Leu	>100	>500	>500	>100	0	0
69.	?	Thr (?)	>100	>500	>500	>100	>100	>100
70.	CAC i		0	50	0	0	0	0
71.	TAC	Tyr	0	0	0	0	0	0
72.	TAC	Tyr	0	0	0	0	0	0
73.	TTG	Leu	>100	>500	>500	25	0	0
74.	ATG	Met	>100	>500	>500	25	0	0
75.	TGG i		0	0	0	0	0	0
76.	AAG	Lys	0	0	0	0	0	0
77.	AAC	Asn	>100	>500	>500	25	0	0
78.	GCC	Ala	>100	>500	>500	25	0	0
79.	GCG	Ala	>100	>500	>500	25	0	0
80.	AAG	Lys	0	0	0	0	0	0
81.	CACC		0	0	0	0	0	0
82.	TCG	Ser	>100	>500	>500	>100	50	0
83.	ATC	Ile	>100	>500	>500	25	0	0
84.	TAC	Tyr	0	0	0	0	0	0
85.	ATG	Met	>100	>500	>500	25	0	0
86.	GCC	Ala	>100	>500	>500	25	0	0
87.	?	Thr (?)	>100	>500	>500	>100	>100	>100
88.	ATG	Met	>100	>500	>500	25	0	0
89.	CTG i		0	0	0	0	0	0
90.	AGC	Ser	>100	>500	>500	>100	50	25
91.	CTG d		0	0	0	0	0	0
92.	TAG	stop	0	0	0	0	0	0
93.	ATG	Met	>100	>500	>500	25	0	0
94.	TCG	Ser	>100	>500	>500	>100	50	0
95.	CAG	Gln	>100	>500	50	0	0	0
96.	AGG d		0	0	0	0	0	0
97.	ATC	Ile	>100	>500	>500	>100	0	0
98.	GCG	Ala	>100	>500	>500	25	0	0

Table I - (continued)

			Benzylpen.	Ampicillin 30°C	Ampicillin 37°C	Aminopen.	Cephalothin 30°C	Cephalothin 37°C
99.	TAC	Tyr	0	0	0	0	0	0
100.	AGG	Arg	0	0	0	0	0	0
101.	GGGC		0	0	0	0	0	0
102.	TAC	Tyr	0	50	0	0	0	0
103.	ACC	Thr	>100	>500	>500	>100	>100	>100
104.	AAG i		0	0	0	0	0	0
105.	TCG	Ser	>100	>500	>500	>100	50	0
106.	GGC	Gly	>100	>500	100	25	0	0
107.	?		>100	>500	>500	>100	0	0
108.	ACC	Thr	>100	>500	>500	>100	>100	>100

*i=insertion, d=deletion in the synthetic fragment.

Table II - Maximal level of resistance of cells containing residue 71 mutant ß-lactamases as determined by gradient plates. Units are in mg/l

Amino Acid	Codons	Ampicillin		Benzylpenicillin		6-Amino-penicillanic Acid	
		30°C	37°C	30°C	37°C	30°C	37°C
Gly	GGC, GGG	>500	trace	>500	50	35	20
Ala	GCG, GCC	>500	75	>500	100	trace	trace
Val	GTG	>500	150	>500	175	75	trace
Leu	TTG, CTC, CTG	>500	100	>500	125	80	trace
Ile	ATC	>500	>500	>500	>500	90	40
Met	ATG	>500	125	>500	150	30	0
Pro	CCG, CCC	>500	200	>500	350	50	40
Phe	TTC	60	0	0	0	0	0
Trp	TGG	0	0	0	0	0	0
Tyr	TAC	0	0	0	0	0	0
His	CAC	>500	>500	>500	>500	>250	>250
Cys	TGC	>500	>500	>500	>500	>250	100
Ser	TCC, AGC, TCG	>500	>500	>500	>500	>250	>250
Thr	ACG, ACC	>500	>500	>500	>500	>250	>250
Asn	AAC	270	trace	210	trace	55	trace
Gln	CAG	130	20	70	trace	trace	0
Asp	GAC	0	0	0	0	0	0
Glu	GAG	200	trace	70	trace	trace	trace
Lys	AAG	0	0	0	0	0	0
Arg	CGC, AGG	0	0	0	0	0	0
Stop	TAG	0	0	0	0	0	0

>A indicates that this value was the highest level of antibiotic tested

CHAPTER III

ROLE OF A DISULFIDE BOND IN STABILIZING RESIDUE 71 MUTANTS OF ß-LACTAMASE

Introduction

Disulfide bonds in proteins can perform a variety of functions such as regulation of the activity of eukaryotic proteins (1), maintainance of the structural integrity of processed proteins (such as insulin) (2), and stabilization of the native conformation of globular proteins (3,4). Whereas stabilization of the native conformation of globular proteins has been the subject of numerous studies, the precise mechanisms by which disulfide bonds can participate in the dynamics of folding and unfolding and, therefore, protein stability remain uncertain. Several studies indicate that proteins with naturally occuring disulfide bonds can assume their native conformation without formation of disulfide bonds (5,6) and can retain their normal activity after reduction of the disulfide bonds (7). These results suggest that disulfide bonds that provide conformational stability generally do not dictate folding and may not affect catalytic activity.

Of the known class A ß-lactamases (8) (EC 3.5.2.6), only the RTEM-1 and RTEM-2 ß-lactamases encoded by plasmids pBR322 and R6K respectively, contain a disulfide bond (9,10); this bond is between residues 77 and 123 according to the numbering of the consensus sequence (8). Complete reduction of this disulfide bond requires harsh conditions that irreversibly denature the protein; consequently, studies comparing activities of the protein with and without the disulfide bond have not been feasible.

Modern techniques for site-specific mutagenesis (11,12) permit one to remove disulfide bonds without chemically reducing the protein. The naturally occurring disulfide bonds were removed from interleukin-2 by replacing cysteine residues with either serine (13) or alanine (14); although the mutant proteins were active, these activities were substantially lower than that of the wild type protein. Site-specific mutagenesis also allows the introduction of disulfide bonds into proteins; disulfide bonds have been successfully introduced into dihydrofolate reductase (15) and T4 lysozyme (16,17). Such experiments require high resolution three-dimensional structures to identify potential sites that will allow formation of a disulfide bond while minimizing undesirable effects on the conformation or activity of the protein.

I have introduced a mutation into the ß-lactamase gene on pBR322 that changes Cys 77 to Ser. The wild type and mutant enzymes confer the same high levels of <u>in vivo</u> resistance to ampicillin. The activities of purified mutant and wild type enzymes are indistiguishable at 30°C using benzylpenicillin as the substrate at pH values between 5 and 8 (Neitzel, J.J. & Richards, J.H., unpublished results). However, the thermal stability of the purified mutant enzyme is reduced at elevated temperatures (above 40°C). The mutant enzyme is also less active than the wild type enzyme in alkaline solutions (above pH 8) (Neitzel, J.J. & Richards, J.H., unpublished results). These results indicate that the

Cys 77 --> Ser mutation affects the activity and conformational stability of the ß-lactamase protein only when exposed to harsh environments.

In the previous chapters, I discussed the production and characterization of mutant RTEM-1 (pBR322) ß-lactamases with all possible amino acid substitutions at residue 71 (18). All of these mutations resulted in proteins with dramatically decreased stability to E. coli proteases at 37°C; however, at 30°C we observed stabilities similar to the wild type enzyme. Fourteen of the mutations exhibited high levels of activity on ampicillin and benzylpenicillin, including the Thr 71 --> Ile mutant (18). However, the mutation corresponding to Thr 71 --> Ile in the S. aureus PC1 ß-lactamase (also a class A ß-lactamase) results in an inactive enzyme (8), probably due to incomplete folding (19). The disulfide bond, which is present in the RTEM-1 enzyme but not in the S. aureus PC1 enzyme may contribute to this difference between these homologous enzymes, perhaps by stabilizing the catalytically active conformation.

To study the role of the disulfide bond in stabilizing ß-lactamases with mutations at residue 71, I introduced the Cys 77 --> Ser mutation into all 19 of these mutants. I analyzed these 19 doubly mutant proteins for their ability to confer resistance to ampicillin in vivo and used electrophoretic blots stained with antibodies against ß-lactamase to estimate the relative quantities of the mutant enzymes in whole cell extracts from E. coli.. The results indicate that

although several of the doubly mutant enzymes retain catalytic activity (including the Thr 71 --> Ile, Cys 77 --> Ser mutant), these mutants are much less stable than the analagous single mutants that contain a disulfide bond. Apparently, the presence of the disulfide bond in ß-lactamase enables the protein to partially overcome instabilities arising from mutations at residue 71. The results presented here also demonstrate that Thr at residue 71 performs a unique and essential role in the structure and stability of ß-lactamase.

Materials & Methods

Enzymes and Chemicals

Restriction enzymes and the large (Klenow) fragment of DNA polymerase I were purchased from Boehringer Mannheim. T4 DNA ligase and T4 polynucleotide kinase were obtained from Bethesda Research Laboratories. All antibiotics were from Sigma. [α-^{32}P]dTTP, 3000 μCi/mmole (1 Ci = 37 GBq), was purchased from Amersham.

DNA

Oligonucleotides were synthesized by using the phosphoramidite chemistry (20) on the Applied Biosystems (Foster City, CA) DNA synthesizer, model 380A and purified by preparative polyacrylamide gel electrophoresis.

Plasmid DNA was prepared from E. coli LS1 (21) cultured in LB medium (22) using standard procedures (23). DNA fragments were purified from 1.2% agarose gels using the International Biotechnologies, Inc. UEA electroelutor. Numbering of pBR322 nucleotides is the one commonly used (22).

Reactions

Individual oligonucleotide strands were phosphorylated using standard procedures (22). The oligonucleotides were annealed by mixing 0.4 pmole/μl of each strand and heating to 95^{o}C in 10 mM $MgCl_2$ / 50 mM Tris-HCl, pH 8, followed by gradual cooling to 20^{o}C over a 45 minute period.

For ligations, approximately 0.04 pmole of each restriction fragment and 0.4 pmole of the synthetic fragment, where appropriate, were mixed in 10 mM $MgCl_2$ / 50 mM Tris-HCl, pH 8

/ 0.5 m<u>M</u> ATP / 5 m<u>M</u> dithiothreitol and 5 units of T4 DNA ligase and incubated at 15°C for 16 hours. The reaction mixture was extracted with phenol and the DNA was precipitated by addition of ethanol. The DNA was redissolved in 20 μl of 1 m<u>M</u> Tris-HCl, pH 8 / 0.1 m<u>M</u> EDTA; an aliquot (5μl) was used to transform competent <u>E. coli</u> LS1. Cell transformations were done using standard procedures (24).

<u>Mutagenesis</u>

I introduced the Cys 77 --> Ser mutation into the ß-lactamase gene by cloning the following synthetic oligonucleotide segment into the Sca I and HincII sites in the ß-lactamase gene (see Figure 1).

```
                  **        *
5'- A-CTG-CTA-TCA-GGC-GCC-GTA-TTA-TCC-CGT-GTT
    T-GAC-GAT-AGT-CCG-CGG-CAT-AAT-AGG-GCA-CAA
```

After ligation and transformation, cells were grown in the presence of ampicillin (200 mg/l).

I combined the Cys 77 --> Ser mutation with the 19 mutations at residue 71 by combining the 2568 b.p. Sty I to Sca I fragment from plasmid pBR322-DCS4 with the 1793 b.p. Sca I to Sty I fragment from the appropriate residue 71 mutant (see Figure 2). After ligation and transformation, cells were grown in the presence of tetracycline (15 mg/l).

<u>DNA Sequencing</u>

The plasmids were digested with Ava I and the resulting 2962 b.p. fragments were isolated as previously described. The fragment was labeled at nucleotide 3973 using [α-^{32}P]dTTP

and the Klenow fragment of DNA polymerase I. The labeled fragments were sequenced by using standard techniques (25).

Antibiotic Concentration Gradients

An antibiotic concentration gradient was generated in L agar plates by elevating one end of a Petri dish (8.5 cm diameter) 5 mm and pouring 15 ml of L agar (22) containing an appropriate concentration of antibiotic into the tilted dish. These were allowed to dry overnight. The plates were then placed on a flat horizontal surface and 15 ml of L agar was poured on top; they were used immediately upon hardening. An aliquot (50 μl) of a $1:10^5$ dilution of a saturated culture (approximately 2×10^9 cells per ml for a saturated culture of E. coli LS1) was spread over half of the plate; the Thr 71 --> Gln mutant (18) was plated on the other half as a standard.

Protein Gels and Electrophoretic Blots

Whole cell extracts were prepared for E. coli. harboring plasmids encoding either mutant or wild type ß-lactamases. The cells were grown to late log-phase (O.D. at 600 nm ~ 1) and the absorbance at 600 nm was measured. The cells were pelleted by centrifugation and resuspended in loading buffer [10% (vol/vol) glycerol / 5% (vol/vol) 2-mercaptoethanol / 3% (wt/vol) sodium dodecyl sulfate / 62.5 mM Tris-HCl, pH 6.8 / 1 mM EDTA] and incubated 10 minutes at 95°C. Cells from cultures with slightly higher or lower densities were resuspended in proportionally greater or lesser volumes of loading buffer. A 20 μl aliquot was loaded onto a 15-cm

12% polyacrylamide gel with a 2 cm 4% stacking gel and electrophoresed at 5 mA for approximately 12 hours.

After electrophoresis, protein was transferred to nitrocellulose according to the Bio-Rad procedure (26) as described in Chapter II. The protein blots were stained according to the Vectastain procedure (27) using antibody raised against ß-lactamase in rabbits as described in Chapter II.

Results

The synthetic segment used to introduce the Cys 77 --> Ser mutation also introduced a silent mutation that results in a Nar I restriction site at 3924; these base changes are designated by asterisks in the synthetic oligonucleotide segment given in the Materials and Methods section. After ligation of the synthetic segment into pBR322-CR7, transformed cells were spread onto L-agar containing 200 mg/l of ampicillin; this selection eliminated plasmids that lacked a synthetic segment or those that contained either inverted or repeated synthetic segments. Plasmids were isolated from cultures derived from nine of the ampicillin resistant colonies. Restriction mapping with Nar I indicated that all nine of these colonies harbored plasmids that had received the synthetic fragment. The sequence was verified by sequencing two of the mutants (pBR322-DC9 = Cys 77 --> Ser and pBR322-DC71 = Thr 71 --> Gln, Cys 77 --> Ser) through the region in which the synthetic fragment was introduced.

After introducing the Cys 77 --> Ser mutation into each of the 19 mutants at residue 71, three tetracycline resistant colonies from each of these 19 sets of double mutants were used to inoculate L-agar plates containing 200 mg/l of ampicillin; these plates were incubated at 30°C and 37°C. For the Cys 77 --> Ser mutants, only cells containing enzyme with Thr at residue 71 grew at 37°C. However, cells containing Cys 77 --> Ser mutants with Val, Thr, Ala, Cys, Leu, Pro, Ile, Ser, His, and Met at residue 71 grew at

30°C; these same strains grew when the plates contained 500 mg/l of ampicillin and were incubated at 30°C and at room temperature. Cys 77 --> Ser mutants with Gly, Lys, Phe, Tyr, Glu, Asp, Arg, Asn, Gln, and Trp did not confer cellular resistance to ampicillin for any of the conditions tested.

More accurate values for the level of antibiotic resistance provided to cells by the mutant ß-lactamases were determined by spreading cells onto an agar plate with a continuous concentration gradient of antibiotic. The maximal concentrations at which colonies were established for each of the mutants are listed in Table I. Cells containing the Thr 71 --> Cys, Cys 77 --> Ser mutant grew at 30°C when picked and inoculated onto L-agar plates containing 500 mg/l of ampicillin, but not when spread onto gradient plates; this probably resulted from the inoculum effect and, consequently, this mutant is designated as having trace activity in Table I.

Whole cell extracts from approximately 2 X 10^8 E. coli. cells harboring the various wild type and mutant plasmids incubated either at 30°C or 37°C was electrophoresed in denaturing sodium dodecyl sulfate / polyacrylamide gels and then transferred to nitrocellulose. Figure 3 shows antibody stains of these blots. All of the mutant ß-lactamases are present, but the quantity is greatly reduced, even at 30°C, when compared to the wild type or Cys 77 --> Ser enzymes.

Discussion

Many factors influence the level of antibiotic resistance observed for cells containing the wild type and mutant plasmids, including: plasmid copy number, rate of transcription, stability and rate of translation of the mutant mRNA's, stability of pre-ß-lactamase in the cytoplasm, rate of processing and secretion, stability of the protein in the periplasm, and, finally, the intrinsic enzymatic activity.

As discussed in Chapter II, previous results (18) suggest that for the conditions used in our experiment, mutations in this region of the plasmid do not alter plasmid copy number, transcription, or translation. For a given amino acid mutant at residue 71 in ß-lactamase, the maximal level of *in vivo* antibiotic resistance conferred by this mutant is not influenced by the use of various codons for that amino acid; similarly, no apparent differences arise from use of a particular codon in the double mutants described here. If mutations in this region of the plasmid did affect these cell processes, one would expect such differences to exist.

Mutations may, however, affect processing and secretion of proteins (28) and thereby alter the observed antibiotic resistance of cells. Mutants such as Thr 71 --> Ser and Ser 70 --> Thr and the double mutant Ser 70 --> Thr, Thr 71 --> Ser are processed and secreted at a normal rate in *Salmonella typhimurium* (29); the Cys 77 --> Ser mutant is also secreted at a normal rate in this system (D. Botstein, personal com-

munication). Previously, we observed abundant quantities of mature ß-lactamase in cells growing at 30°C that contained any of the 19 mutations at residue 71 (18); in Figure 3, abundant quantities of the Cys 77 --> Ser mutant are present. Figure 3 also shows no buildup of pre-ß-lactamase for any of the mutant enzymes. Together, these facts suggest that the rate of processing and secretion of the various mutant ß-lactamases described here do not differ significantly from each other or from the wild type enzyme.

However, stability of the wild type and mutant ß-lactamase protein in the periplasm, which will affect the quantity of enzyme capable of catalysis, obviously affects the antibiotic resistances observed here. Figure 1 shows that for the Cys 77 --> Ser mutant, all of the mutations at residue 71 result in a dramatic decrease in proteolytic stability for the protein, even at 30°C. The stability of these mutants apparently is decreased greatly from that of the single mutants at residue 71 (with the disulfide bond intact) (18) and from that of the Cys 77 --> Ser (with Thr at residue 71); the Cys 77 --> Ser mutant is apparently more stable than any of the residue 71 mutants. The apparent order of stability, therefore, is: wild type > Cys 77 ⟶ Ser > single mutants with Thr 71 ⟶ any other amino acid > double mutants with Cys 77 ⟶ Ser, Thr 71 ⟶ any other amino acid.

The intrinsic enzymatic activity will also affect the antibiotic resistance of cells containing the mutant ß-lactamases. Precise values can be obtained only with purified

protein; however, the instability of these mutants makes purification and characterization difficult. The data in Table I show that several of the mutant ß-lactamases retain catalytic activity when the disulfide bond is removed. Mutant enzymes with Gly, Asn, Gln, Glu, and Phe at residue 71, although active when the disulfide bond is present (18), do not confer ampicillin resistance to cells when the Cys 77 --> Ser mutation is introduced. Mutant enzymes with Ala, Val, Leu, Ile, Met, Pro, His, and Ser at residue 71 all have low levels of activity on ampicillin; the Thr 71 --> Cys, Cys 77 --> Ser mutant has only trace activity. Apparently, these mutant enzymes at least partially maintain a population of active conformations; however, from these experiments we cannot determine whether the k_{cat} and/or K_m values for the mutant enzymes are intrinsically altered or whether the relative amount of enzyme in active and inactive conformations is changed.

Ampicillin resistance provided by the Thr 71 --> Ile, Cys 77 --> Ser mutant is greater than that provided by double mutants with the other nonpolar amino acids at residue 71; this may result from slightly greater stability, which is similar to observations for the Thr 71 --> Ile single mutant (18). However, this result is in contrast to the S. aureus PC1 ß-lactamase in which Thr 71 --> Ile results in an inactive enzyme (18). Structural studies of the PC1 mutant indicate that the enzyme is not folded properly; the authors propose that this mutation blocks a late step in the folding

process such that the enzyme cannot assume its active conformation (19). For the RTEM-1 Thr 71 --> Ile, Cys 77 --> Ser mutant, either a different folding mechanism is operating or the structure of this enzyme allows it to accommodate mutations at residue 71 more readily than the PC1 enzyme.

Interestingly, the mutant with Thr 71 --> His, Cys 77 --> Ser is resistant to higher levels of ampicillin than the mutant with Thr 71 --> Ser, Cys 77 --> Ser; the single mutants Thr 71 --> His and Thr 71 --> Ser are phenotypically indistinguishable on ampicillin (18). One would not expect His to be a suitable replacement for Thr structurally or catalytically, certainly not comparable to or better than Ser.

The results presented here suggest two general conclusions. First, the disulfide bond between Cys 77 and Cys 123 in the RTEM-1 ß-lactamase encoded by pBR322 can stabilize the enzyme when mutations are introduced at residue 71. One could speculate that since the disulfide bond is separated from Thr 71 by only five residues, instabilities that might propagate along the peptide chain could be limited by the disulfide cross-link.

A high resolution crystallographic determination of the structure of a class A ß-lactamase has not yet been accomplished. However, the structure of D-alanyl carboxypeptidase-transpeptidase from Streptomyces R61 has been solved at 2.8 A resolution (30). This protein is probably structurally related to the class A ß-lactamases (8) and, indeed,

shows a high degree of structural homology to Bacillus licheniformis 749/C ß-lactamase (31). In the structure of this carboxypeptidase, the residue seven amino toward the carboxy terminal end from the nucleophilic Ser (Ser 70 in ß-lactamase) is positioned in the A α-helix that immediately follows (toward the carboxy terminal end) the loop containing residues believed to be analagous to Ser 70 and Lys 73 in ß-lactamase. By analogy with this structure, Cys 123 would then be positioned in a long (95 amino acid) random coil region that extends from the end of the A helix and connects to the beginning of a ß sheet region (ßA). Interestingly, this random coil segment passes near the region in which Thr 71 would be located. Possibly, therefore, Thr 71 and the disulfide bond may be important for positioning this section of the polypeptide chain.

The second general conclusion suggested by these experiments is that the threonine side chain can perform a unique and essental role in the conformational stability of proteins; apparently both the hydroxyl and methyl groups of Thr 71 are necessary for proper structural stability in ß-lactamase. Even Ser, whose side chain differs from Thr by only a single methyl group, cannot adequately replace Thr. The experiments presented here should reflect more accurately the structural role of Thr 71; in previous experiments with mutations at residue 71 (18) the disulfide bond stabilizes the mutant enzymes. The decrease in stability that results from replacing Thr with any other amino acid is much more

striking in the absence of a disulfide bond.

Conclusion

The results presented here demonstrate that the presence of a disulfide bond stabilizes the residue 71 mutants of ß-lactamase. Of the class A ß-lactamases, only the RTEM enzymes contain a disulfide bond and, interestingly, the RTEM enzymes are the least homologous to other class A ß-lactamases; one could speculate that the presence of a disulfide bond could allow greater freedom for structural variation arising from mutations.

Site saturation of residue 71 in the Cys 77 --> Ser mutant of ß-lactamase has again produced some surprising results. For example, His apparently can replace Thr more effectively than Ser at residue 71; the Thr 71 --> Cys, Cys 77 --> Ser mutant exhibits only very low levels of activity, unlike its singly mutant counterpart; and the Thr 71 --> Ile, Cys 77 --> Ser mutant again exhibits higher levels of activity than substitution with other nonpolar residues.

References

1. Ziegler, D.M. (1985) Ann. Rev. Biochem. 51, 305-329.
2. Givol, D., DeLorenzo, F., Goldberger, R.F. & Anfinsen, C.B. (1965) Proc. Natl. Acad. Sci. USA 53, 676-684.
3. Thorton, J.M. (1981) J. Mol. Biol. 151, 261-287.
4. Richardson, J. (1981) Adv. Prot. Chem. 34, 167-339.
5. Haber, E. & Anfinson, C.B. (1962) J. Biol. Chem. 237, 1839-1844.
6. Creighton, T.E. (1978) Progr. Biophys. Mol. Biol. 33, 231-297.
7. Plunkett, G. & Ryan, C.A. (1980) J. Biol. Chem. 255, 2752-2755.
8. Ambler, R.P. (1980) Phil. Trans. Roy. Soc. London 289, 321-331.
9. Pollitt, S. & Zalkin, H. (1983) J. Bacteriol. 153, 27-32.
10. Barthelemy, M., Peduzzi, J. & Labia, R. (1985) Ann. Inst. Pasteur/Microbiol. 136A, 311-321.
11. Dalbadie-McFarland, G. & Richards, J.H. (1983) Ann. Rep. Med. Chem. 18, 237-245.
12. Smith, M. (1985) Ann. Rev. Genet. 19, 423-462.
13. Wang, A., Lu, S.-D. & Mark, D.F. (1984) Science 224, 1431-1433.
14. Liang, S.-M. Thatcher, D.R., Liang, C.-M. & Allet, B. (1986) J. Biol. Chem. 261, 334-337.
15. Villafrance, J.E., Howell, E.E., Voet, D.H., Strobel, M.S., Ogden, R.C., Abelson, J.N. & Kraut, J. (1983) Science 222, 782-788.

16. Perry, L.J. & Wetzel, R. (1984) Science 226, 555-557.

17. Perry, L.J. & Wetzel, R. (1986) Biochemistry 25, 733-739.

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

19. Craig, S., Hollecker, M., Creighton, T.E. & Pain, R.H. (1985) J. Mol. Biol. 185, 681-687.

20. Beaucage, S.L. & Caruthers, M.H. (1981) Tet. Lett. 22, 1859-1862.

21. Boyer, H.W., & Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 459-472.

22. Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratories, Cold Spring Harbor NY).

23. Ish-Horowicz, D. & Burke, J.F. (1981) Nucl. Acids Res. 9, 2989-2998.

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

25. Maxam, A.M. & Gilbert, W. (1980) Methods in Enzymol. 65, 499-560.

26. Bio-Rad Laboratories (1984) Trans-Blot Cell Operating Instructions (Bio-Rad Laboratories, Richmond, CA).

27. Vector Laboratories (1984) Immunodetection of Antigens on Nitrocellulose using Vectastain ABC Kit (Vector Laboratories, Burlingame, CA).

28. Kadonaga, J.T., Gautier, A.E., Straus, D.R., Charles, A.D., Edge, M.D. & Knowles, J.R. (1984) J. Biol. Chem. 259, 2149-2154.

29. Dalbadie-McFarland, G., Neitzel, J. & Richards, J.H. (1986) Biochemistry 25, 332-338.

30. Kelly, J.A., Knox, J.R., Moews, P.C., Gilbert, J.H., Bartolone, J.B., Zhao, H., Joris, B., Frere, J.-M. & Ghuysen, J.-M. (1985) J. Biol. Chem. 260, 6449-6458.

31. Kelley, J.A., Dideberg, O., Charlier, P., Wery, J.P., Libert, M., Moews, P.C., Knox, J.R., Duez, C., Fraipont, Cl., Joris, B., Dusart, J., Frére, J.M. & Ghuysen, J.M. (1986) Science 231, 1429-1431.

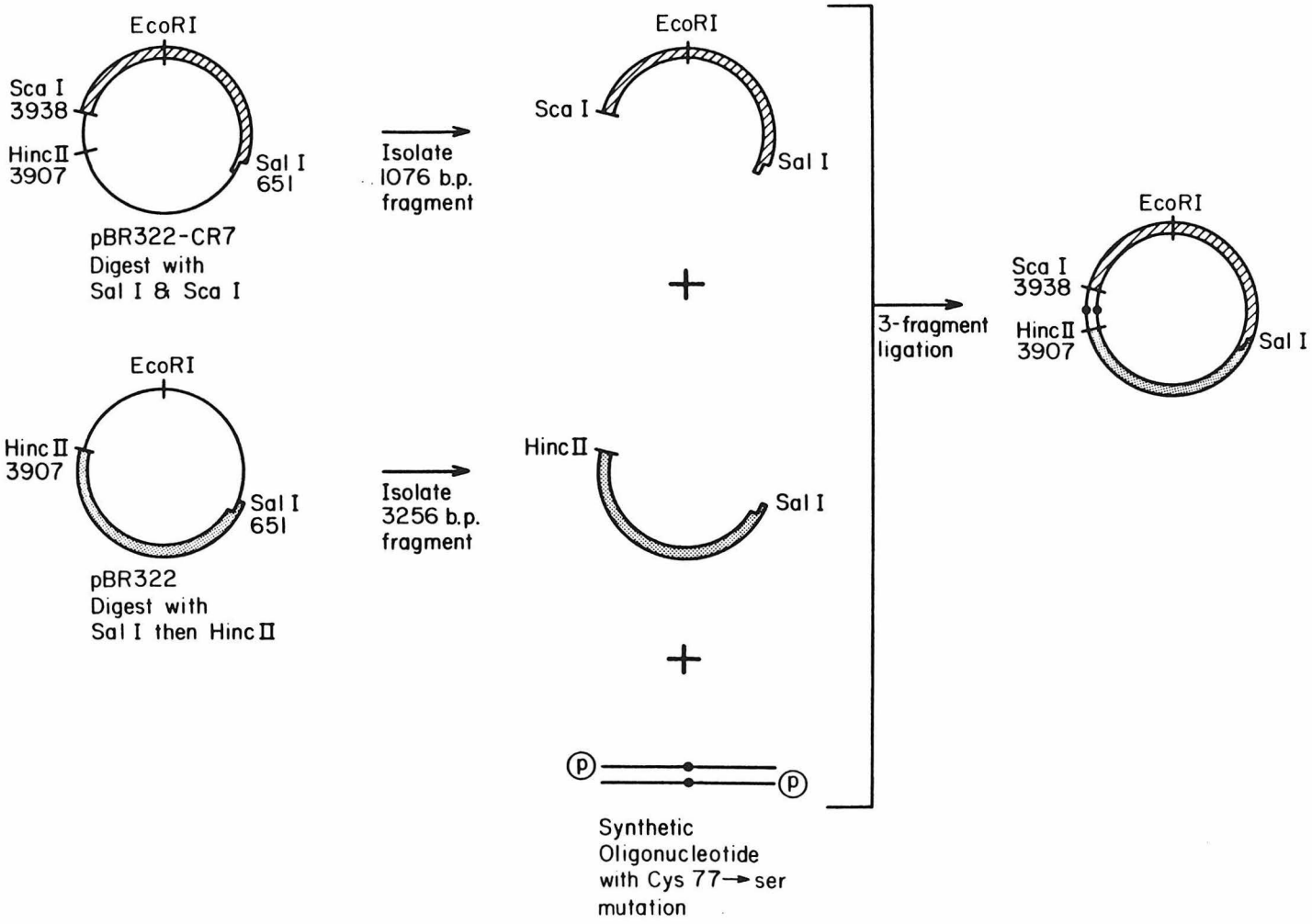


Figure III-1. Three fragment ligation used to introduce the Cys 77 --> Ser mutation into the ß-lactamase gene.

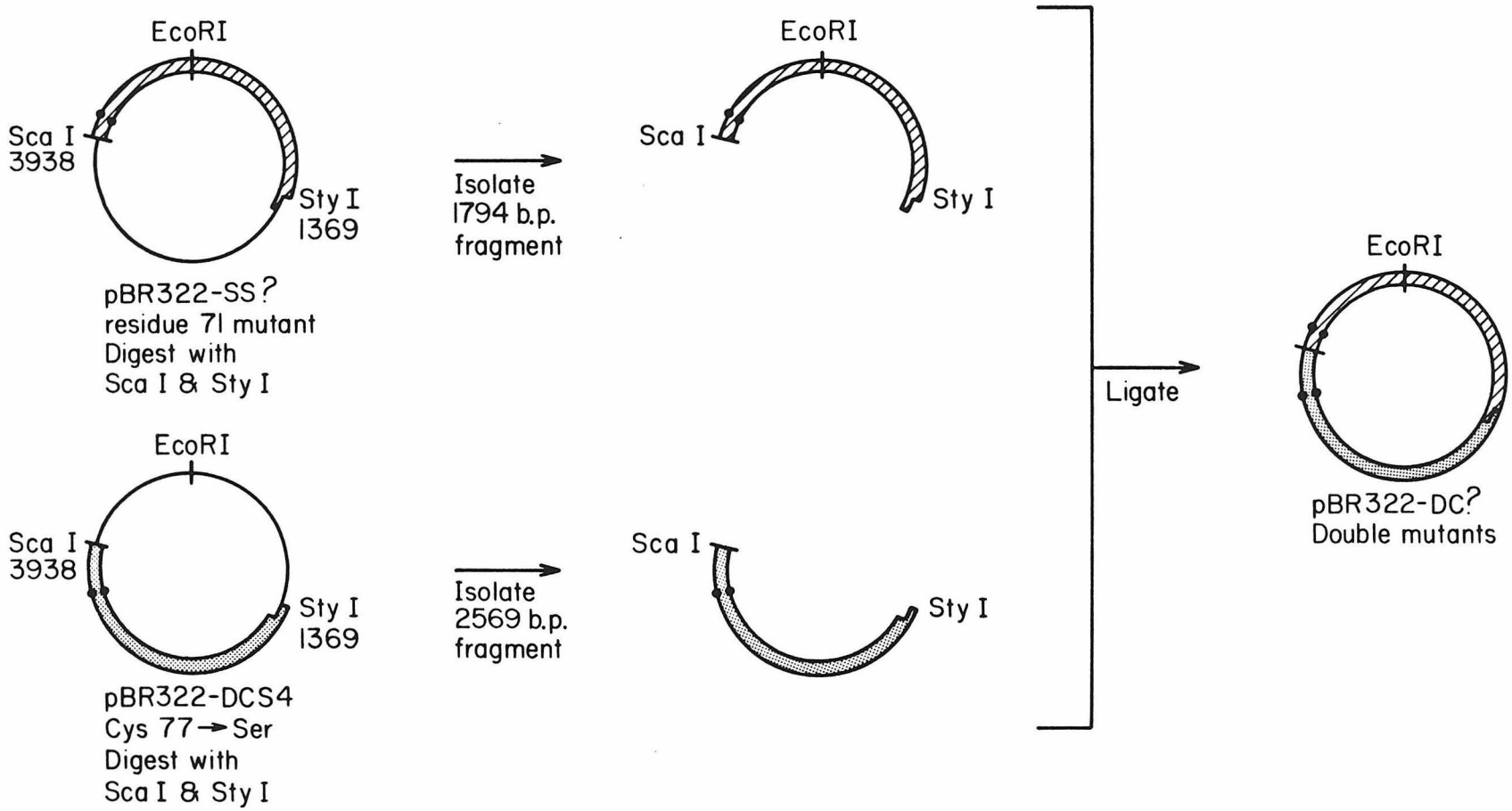


Figure III-2. Design of ligation used to introduce the Cys 77 --> Ser mutation into plasmids containing mutations at the codon for residue 71.

Figure III-3. Electrophoretic blots of the Cys 77 --> Ser, Thr 71 --> ? double mutants of ß-lactamase.

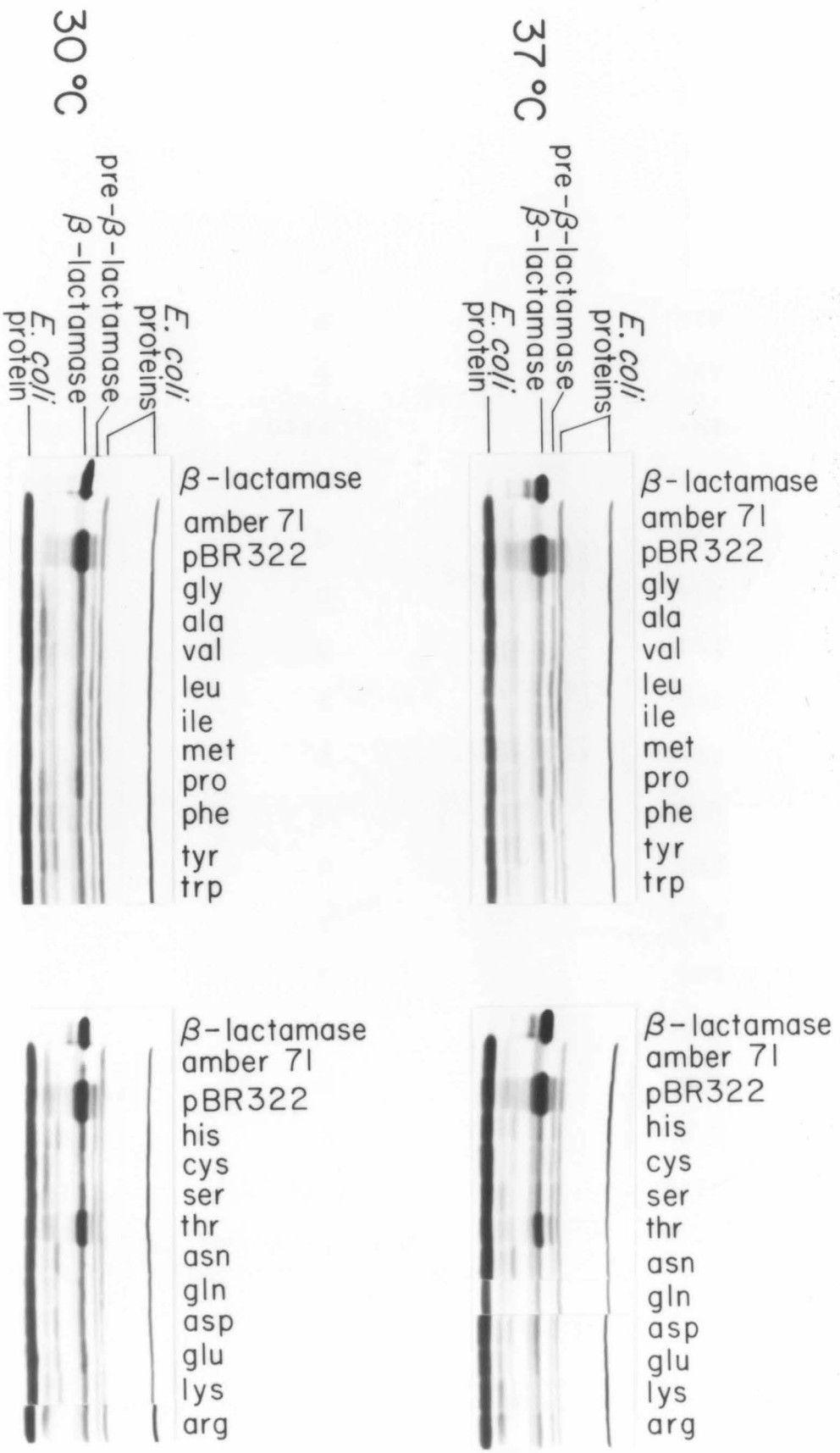

30 °C
E. coli protein
β-lactamase
pre-β-lactamase
E. coli proteins
β-lactamase
amber 71
pBR322
gly
ala
val
leu
ile
met
pro
phe
tyr
trp
37 °C
E. coli protein
β-lactamase
pre-β-lactamase
E. coli proteins
β-lactamase
amber 71
pBR322
gly
ala
val
leu
ile
met
pro
phe
tyr
trp
β-lactamase
amber 71
pBR322
his
cys
ser
thr
asn
gln
asp
glu
lys
arg
β-lactamase
amber 71
pBR322
his
cys
ser
thr
asn
gln
asp
glu
lys
arg

Table I - Maximal level of ampicillin resistance of ß-lactamase Cys 77 --> Ser strains with mutations at residue 71.

	Maximal antibiotic concentration (mg/l) at which colonies grew	
Amino acid	37°C	30°C
Gly	0	0
Ala	0	40
Val	0	70
Leu	0	70
Ile	0	180
Met	0	85
Pro	0	85
Phe	0	0
Trp	0	0
Tyr	0	0
His	0	400
Cys	0	trace
Ser	0	250
Thr	>500	>500
Asn	0	0
Gln	0	0
Asp	0	0
Glu	0	0
Lys	0	0
Arg	0	0

APPENDIX I

CHEMICAL SYNTHESIS OF OLIGONUCLEOTIDES USING THE PHOSPHORAMIDITE METHOD

Abbreviations

DCC - Dicyclohexylcarbodiimide

DCA - Dichloroacitic acid

dA - Deoxyadenosine

dC - Deoxycytosine

dG - Deoxyguanine

dT - Deoxythymidine

DMT - Dimethoxytrityl

TEAoAc - Triethylammonium acetate

Introduction

Several diverse areas of biochemical research such as studies of the structure and dynamics of the DNA double helix (1,2,3), cloning and sequencing of various genes (4,5), oligonucleotide-directed mutagenesis (6,7), and gene synthesis (8,9,10) require a convenient and reliable source of oligonucleotides with specified sequences. Consequently, intense effort has focused on developing efficient, reliable techniques for chemical synthesis of oligonucleotides. The synthesis essentially involves creating a phosphate diester linkage between the 3'-hydroxyl of one deoxyribonucleoside and the 5'-hydroxyl of another. However, since one must repeat the reactions many times to build an oligonucleotide sequence, minor inefficiencies such as incomplete reactions or side reactions (such as reactions with the nucleotide bases) will quickly limit the length and quality of oligonucleotide sequences that one can obtain from chemical synthesis.

Khorana and co-workers developed the first workable oligonucleotide synthesis in the early 1960's (11). The reaction was a DCC or $ArSO_2Cl$ catalyzed esterification of the 3'-hydroxyl group of one nucleoside with the 5'-phosphate of a second nucleotide (Figure 2). Protecting groups prevented reaction at other nucleophilic sites on the purine or pyrimidine base and the deoxyribose ring (Figure 1) (11); current methods for oligonucleotide synthesis utilize these same protecting groups. Although Khorana's method was successful,

each coupling required hours or even days to produce reasonable yields; also, the charged phosphate backbone severely limited the selection of solvents making the chemistry rather difficult.

Letsinger and co-workers solved the latter problem in 1967 with the "phosphate triester" method (12) in which an additional organic ester at the phosphate group resulted in a noncharged backbone. K. Itakura has improved several aspects of this method including the use of 1-(mesitylene-2-sulfonyl)-3-nitro-1,2,4-triazolide as a coupling agent (13) for faster, more efficient reactions; 30-60 minutes produced 90+% yields.

Letsinger and co-workers introduced a faster, more efficient method in 1974 (14) that utilized highly reactive phosphochloridate derivatives of the nucleosides. This "phosphite" method (Figure 2) allowed one to couple two deoxyribonucleotides with greater than 90% yields in less than 5 minutes. However, the instability of the phosphochloridate compounds made this method difficult and unreliable.

Serge Beaucage, working in M. Caruthers group, developed a more reliable method (15) by utilizing phosphoramidite derivatives (more specifically N,N-dimethylaminomethoxyphosphine derivatives) of the nucleosides (Figure 2). These phosphoramidites are stable in basic conditions but become activated in the presence of weak acids and couple with efficiencies greater than 90% in only 3-5 minutes. This phosphoramidite method for synthesis of oligonucleotides has become the method of choice in most groups; although

Itakura's improved triester method is also popular.

The development of methods for solid-phase synthesis of oligonucleotides (16,17,18) was essential for an effective oligonucleotide synthesis. By coupling the first deoxyribonucleotide to solid supports such as polystyrene or silica gel, one could remove reagents by filtration and thereby avoid losses due to purification and isolation at each step of the synthesis.

The general scheme for solid phase synthesis of oligonucleotides using the phosphoramidite chemistry is presented in Figure 3. Briefly, the reactions for each cycle of synthesis are: (i) removal of the 5'-dimethoxytrityl group with a weak acid, (ii) condensation of a new nucleoside phosphoramidite to the 3'-end of the growing oligonucleotide chain (which is bound to the solid support), (iii) capping of unreacted 5'-hydroxyl groups using acetic anhydride, (iv) oxidation of the phosphine backbone with iodine.

I have improved several aspects of the oligonucleotide synthesis using the phosphoramidite chemistry. First, Suzanne Horvath and I have improved the stability of the phosphoramidites (stored under argon as a dry powder) from 2-3 weeks to 8+ months. Second, I have reduced the cycle time for manual sythesis to 35 minutes per cycle. Third, I have increased the efficiency of the coupling reaction to an average of 97 or 98% per cycle. I have synthesized several oligonucleotides for use in a variety of experiments. I have also developed an HPLC method for purifying μmole quantities

of oligonucleotides.

Presently, manual synthesis of oligonucleotides is necessary only when large amounts are required or non-routine syntheses are performed. Several companies, such as Applied Biosystems, Beckman, and Vega have developed extremely efficient machines that routinely synthesize oligonucleotide sequences up to 100 bases in length. The Applied Biosystems machine uses the phosphoramidite chemistry on a solid support consisting of controlled pore glass. Yields above 98% per cycle are routine and very high quality product can be obtained when the machine is properly maintained and operated (19).

Materials & Methods

All solvents and reagents not specified here were from Aldrich Chemical Co.; Aldrich gold label chemicals were purchased if available. Sodium sulfate, methanol, and tetrahydrofuran were from Baker Chemical Co.. Triethylamine was from Kodak (Eastman-Kodak). The silica gel was either Vydac$^{T.M.}$ A (from The Separations Group, Hesperia, CA) or Fractosil (from American Scientific, 300-700 A pore size). The 3 A molecular sieves were from Aldrich and the 4 A molecular sieves were from Linde (Union Carbide). N,N-Dimethylaminotrimethylsilane was from Petrarch Systems, Inc. or Aldrich. Dichloromethoxyphosphine was from Aldrich (this brand was the most reliable). The alumina was Woelm N, Akt 1 (Woelm Pharma, West Germany). Bio-Gel P-4, polyacrylamide, and N,N-methylenebisacrylamide were purchased from Bio-Rad.

Preparation of Solvents

The solvents listed below were prepared as follows:

Toluene	Distilled from calcium hydride
Nitromethane	Distilled from calcium hydride
Acetonitrile	Distilled from calcium hydride
Dioxane	Distilled from calcium hydride
Pyridine	Distilled from p-toluensulfonyl chloride then calcium hydride
2,6-Lutidine	Distilled from p-toluensulfonyl chloride then calcium hydride
Diisopropylethylamine	Distilled from Potassium Hydroxide then calcium hydride

Chloroform	Distilled from phosphorus pentoxide, passed over alumina immediately prior to use
Tetrahydrofuran	400 ml was refluxed over 4.6 g of sodium and 10 g benzophenone until this solution was purple, then distilled
Methanol	Dried over molecular sieves
Dimethylformamide	Dried over molecular sieves

Methanol, chloroform, acetonitrile, and nitromethane were stored over 3 A molecular sieves and toluene, dioxane, puridine, 2,6-lutidine, diisopropylethylamine, and tetrahydrofuran were stored over 4 A molecular sieves.

Derivatization of Silica Gel (18)

1. Preparation of "amino" silica gel.

Vydac A silica gel (12 g) was shaken at room temperature with 13 ml (55.3 mmoles) of 3-aminopropyltriethoxysilane in 120 ml of toluene for 12 hours. Boiling stones were added and the mixture was refluxed for 18 hours. The silica gel was filtered and washed with: 4x50 ml of toluene, 3x50 ml of methanol, and 2x50 ml of 1:1 methanol/water. Methanol/water (1:1, 50 ml) was added and the mixture was shaken overnight. The silica gel was filtered, washed with 2x50 ml of methanol and 2x50 ml of ether and dried overnight in a vacuum desicator. Anhydrous pyridine (50 ml) and trimethylsilyl chloride (7 ml, 55.2 mmoles) were added and the mixture was shaken overnight. The silica gel was filtered, washed with 5x50 ml of methanol and 3x50 ml of ether and dried overnight in a vacuum desiccator.

2. Synthesis of 3'-succinyl nucleosides.

Protected and tritylated nucleoside (5 mmoles) was dissolved in 10 ml of anhydrous pyridine at room temperature; dimethylaminopyridine (0.6g, 4.9 mmoles) and succinic anhydride (0.7g, 7 mmoles) were added and the solution was stirred overnight at room temperature. The pyridine was evaporated (roto-vac) to give a gummy product. Toluene (10 ml) was added and then evaporated (roto-vac) to give a foam. The foam was dissolved in 20 ml of dichloromethane and precipitated in 450 ml of 1:1 hexane/ether at room temperature. The product (a white powder) was filtered and dried overnight in a vacuum desiccator.

3. Coupling of succinylated nucleosides to amino silica gel.

The p-nitrophenyl ester of the succinylated nucleoside was prepared by dissolving succinylated nucleoside (3.5 mmoles), DCC (3.5 mmoles), and p-nitrophenol (3.5 mmoles) in 12 ml of dioxane and 1 ml of pyridine. The solution was shaken for 2 hours at room temperature. Dicyclohexylurea from the reaction was filtered and rinsed with 3 ml of pyridine. The filtrates were combined and added to a suspension of 5 g of amino silica gel in 15 ml of DMF; this mixture was shaken 4 hours at room temperature. The silica gel was filtered and washed with 2x30 ml of DMF, 2x30 ml of dioxane, 5x30 ml of methanol, and 2x30 ml of ether. The derivatized silica gel was dried overnight in a vacuum desiccator.

Synthesis of Nucloeside Phosphoramidites (15)

1. Synthesis of Chloro-N,N-dimethylaminomethoxyphosphine

Dichloromethoxyphosphine (25 ml, 0.266 moles) was stirred in an ice-acetone bath (-10° to -15°C). Dimethylaminomethylsilane (42.1 ml, 0.289 moles) was added dropwise over a two hour period. The reaction was allowed to warm to room temperature and stirred for 1 hour. The product was distilled at 15 torr (b.p. 44-44.5°C). ^{31}P-NMR was used to assess the purity (smoky = -179.5 ppm, dichloromethoxyphosphine = -189.56 ppm in deuterochloroform with respect to 5% aqueous phosphoric acid. Yields were 60-70% of 98+% purity.

2. Synthesis of nucleoside phosphoramidites.

Protected 5'-DMT nucleoside (4 mmoles) was dissolved in 12 ml of chloroform / 3.2 ml of diisopropylethylamine. Smoky (1 ml) was added dropwise over a 10 minute period and the reaction was stirred at room temperature for 15-30 minutes. Ethyl acetate (40 ml) was added and the mixture was washed quickly with 4x60 ml of saturated NaCl. The organic phase was removed, dried over anhydrous sodium sulfate, and evaporated (roto-vac) to give a white foam. The foam was dissolved in 12 ml of toluene and precipitated in 200 ml of hexane in a dry ice-acetone bath. The white powder was dried overnight in a vacuum desiccator. Yields were 60-80%.

The Condensation Cycle (19)

1. Cleavage of the dimethoxytrityl group.

A saturated solution of zinc bromide / 95:5 nitrome-

thane:methanol was used for removal of the dimethoxytrityl group. For a 4 μmole reaction, 1.5 ml of the zinc bromide solution was added three times and shaken for 3 min., 7 min., and 12 min. For 20 μmole reactions, 2.0 ml of the solution was used.

A solution of 3% dichloroacetic acid was used for faster, more efficient removal of the dimethoxytrityl group. For a 4 μmole reaction, 1.5 ml of the solution was twice added and filtered immediately followed by a 1.5 ml, 30 second reaction. A solution of 5% diisopropylethylamine (1 ml) was added immediately after this DCA solution was removed.

2. Condensation reaction.

A 0.5 M solution of tetrazole in acetonitrile was made fresh every week. The nucleoside phosphoramidite (40 μmoles) was dissolved in 0.4 ml of this tetrazole solution immediately prior to condensation. This solution was added to the 4 μmole reaction and shaken for five minutes. For the 20 μmole reactions, 200 μmoles of the nucleoside phosphoramidite was dissolved in 2 ml of the 0.5 M tetrazole solution. Yields ranged from 90-105% as measured by the DMT cation absorbance at 498 nm.

3. Capping reaction.

Acetic anhydride (0.1 ml) was added to 1 ml of a solution of dimethylaminopyridine in tetrahydrofuran / 2,6-lutidine (3.5g DMAP, 20 ml lutidine, 80 ml THF). This was added immediately to the 4 μmole reaction and shaken for 2 minutes.

4. Oxidation

A solution of 0.2 M iodine in THF:2,6-lutidine:water (2:1:1) was added to the reaction and was shaken for 2 minutes; 1.5 ml was used for 4 μmole reactions and 2 ml was used for 20 μmole reactions.

5. Cycle schedule.

Zinc Bromide cycle

1.	Nitromethane wash	2 x 30 sec.
2.	Zinc Bromide cleavage	1 x 3, 7, 12 min.
3.	Methanol wash	3 x 30 sec.
4.	Acetonitrile wash	5 x 30 sec.
5.	Coupling	10 min.
6.	Acetonitrile wash	2 x 30 sec.
7.	Capping	5 min.
8.	Acetonitrile wash	2 x 30 sec.
9.	Oxidation	5 min.
10.	Methanol wash	5 x 30 sec.

DCA cycle

1.	Toluene wash	2 x 30 sec.
2.	DCA cleavage	2 rinses, 1 x 30 sec.
3.	DIPEA wash	30 sec.
4.	Acetonitrile wash	5 x 30 sec.
5.	Coupling	5 min.
6.	Acetonitrile wash	2 x 30 sec.
7.	Capping	2 min.
8.	Acetonitrile wash	2 x 30 sec.
9.	Oxidation	2 min.

10. Methanol wash 5 x 30 sec.

Deprotection (20)

Silica gel with 0.25 μmoles of the desired sequence was shaken for 45 min. in 5 ml of a solution of thiophenol:triethylamine:dioxane (1:2:2). The mixture was centrifuged in a clinical centrifuge and the supernatant was removed. The silica gel was washed with 1x5 ml dioxane, 5x5 ml methanol, and 1x5 ml ether. The silica gel was allowed to dry for 30 minutes. A concentrated ammonium hydroxide solution (1.5 ml) was added and the mixture was shaken for 4 hours. The suspension was centrifuged and the supernatant was removed to another tube; the silica gel was washed with 1 ml of concentrated ammonium hydroxide and this was added to the supernatant; the resulting solution was incubated at 50°C for 15 hours. The ammonium hydroxide solution was removed by evaporation at reduced pressure. Approximately 17 drops of 80% acetic acid was added to the oligonucleotide and this solution was incubated at room temperature for 15 minutes; the acetic acid was evaporated at reduced pressure. The oligonucleotide was redissolved in 0.6 ml of 0.1 M triethylammonium acetate (pH 7.0) and this solution was aliquoted into five tubes, frozen on dry ice, and lyophilized.

Purification

1. Preparative Polyacrylamide Gel Electrophoresis.

The preparative gels were 4 mm thick and approximately 40 cm long; the wells were 2 cm wide. The concentration of

acrylamide used depended on the length of the oligonucleotide: 20% (weight/volume) for less than 30 bases, 18% for 30 to 50 bases, 15% for over 50 bases. Standard 20% polacrylamide-urea gels contained: 150 g urea (50% weight/volume), 150 ml of a 38% acrylamide : 2% bis-acrylamide solution (weight/volume), 30 ml of 1 M TBE, 3 ml of a 10% ammonium persulfate solution (weight/volume) -- this mixture was diluted to 300 ml and 150 μl of TEMED (N,N,N',N'-Tetramethylethylenediamine) was added just before the gel was poured. Approximately 0.05 μmole of the oligonucleotide was dissolved in 40 μl of loading buffer (80% formamide, recrystallized, 50 mM Tris-borate, pH 8.3, 1mM EDTA, 0.1% bromophenol blue) and loaded into one well. The gel was run at 500 volts for 12 to 15 hours or 700 to 1200 volts until the bromophenol blue was at the bottom of the gel (approximately 6 to 8 hours).

The bands were visualized by UV shadowing in which a silica gel plate that contained a florescent indicator (240 nm) was placed beneath the gel and illuminated with a small, hand held short-wave UV light. The desired band was cut out with a razor blade (quickly to avoid UV damage to the DNA) and crushed between weighing paper. The pieces of gel were placed in a 1.5 ml eppendorf tube and covered with approximately 1.0 ml of NACS buffer A (0.2 M NaCl; 10 mM Tris-HCl, pH 7.2; 1 mM EDTA); the oligonucleotides were eluted overnight at 20°C. This mixture was placed in a quick-sep funnel and spun in a clinical centrifuge at setting 5; the supernatant was collected. This solution was loaded

onto a NACS column that was prepared according to the instruction manual (21). The oligonucleotide was eluted 3 times with 200 μl of solution B (1.0 M NaCl in 10 mM Tris-HCl, pH 7.2 / 1 mM EDTA). The eluant was loaded onto a sephadex G-25 spin column equilibrated in 1/10 TE (1.0 mM Tris-HCl, pH 8/ 0.1 mM EDTA) and prepared as follows: sephadex G-25-50 was autoclaved in 1/10 TE and the resin was pipetted into a quick-sep column to give a wet-bed volume of 4 ml; these columns were spun in a clinical centrifuge at setting 3 for 2 minutes and at setting 5 for 3 minutes; the oligonucleotide solution was loaded onto the column and spun at setting 5 for 3 minutes; the column was rinsed by adding 200 μl of 1/10 TE and spinning at setting 5 for 2 minutes; the eluant was then loaded onto a second spin column and eluted as described above. The 200 μl wash from this second column brought the final volume to approximately 1.0 ml.

The absorbance of this solution was measured at 260 and 280 nm. The concentration was calculated as follows (22):

$$\frac{X \text{ O.D. (260 nm)}}{\text{ml}} \times \frac{20\ \mu\text{g}}{\text{ml}\ \ \text{O.D.}} \times \frac{1 \text{ mole base}}{330 \text{ g / mole}} \times \frac{1 \text{ oligo}}{Y \text{ bases}} = \frac{Z \text{ nmoles}}{\text{ml}}$$

2. HPLC purification.

Two types of HPLC purification were developed: one for μmole quantities of double and single stranded oligonucleotides 10 bases in length and a second for sub-μmole quantities of single stranded segments ranging from 14 to 29 bases in length. Both were done on a Waters μbondapak C_{18} reverse phase column (7.8 mm X 30 cm). Buffer A was 0.1 M

triethylammonium acetate, pH 7.0 and buffer B was Aldrich, Gold label acetonitrile. The flow rate was 2.0 ml/minute. All equipment was from Waters: model 6000A pumps, model 440 absorbance detector, model 660 solvent programmer, and model U6K injector.

a. Purification of μmole quantities of oligonucleotides.

Trityl groups were not removed after the deprotection as described previously. Approximately 2 μmole was loaded per injection. A linear concentration gradient of 10 --> 28 % acetonitrile over a 40 minute period gave the best separation. Retention time for the desired double stranded DMT-DNA was approximately 60 minutes; retention times for single stranded DMT-DNA was 35-50 minutes. The desired peak was collected and the solvent was removed by evaporation (roto-vac).

After this initial purification, the oligonucleotide segments were detritylated with 80% acetic acid as described previously. Non-selfcomplementary oligonucleotides were now combined with their complementary strands. Approximately 1.5 μmole of the detritylated, double stranded oligonucleotide was injected into the system and eluted with a linear concentration gradient of approximately 8 --> 12 % acetonitrile over a 60 minute period; the concentration gradient that produced optimal separations had to be determined experimentally for each sequence, however. The desired peak was again collected and the solvent was removed by evaporation (roto-vac).

The DNA was redissolved in 1 ml of 10 mM methyl phosphonate, pH 7.5 and loaded onto a Bio-gel P-6 minus 400 mesh column (1.5 cm O.D. X 60 cm) that had been equilibrated with 10 mM methyl phosphonate, pH 7.5. The column was run under mild air pressure at 6-10 ml per hour. The double-stranded fragment (10 b.p.) eluted after 23 ml of buffer had passed through the column and was contained in a total volume of approximately 2 ml.

b. Purification of sub-μmolar quantities of oligonucleotide.

Approximately 0.05 μmole (one aliquot of the crude mixture from the deprotection as described previously) was purified per injection. A linear gradient of 10 --> 12 % acetonitrile in 60 minutes was generally successful, but precise conditions for optimal separations had to be determined experimentally for each sequence. Retention times varied from 30 to 50 minutes. The desired peak was collected and the solvent was removed by evaporation in vacuo.

Results

While working with Dr. Suzanne Horvath, eighteen oligonucleotides were synthesized. These oligonucleotides varied from 10 to 17 nucleotides long, and from single sequences to mixtures of 64 sequences. Most of these have been used successfully (23,24); in cases where results were not obtained, factors other than the oligonucleotide were at fault. All of the sequences synthesized in Suzanne's lab were made using the $ZnBr_2$ cycle described previously. This method reliably produced highly pure, biologically active oligonucleotides.

Since then, I have synthesized several oligonucleotides in our lab for use by our research group. These are classified by cycle ($ZnBr_2$ or DCA cycle) and scale (4 μmolar or 20 μmolar).

<u>$ZnBr_2$ cycle-4 μmole scale</u>

1. ßL-15 Cys_{77} --> Ser CCACTTAGCAGAACT

Overall yield = 25.4 %
Average yield = 91 %

This sequence was used to convert Cys 77 to a Ser in ß-lactamase (25). The sequence was made with protected DMT-nucleosides from Biosearch. The capping reaction was not used. An analytical polyacrylamide gel showed lighter than normal bands for the sequence and a smear appeared above the band for the 15-mer; this suggests that the capping reaction may be necessary for reasons besides capping unreacted chains.

2. ßL-15 Thr_{71} --> Cys ATGATGAGCTGTTT

Overall yield = 27.9 %
Average yield = 91 %

This sequence was made for converting Thr 71 to a Cys in ß-lactamase. The sequence was made with protected DMTnucleosides from Biosearch. The oligonucleotide showed a normal banding pattern on the gel.

3. ßL-15 Ser_{70} --> Arg ATGATGCGCACTTTT

Overall yield = 46 %
Average yield = 95 %

This sequence was used to convert Ser 70 to an Arg (26) in ß-lactamase. The sequence was made with protected DMT-nucleosides from American Bionuclear. More than the normal amount of product (15-mer) was observed on an analytical polyacrylamide gel.

4. α-LP-14^{16}

```
      A
AACGCCTGCATGG
  T  G  T
     T
```

Overall yield = 19.3 %
Average yield = 89 %

This sequence was used in attempting to isolate the gene for α-lytic protease from Lysobacter enzymogenes. The sequence was made using protected DMT-nucleosides from American Bionuclear. For sites with mixed compositions of bases, the silica gel was split into equal portions and the condensations were done separately. The ratio of bases from these reactions was corrected to be within 5 % of an equimolar mixture when the silica gel was recombined. The

reason for lower than normal yield was an unusually low yield (55 %) in condensation of the G closest to the 5' end of the sequence. I have no explanation for this low yield and it has not occurred before or since.

5. α-LP-11[8] GGCGTAACCAT
T G T

Overall yield = 58 %
Average yield = 94 %

This sequence was also used in attempting to isolate the gene for α-lytic protease from Lysobacter enzymogenes. The sequence was made using protected DMT-nucleosides from American Bionuclear. The sites with mixed compositions of bases were made by mixing equal amounts of the appropriate bases.

$ZnBr_2$ cycle - 20 μmole scale

The following sequences were made using protected DMT-nucleosides from Biosearch, which have since proven impure. These sequences were made for use in structural studies of the DNA double helix using NMR.

		Overall yield	Average yield
6.	GCGCGCGCGC	50.1 %	92 %
7.	GCGCATGCGC	53.9 %	93 %
8.	TATATATATA	62.7 %	95 %
9.	ACACACACAC	52.6 %	92 %

DCA cycle -- 4 μmole scale

The following sequences were made using protected DMT-nucleosides from American Bionuclear.

10. ßL-15 Thr_{71} --> Met GATGAGCATGTTTAA

Overall yield = 65 %
Average yield = 97 %

This sequence was made with Yie-Hwa Chang and used to convert Thr to a Met in ß-lactamase (27).

11. ßL-14 Lys_{73} --> Arg AGCACTTTTAGAGT

Overall yield = 51 %
Average yield = 95 %

This sequence was made with Jim Neitzel and used to convert Lys 73 to and Arg in ß-lactamase (28).

12. C5A-33 GGCGCGCTAACATCTCCCACAAAGACATGCCAA

Overall yield = 25.1 %
Average yield = 96 %

This sequence was made with Michael Emerling and was used in attempting to synthesize the gene for the complement factor C5a (29).

13. C5A-30 CACTCGGGTCAACGGCCCAGTTGCATGTCT

Overall yield = 31.2 %
Average yield = 96 %

This sequence was made with Marilyn Tomich and was used in attempting to synthesize the gene for the complement factor C5a (29).

14. PLP1AT-29 AATTGGTCGACATGATCGACGTTCTGCTG

Overall yield = 23.9 %
Average yield = 95 %

This sequence was made with Sheila Iverson and was used in attempting to synthesize the gene for plastocyanin (30).

15. PLP1AB-25 TCCGTCGTCAGCACCCAGCAGAACG

Overall yield = 39.4 %
Average yield = 96 %

This sequence was made with Sheila Iverson and was used in attempting to synthesize the gene for plastocyanin (30).

16. SB-24 GGCTCGAGATTGACGGCGTAGTAC

Overall yield = 37 %
Average yield = 96 %

This sequence was used in attempting to introduce a Xma III restriction site near the 3'-end of Sindbis virus (31).

17. SB-29 GCGTCTAGATTTTTTTTTTTTTTTTTT

Overall yield = 34.2 %
Average yield = 96 %

This sequence was made with Reef Hardy and was used in attempting to introduce a Sal I site near the 5'-end of Sindbis virus (31).

DCA cycles -- 20 μmole scale -- The following sequences were made using protected DMT-nucleosides from American Bionuclear. These sequences were made for use in structural studies of the DNA double helix using NMR.

		Overall yield	Average yield
18.	GCGCGCGCGC	70 %	96 %
19.	GCGTATACGC	82 %	98 %
20.	GCGTGCACGC	80 %	97 %
21.	GCGCACGCGC	73 %	96 %
22.	GCGCGTGCGC	75 %	97 %
23.	CGCGTACGCG	79 %	97 %
24.	CACACACACA	87 %	98 %
25.	TGTGTGTGTG	79 %	97 %

Purification

Oligonucleotides purified by preparative poyacrylamide gel electrophoresis consistantly gave one dark band of appropriate length on analytical gels. This method is useful

for purifying sub-μmole amounts of DNA for use in biological experiments and is now routine.

For purifying μmole quantities of oligonucleotides, a preparative HPLC method was developed. Chromatograms (measured by absorbance at 254 nm) are presented in Figures 4 to 9. Figure 4 shows the chromatogram of an analytical quantity of tritylated $(GC)_5$ used in determining conditions for optimal separations. Figures 5 and 6 show chromatograms of samples injected for purification of tritylated and detritylated $(GC)_5$. The tritylated samples were the crude reaction mixtures obtained from deprotection and the detritylated samples were obtained by treating the purified tritylated samples with 80% acetic acid as described previously. Sequences purified in this manner showed a single band of appropriate length on analytical gels. Figures 7, 8 and 9 show chromatograms of HPLC purifications of 0.05 μmole quantities of ßL-15 Thr_{71} --> Met, ßL-14 Lys_{73} --> Arg, and SB-24. These purified sequences showed a dark band of appropriate length on analytical gels, but several, very faint bands were also present.

Discussion

The general scheme for solid phase synthesis of oligonucleotides using the phosphoramidite chemistry is presented in Figure 3. We have focused primarily on three aspects of the synthesis: improving the stability of phosphoramidite derivatives of deoxyribonucleosides, decreasing the time required for the reaction cycle, and increasing the efficiency of each reaction in the cycle.

Previously, the stability of deoxyribonucleoside phosphoramidites was unpredictable; one to five weeks after synthesis, the deoxyribonucleoside phosphoramidites would suddenly become less reactive, resulting in low yields in the coupling reaction, and within two days these compounds would be inactive. This rapid decay in activity was a source of extensive frustration and confusion. Since extreme care in handling and storage of the phosphoramidite derivatives apparently had not improved their stability, we decided to investigate the quality of the nucleoside phosphoramidites and the reagents used in their synthesis.

Dichloromethoxyphosphine was purchased from several sources (including Aldrich, American Bionuclear, and Specialty Organic Chemicals) and was found to contain from 2 to 20 % PCl_3. We carefully distilled the dichloromethoxyphosphine until it was 99 % pure (by NMR). This purified dichloromethoxyphosphine was then used to synthesize the chloro-N,N-dimethylaminomethoxyphosphine as described in the experimental section of this chapter. Careful fractional distillation of

chloro-N,N-dimethylaminomethoxyphosphine at reduced pressure (15 mm Hg) immediately after it was synthesized produced greater than 98 % purity (by ^{31}P-NMR). Nucleoside phosphoramidites synthesized with this chloro-N,N-dimethylaminomethoxyphospine were stable for eight or more months. The loss of activity that previously had plagued the oligonucleotide synthesis has not been observed since the purity of the chloro-N,N-dimethylaminomethoxyphosphine has been established. We believe that a by-product produced during or carried through the synthesis of this compound, and which has a boiling point very near that of chloro-N,N-dimethylaminomethoxyphosphine was responsible for the instabilities previously observed - perhaps dichloro-N,N-dimethylaminophosphine or some compound that could readily form an acid which will activate the phosphoramidites and result in hydrolysis or other undesirable side reactions.

More recently, several groups have developed phosphoramidite derivatives that are more stable than the dimethylamino compound. The N,N-diisopropylamino phosphoramidites (32) have replaced the N,N-dimethylamino derivatives in most laboratories. A ß-cyanoethyl derivative (33) is also more stable than the dimethyamino phosphoramidite and also makes the thiophenol reaction in deprotection of the oligonucleotide unnecessary; slightly longer coupling times are required for this derivative, however.

We have also studied methods for optimizing each step of the reaction cycle. The first step, cleavage of the dime-

thoxytrityl (DMT) group from the 5' end of the preceeding base, previously utilized the Lewis acid $ZnBr_2$ (34), since strong protic acids can lead to depurinization of oligonucleotides. Although removal of the DMT group with $ZnBr_2$ is sufficiently effective, the reaction is slow, requiring 15 to 25 minutes for a complete reaction. Trichloracetic acid is fast and has been used extensively, but substantial amounts of depurinization can occur during prolonged exposures (35). We have used dichloroacetic acid (pK_a's: TCA = 0.64, DCA = 1.26) which cleaves the trityl group in less than 1 minute. A solution of 5% diisopropylethylamine in toluene was added immediately after the cleavage reaction to quickly neutralize any remaining acid. Treatment of the αLP-14^{16} sequence with 3% DCA in toluene for one hour produced no detectable change in the banding pattern when the oligonucleotide was deprotected and run on an analytical polyacrylamide gel. This method of cleaving the DMT group has greatly reduced the cycle time and also has resulted in slightly greater yields per cycle.

For the coupling reaction, tetrazole, a very weak acid is added to activate the phosphoramidite immediately prior to coupling (15); this acid probably protonates the lone pair of electrons on the amino nitrogen. The coupling reaction is fast and efficient; it is complete within 3 minutes. Slightly longer coupling times were used to allow for inconsistancies in mixing. The first two couplings frequently required longer (10-15 minutes) coupling times; the source of

this effect is unknown.

The capping reaction was thought to be optional. However, our results indicate that besides capping unreacted 5'-hydroxyl groups, this reaction may serve an additional function. When this reaction was omitted, an analytical polyacrylamide gel showed a dramatic decrease in the quantity of oligonucleotide (despite normal reaction yields) and a large smear above the expected position of the oligonucleotide. These results indicate that the capping reaction is suppressing side reactions, perhaps some that result in branching of the oligonucleotide chain.

The oxidation reaction is complete within one minute. The only problem encountered with this reaction occurred when using an iodine solution that was several months old; although the cycles had produced normal yields, no oligonucleotide was observed on an analytical gel after deprotection. The unoxidized phosphate backbone was probably cleaved before or during the deprotection. Treatment of non-deprotected oligonucleotide with fresh iodine solution did not salvage the sequences suggesting either cleavage had occured during the reactions or an intermediate complex that is not easily oxidized had formed. Subsequently, fresh iodine solution was made every month.

The efficiency of coupling in each reaction cycle, as measured by the absorbance of the trityl cation after each cycle, averaged 91% using the initial $ZnBr_2$ procedure (19). All attempts to increase the yields using this pro-

cedure were unsuccessful. However, after several problems with protected nucleosides from Biosearch, protected nucleosides from American Bionuclear were purchased and used in the synthesis. Immediately, average yields of 95% per cycle were obtained. Development of the DCA cycle further increased the yields to 97-98% per cycle. Recently, by using controlled pore glass as the solid support, yields of 99% per cycle have been obtained (19,32).

We have used two methods to purify oligonucleotide sequences synthesized in our lab: preparative polyacrylamide gel electrophoresis and HPLC. For amounts of oligonucleotides less than 1 μmole, both preparative gels and HPLC have been used successfully. Examples of HPLC purification of such quantities are presented in Figures 7 thru 9. Oligonucleotides purified by this method showed a dark band of expected length on analytical polyacrylamide gels, but trace amounts of oligonucleotides with shorter lengths could also be seen. These shorter oligonucleotides were very minor components (well below 1%) and did not interfere with the experiments (27,29,30). Oligonucleotides purified by preparative gel electrophoresis showed no trace of these smaller sequences.

For purifying μmole quantities of oligonucleotides, only the HPLC method is practical; examples of chromatograms from these purifications are shown in Figures 5 and 6. These sequences were first purified while still containing the trityl groups on the 5'-ends of the oligonucleotide to sepa-

rate them from any sequences that were capped or otherwise terminated during the synthesis. The trityl groups were removed after this initial purification and the sequences were again purified by HPLC. Sequences purified in this manner showed only one dark band on analytical polyacrylamide gels.

The success of purification by HPLC requires very high quality instrumentation. C_{18} reverse phase columns were the most successful; other columns including DEAE, SAX, and C_4 separated less well in our system. The affinity of oligonucleotide sequences for the C_{18} resin is very narrowly defined and, consequently, separation of similar sequences is successful only in a very narrow solvent range. High quality pumps are required for such precise gradients. The most frequent problem encountered with our system resulted from minor failures in the check valves; although such valves functioned sufficiently well when using rapid gradients, minor fluctuations in the very narrow gradient required here produced unworkable results.

Presently, manual synthesis of oligonucleotides is necessary only when large amounts are required or non-routine syntheses are performed. Several companies, such as Applied Biosystems, Beckman, and Vega, have developed extremely efficient machines that routinely synthesize of oligonucleotide sequences up to 100 bases in length. The Applied Biosystems machine uses the phosphoramidite chemistry on a solid support consisting of controlled pore glass. Yields above

98% per cycle are routine and very high quality product can be obtained when the machine is properly maintained and operated (19).

REFERENCES

1. Dikerson, R.E., Drew, H.R., Conner, B.N., Wing, R.M., Fratini, A.V., & Kapka, M.L. (1982) Science 216, 475-485.

2. Wang, W.H.-J., Quigley, G.J., Kolpak, F., Crawford, J.L., van Boom, J.H., van der Marel, G. & Rich, A. (1979) Nature (London) 282, 680-686.

3. Weiss, M.A., Patel, D.J., Sauer, R.T. & Karplus, M. (1984) Proc. Natl. Acad. Sci. USA 81, 130-134.

4. Wallis, S.C., Rogne, S., Gill, L., Markam, A., Edge, M., Woods, D., Williamson, R. & Humphries, S. (1983) EMBO J. 2, 2369-2373.

5. Anderson, S. & Kingston, I.B. (1983) Proc. Natl. Acad. Sci. USA 80, 6838-6842.

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

7. Winter, G., Fersht, A.R., Wilkinson, A.J., Zoller, M. & Smith, M. (1982) Nature (London) 299, 756-758.

8. Itakura, K., Hirose, T. Crea, R., Riggs, A.D., Heyneker, H.L., Bolivar, F. & Boyer, H.W. (1977) Science 198, 1056-1063.

9. Rossi, J.J., Kierzek, R., Haung, T., Walker, P.A. & Itakura, K. (1982) J. Biol. Chem. 257, 9226-9229.

10. Brown, E.L., Belagaje, R., Ryan, M.J. & Khorana, H.G. (1979) Meth. Enzymol. 68, 109-151.

11. Khorana, H.G. (1978) Bioorg. Chem. 7, 351-393.

12. Letsinger, R.L. & Ogilvie, K.K. (1967) J. Am. Chem Soc. 89, 4801-4803.

13. Hirose, T., Crea, R. & Itakura, K. (1978) Tet. Lett. 28, 2449-2452.

14. Letsinger, R.L., Finnan, J.L., Heavner, G.A., Lunsford, W.B. (1975) J. Am. Chem. Soc. 97, 3278-3279.

15. Beaucage, S.L. & Caruthers, M.H. (1981) Tet. Lett. 20, 1859-1862.

16. Gait, M.J. (1980) in Polymer-supported Reactions in Organic Synthesis (Hodge, P. & Sherrington, C.C., eds) John Wiley, p. 435-440.

17. Wallace, R.B. & Itakura, K. (1983) Chem. Anal. 66, 631-633.

18. Matteucci, M.D. & Caruthers, M.H. (1980) Tet. Lett. 21, 719-722.

19. Suzanne Horvath, personal communication.

20. Matteucci, M.D. & Caruthers, M.H. (1981) J. Am. Chem. Soc. 103, 3185-3191.

21. NACS Prepac™ Instruction Manual (Bethesda Research Laboratories, Bethesda, Maryland)

22. Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

23. Steinmetz, M., Minard, K., Horvath, S., McNicholas, J., Frelinger, J., Wake, C., Long, E., March, B. & Hood, L. (1982) Nature (London) 300, 35-42.

24. Roach, A., Boylan, K., Horvath, S., Prusiner, S. & Hood, L. (1983) Cell 34, 799-806.

25. Gloria Dalbadie-McFarland, unpublished results.

26. Foster, P. & Dalbadie-McFarland, G., unpublished results.

27. Yie-Hwa Chang, unpublished results.

28. James J. Neitzel, unpublished results.

29. Michael Emerling, unpublished results.

30. Sheila Iverson, unpublished results.

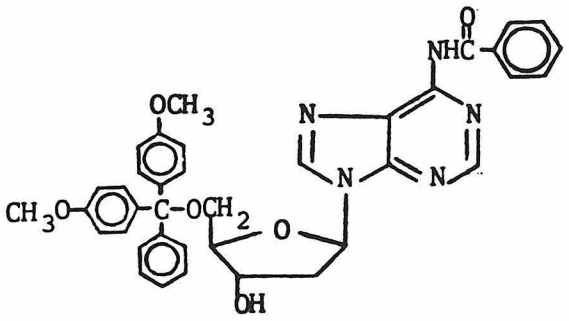
31. Winters Reef Hardy, unpublished results.

32. Adams, S.P., Kavka, K.S., Wykes, E.J., Holder, S.B. & Galluppi, G.R. (1983) J. Am. Chem. Soc. 105, 661-663.

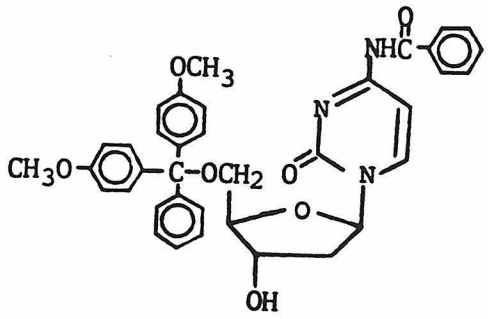
33. Sinha, N.D., Biernat, J., McManus, J. & Koster, H. (1984) Nucl. Acids Res. 12, 4539-4557.

34. Tanaka, T. & Letsinger, R. (1982) Nucl. Acids Res. 10, 3249-3260.

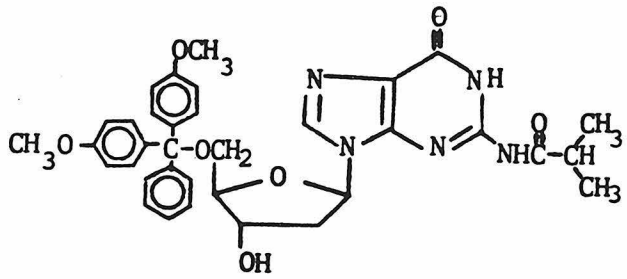
35. Duckworth, M.L., Gait, M.J., Goelet, P., Hong, G.F., Singh, M. & Titmas, R.C. (1981) Nucl. Acids Res. 9, 1691-1706.



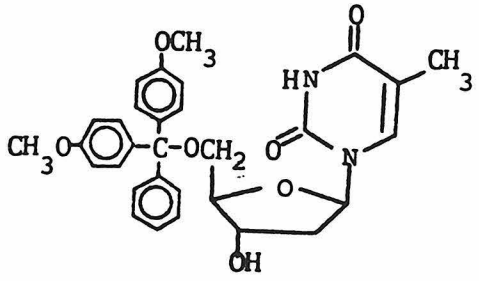
DMT-dA^{bz}



DMT-dC^{bz}

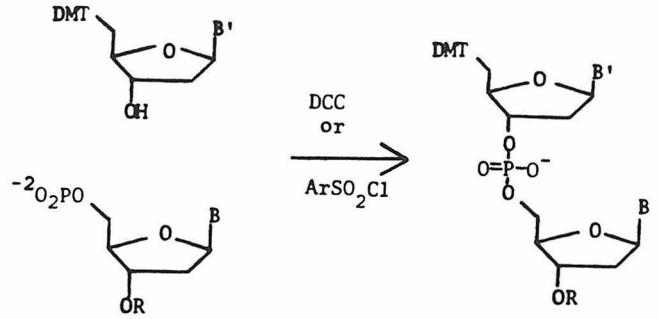


DMT-dG^{iBu}

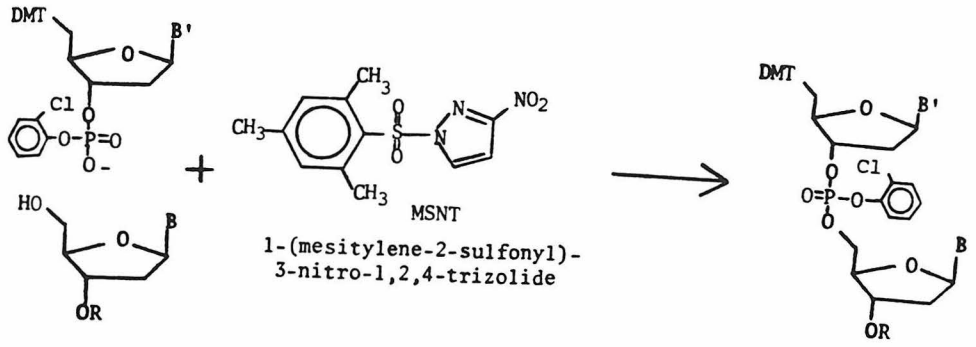


DMT-dT

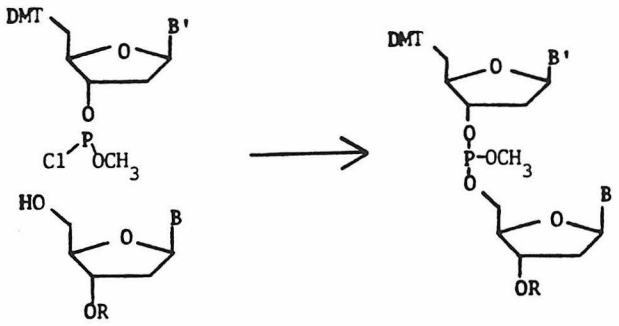
Figure AI-1. Protecting groups of the nucleoside bases.



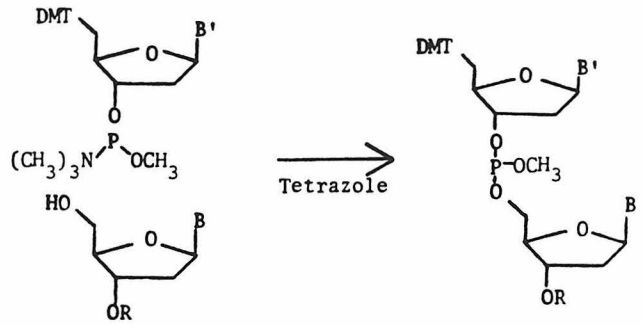
Khorona's coupling.



Triester coupling.



Phosphite coupling.



Phosphoramidite coupling.

Figure AI-2. Schemes for DNA synthesis.

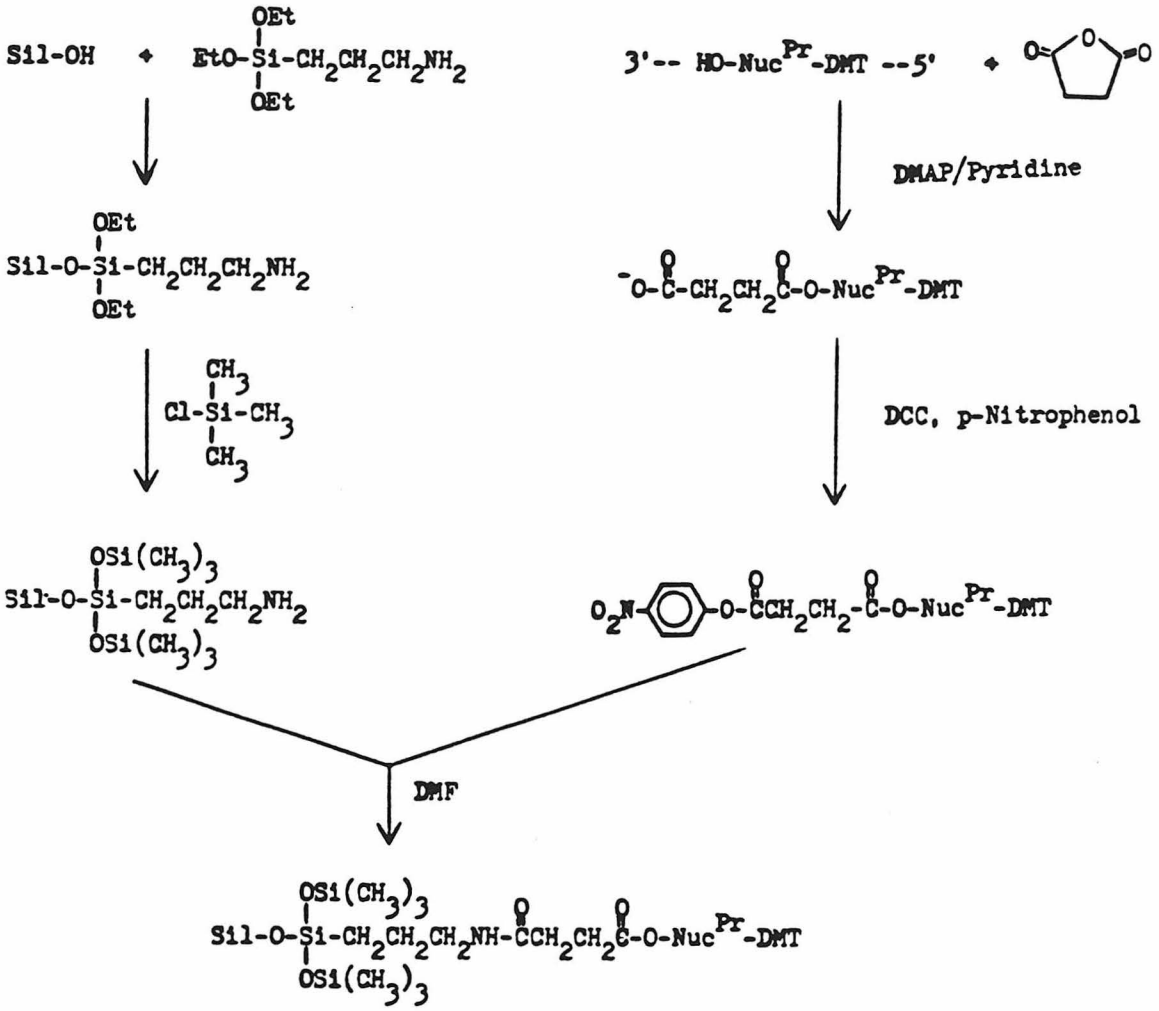
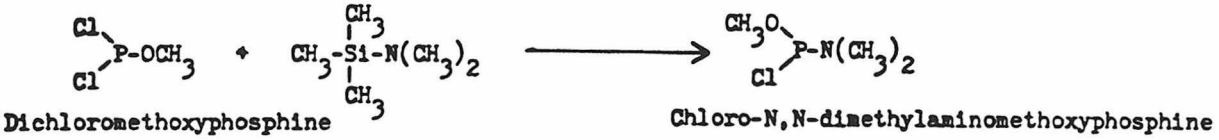


Figure AI-3. Scheme for solid phase synthesis of oligonucleotides using the phosphoramidite chemistry.

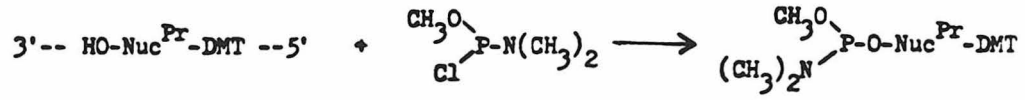
Figure AI-3 (continued)

Synthesis of Deoxynucleoside "Phosphoramidites"



Dichloromethoxyphosphine

Chloro-N,N-dimethylaminomethoxyphosphine



Condensation Cycle

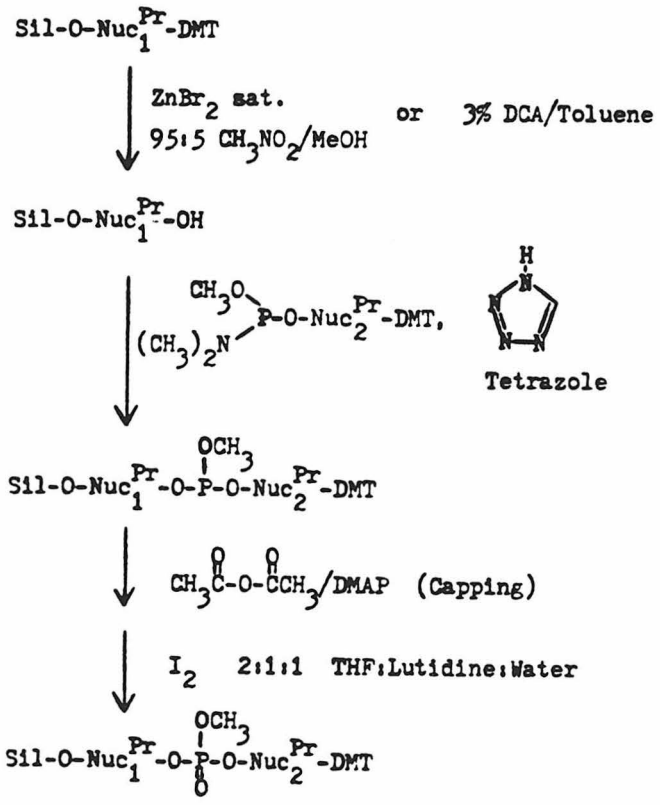
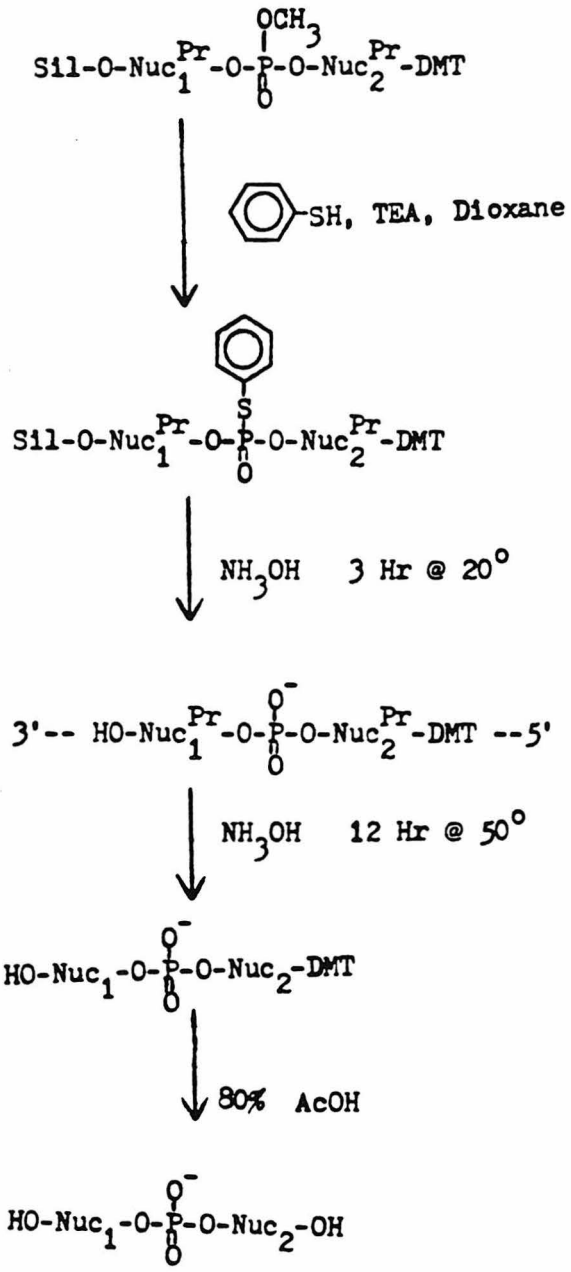


Figure AI-3 (continued)



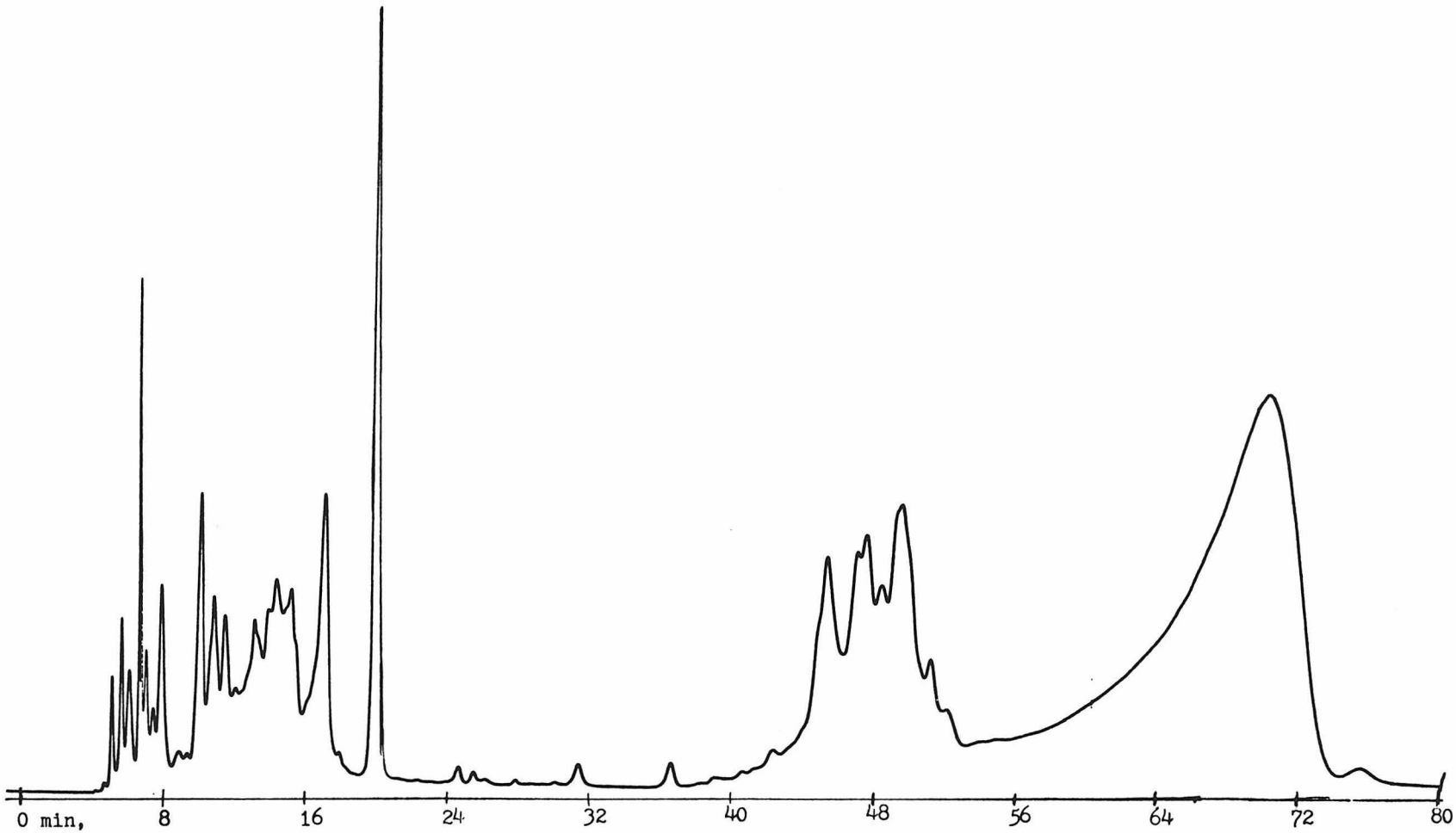


Figure AI-4. HPLC chromatogram of an analytical quantity of 3'-DMT-$(GC)_5$. Gradient was 10 --> 29% acetonitrile/0.1M TEAoAc in 40 minutes.

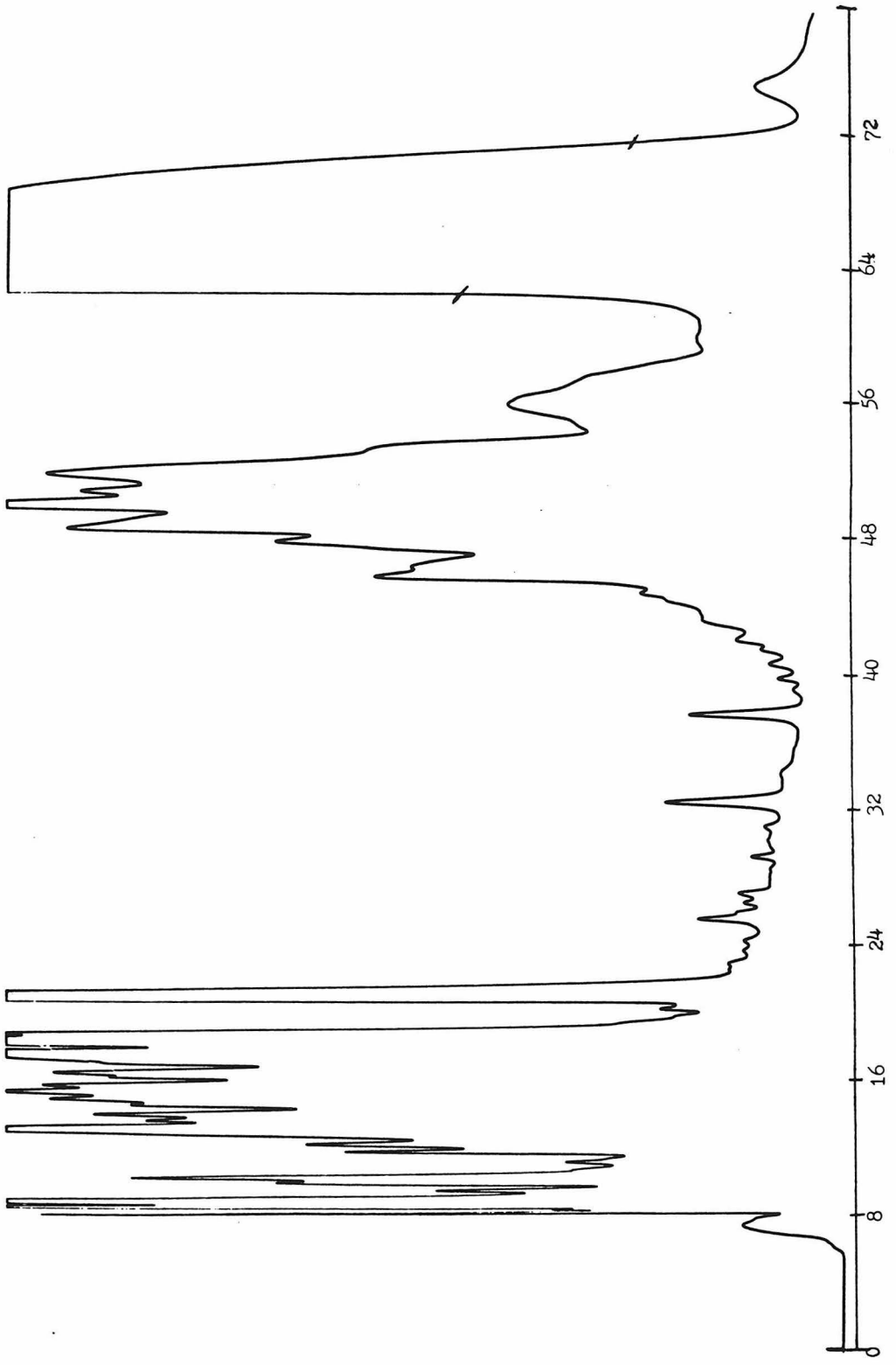


Figure AI-5. HPLC chromatogram of preparative quantity of 3'-DMT-$(GC)_5$. Gradient was 10 --> 29% acetonitrile / 0.1 M TEAoAc in 40 minutes, sensitivity was 2.0.

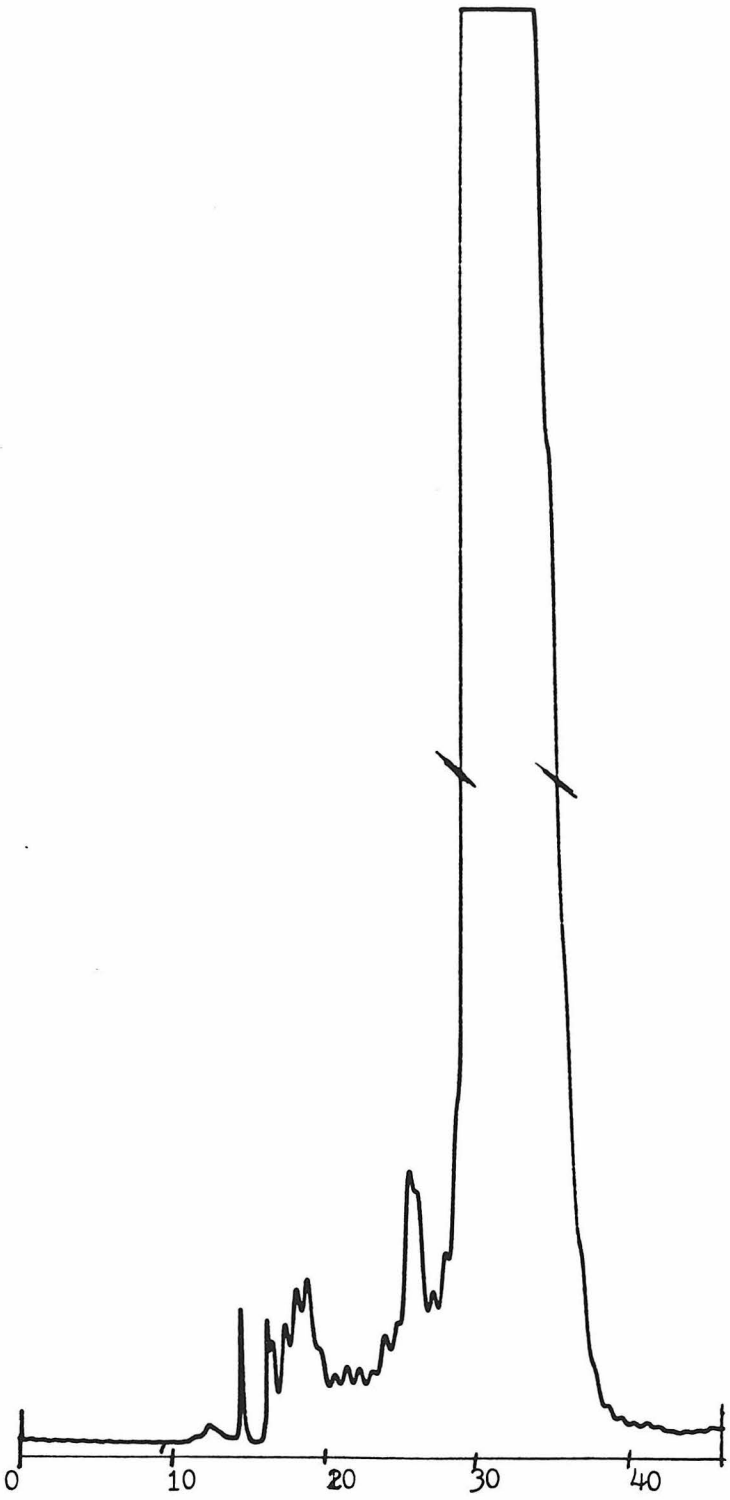


Figure AI-6. HPLC chromatogram of preparative quantity of $(GC)_5$. Gradient was 8 --> 12% acetonitrile / 0.01 M TEAoAc in 50 minutes.

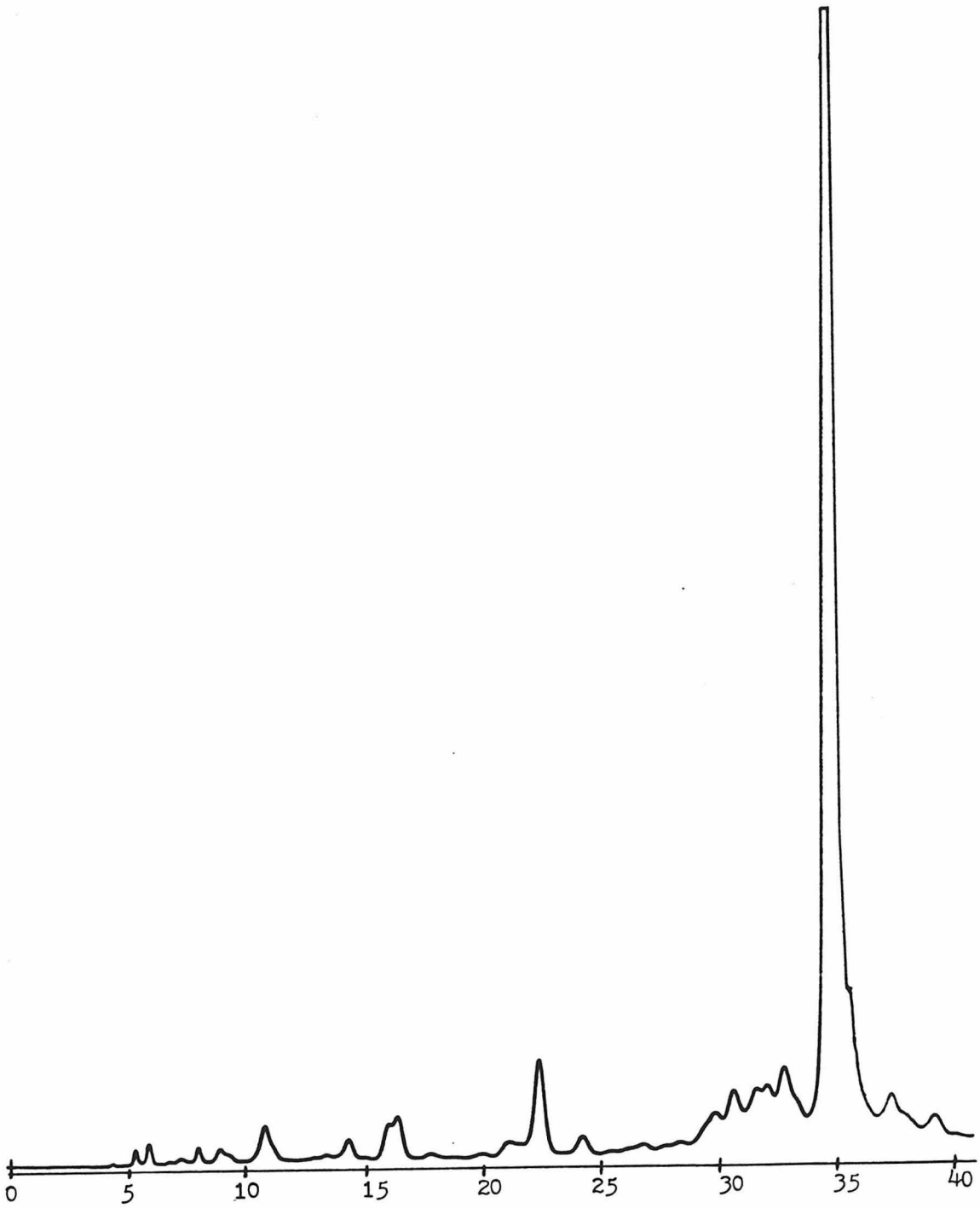


Figure AI-7. HPLC chromatogram of βL-15 Thr_{71}--> Met. Gradient was 10 --> 12% acetonitrile / 0.1 M TEAoAc in 30 minutes.

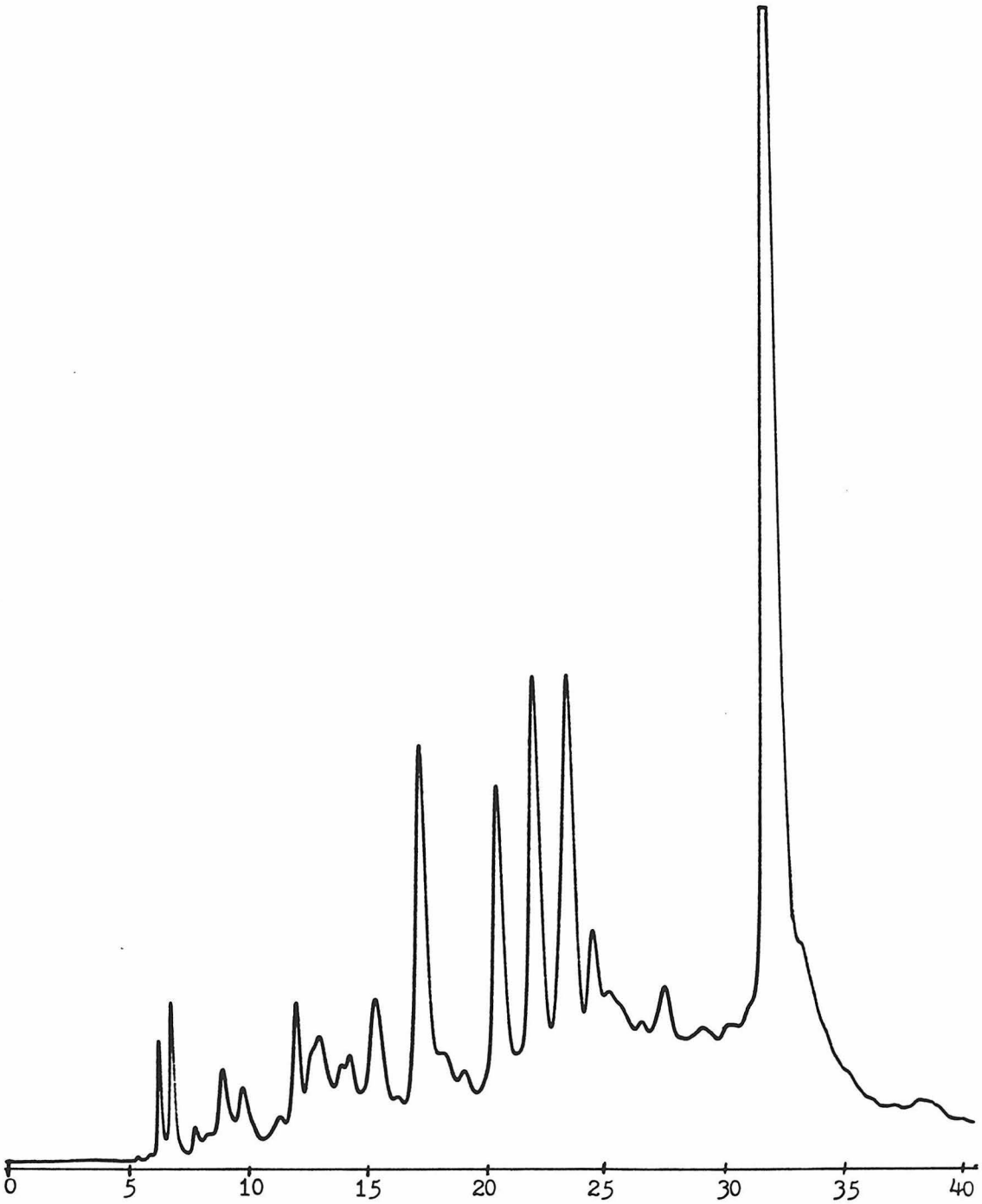


Figure AI-8. HPLC chromatogram of SB-24. Gradient was 10 --> 12% acetonitrile / 0.1M TEAoAc in 20 minutes.

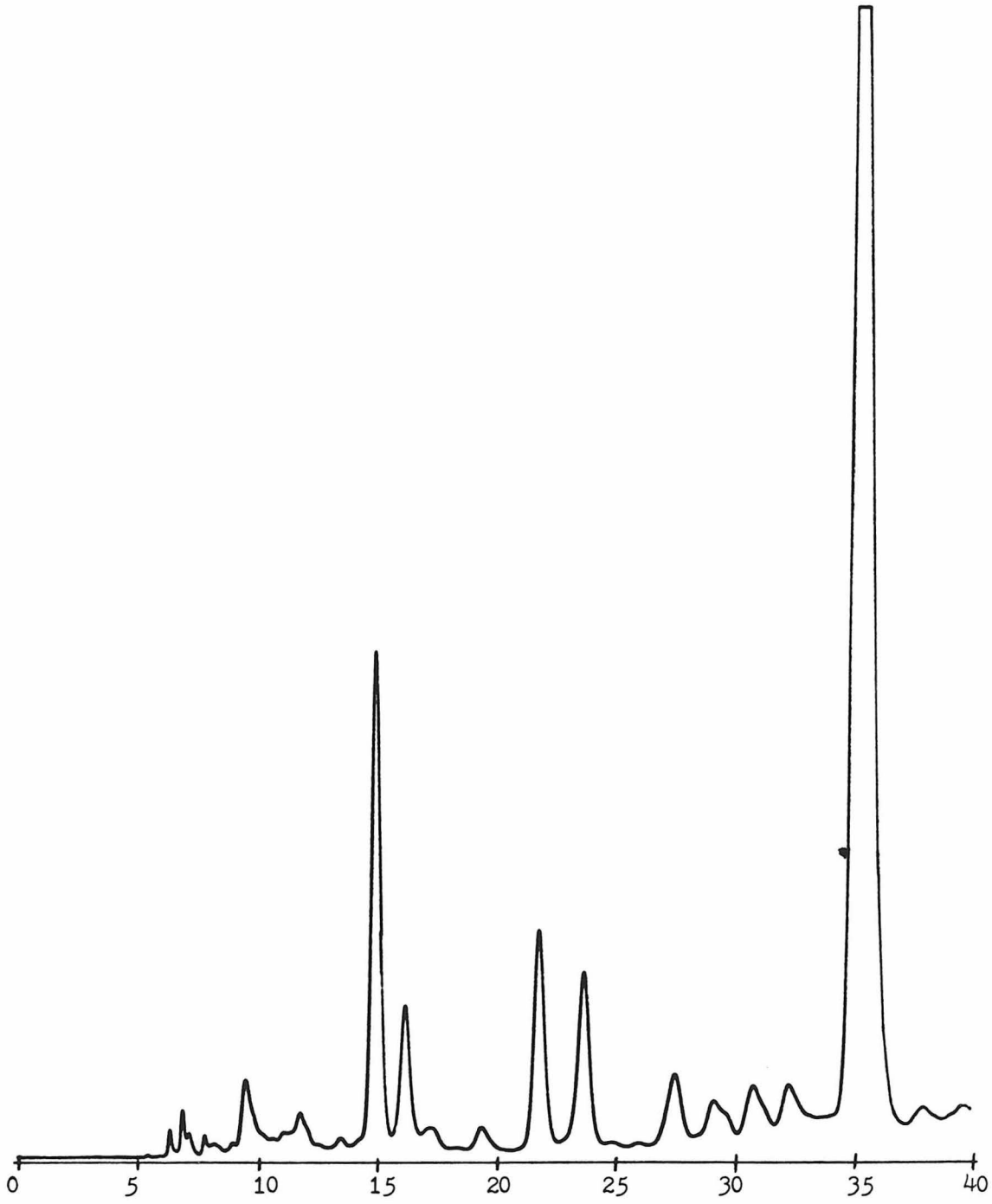


Figure AI-9. HPLC chromatogram of ßL-14 Lys_{73} --> Arg. Gradient was 10 --> 12% acetonitrile / 0.1 M TEAoAc in 20 minutes.

APPENDIX II

SITE SATURATION STUDIES OF ß-LACTAMASE: PRODUCTION AND CHARACTERIZATION OF MUTANT ß-LACTAMASES WITH ALL POSSIBLE AMINO ACID SUBSTITUTIONS AT RESIDUE 71

Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 1588–1592, March 1986
Biochemistry

Site-saturation studies of β-lactamase: Production and characterization of mutant β-lactamases with all possible amino acid substitutions at residue 71

(mutagenesis/enzymatic catalysis/protein structure–function relationships/protein stability)

STEVE C. SCHULTZ AND JOHN H. RICHARDS

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

Communicated by Norman Davidson, October 29, 1985

ABSTRACT **A mutagenic technique that "saturates" a particular site in a protein with all possible amino acid substitutions was used to study the role of residue 71 in β-lactamase (EC 3.5.2.6). Threonine is conserved at residue 71 in all class A β-lactamases and is adjacent to the active site Ser-70. All 19 mutants of the enzyme were characterized by the penam and cephem antibiotic resistance they provided to *Escherichia coli* LS1 cells. Surprisingly, cells producing any of 14 of the mutant β-lactamases displayed appreciable resistance to ampicillin; only cells with mutants having Tyr, Trp, Asp, Lys, or Arg at residue 71 had no observable resistance to ampicillin. However, the mutants are less stable to cellular proteases than wild-type enzyme is. These results suggest that Thr-71 is not essential for binding or catalysis but is important for stability of the β-lactamase protein. An apparent change in specificity indicates that residue 71 influences the region of the protein that accommodates the side chain attached to the β-lactam ring of the substrate.**

The creation of mutant proteins with specific changes in amino acid sequence by oligonucleotide-directed mutagenesis (1, 2) affords a general method for studies of the relationship between structure and function. Recently, this approach has been applied to several enzymes, including trypsin (3), dihydrofolate reductase (4), aspartate carbamoyltransferase (5), tyrosyl-tRNA synthetase (6, 7), triosephosphate isomerase (8, 9), lysozyme (10), and β-lactamase (11, 12). Insights obtained from x-ray crystallography, computer modeling, sequence homologies, and catalytic mechanism greatly assist in predicting the consequences of a particular mutation on protein function. However, even then a particular amino acid substitution often causes a surprising change in activity; a thorough study of the role of an important residue requires substitution with all 19 other amino acids (site saturation). Such a study was recently reported (13) for residue 12 in the c-Ha-*ras* 1 gene; in this case the mutants were generated by oligonucleotide-directed mutagenesis. All 19 amino acid substitutions were also produced for residue 222 in subtilisin by a technique similar to the one used here (14).

As an approach to structure–function studies, site saturation becomes particularly attractive under two conditions: (*i*) a procedure exists for efficiently generating the appropriate mutants and (*ii*) the protein in question affects the phenotype of cells such that one can easily screen for mutants that perform a particular function. One can extend this approach to the simultaneous saturation of two or more sites in a protein to assess the effect of combinations of residues on protein function or, perhaps, to create a novel activity.

Site saturation can be accomplished by introducing into the gene at the codon for the residue of interest a mixture of nucleotides: A, T, G, and C at the first two positions, G and C at the third position. This mixture of oligonucleotides, which codes for all 20 amino acids and the amber codon, can be readily introduced into the gene as a "cassette" when unique restriction sites are conveniently nearby. (If desired, appropriate mixtures of oligonucleotides can be used to encode a particular subset of amino acids.)

We have applied site saturation to the study of pBR322-encoded β-lactamase (EC 3.5.2.6). This is an RTEM-1 enzyme (12, 15) originally taken from R factor R1 (16, 17) that catalyzes the hydrolysis of the β-lactam ring of penam and cephem antibiotics (18). This activity confers resistance to these antibiotics on cells that produce the enzyme and thereby provides a convenient screening procedure to assess the activities of mutants. The class A β-lactamases contain a conserved triad (Ser-Thr-Xaa-Lys) (19) at the catalytic site; this site includes Ser-70, whose hydroxyl group opens the β-lactam ring (20, 21). This study focuses on the conserved Thr-71, whose role in the activity of β-lactamase is unknown.

Restriction sites that flank the codon for Thr-71 were introduced into pBR322 by oligonucleotide-directed mutagenesis; an *Ava* I site at nucleotide 3972 and a *Sca* I site at nucleotide 3937 were introduced. The DNA fragment between nucleotides 3972 and 3937 (Pro-62 to Val-74 in β-lactamase) was removed and replaced with a mixture of synthetic double-stranded oligonucleotides that included 32 codons for residue 71 (codons for 20 amino acids and the amber codon). The resulting mixture of plasmids was used to transform *Escherichia coli* and the colonies were screened for their sensitivity to penam and cephem antibiotics. Surprisingly, cells containing any of 14 of the mutant β-lactamases displayed appreciable resistance to ampicillin; only cells containing mutants with Tyr, Trp, Asp, Lys, or Arg at residue 71 had no observable resistance to ampicillin. The mutant proteins show significantly increased sensitivity to proteolysis.

MATERIALS AND METHODS

Restriction enzymes and the large (Klenow) fragment of DNA polymerase I were purchased from Boehringer Mannheim. The T4 DNA ligase and T4 DNA kinase were obtained from Bethesda Research Laboratories. All antibiotics were from Sigma. The [α-^{32}P]dTTP, 3000 μCi/mmol (1 Ci = 37 GBq), was purchased from Amersham and [γ-^{32}P]ATP was from ICN.

Oligonucleotides were synthesized by using the phosphoramidite chemistry (22) on the Applied Biosystems (Foster City, CA) DNA synthesizer, model 380A. Mixed oligonucleotides were produced by using an equimolar mixture of the four nucleotide phosphoramidites or an equimolar mixture of G and C nucleotide phosphoramidites in a normal

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

coupling reaction. The oligonucleotides were purified by preparative polyacrylamide gel electrophoresis.

Bacteriophage M13 mp 8 was grown in *E. coli* JM 103 (23) in 2× YT medium and prepared according to standard procedures (24). Plasmid pBR322 was grown in *E. coli* LS1 (25) in LB medium (28). Plasmid purification (26) and cell transformations (27) were done by using standard procedures. Selection involved growth in the presence of tetracycline at 15 mg/liter or ampicillin at 100 mg/liter unless otherwise stated. Numbering of pBR322 nucleotides is the one commonly used (28).

Oligonucleotide-directed mutagenesis was carried out on M13 mp 8 containing the 752-base-pair *Pst* I–*Eco*RI fragment of pBR322. An *Ava* I site was introduced by changing T to C at residue 3972 and a *Sca* I site was introduced by changing A to T at residue 3939. Procedures used for the mutagenesis are described elsewhere (29, 30). To remove the preexisting *Sca* I site at 3846 on pBR322, the 61-base-pair segment between *Hin*cII at 3907 and *Sca* I at 3846 was removed and replaced by a 61-base-pair fragment that changed G to A at 3846.

To introduce the 20 amino acids at site 71, the following mixture of oligonucleotides was used.

$$5'\text{-CC-GAG-GAA-CGT-TTT-CCA-ATG-ATG-AGC}\begin{bmatrix}\text{AA}\\ \text{CCC}\\ \text{GGG}\\ \text{TT}\end{bmatrix}\overset{*}{\text{TTC}}\text{-AAA-GT}$$
$$\text{C-CTT-GCA-AAA-GGT-TAC-TAC-TCG}\begin{bmatrix}\text{AA}\\ \text{CCC}\\ \text{GGG}\\ \text{TT}\end{bmatrix}\text{AAG-TTT-CA-5'}$$

The asterisk indicates a silent mutation that removes a *Dra* I restriction site at 3943 and also decreases the self-complementarity of 6 (and possibly as many as 10) bases near this end of the individual strands. After purification, as described above, the mixtures of oligonucleotides were individually phosphorylated. They were then annealed by mixing 0.4 pmol/μl of each strand and heating to 95°C in 10 mM $MgCl_2$/50 mM Tris·HCl, pH 8, followed by gradual cooling to 20°C over a 45-min period. The resulting mixture of double-stranded oligonucleotides (0.4 pmol) was mixed with the *Sal* I to *Sca* I fragment (0.04 pmol) and the *Ava* I to *Sal* I fragment (0.04 pmol) from pBR322-CR7 and pBR322-XN1, respectively (see Fig. 1). This mixture was ligated in 10 mM $MgCl_2$/50 mM Tris·HCl, pH 8/0.5 mM ATP/5 mM dithiothreitol containing 10 units of T4 DNA ligase in a total volume of 50 μl at 15°C for approximately 18 hr. The reaction mixture was extracted with phenol and the DNA was precipitated by addition of ethanol. The DNA was redissolved in 20 μl of 1 mM Tris·HCl, pH 8/0.1 mM EDTA; an aliquot (5 μl) was used for transformation of *E. coli* LS1 cells. After the transformation, 1/10th of the cells were plated on tetracycline, producing 423 colonies. From these, 108 individual colonies were picked and inoculated onto tetracycline-containing plates and were also tested for activity on the various substrates of β-lactamase. Plasmids derived from these 108 colonies were sequenced.

For sequencing, the plasmids were digested with *Ava* I and the resulting 2962-base-pair fragment was isolated from a 1.2% agarose gel by using DEAE-paper or the International Biotechnologies (New Haven, CT) UEA electroelutor. The fragment was labeled at T-3973 by using [α-^{32}P]dTTP and the Klenow fragment of DNA polymerase I. The labeled fragments were sequenced by using standard techniques (31).

In L agar plates (28), an antibiotic concentration gradient was generated by raising one end of a Petri dish (8.5-cm diameter) 5 mm and pouring 15 ml of L agar containing an appropriate concentration of antibiotic into the tilted dish. These were allowed to dry overnight. The plates were then placed on a flat horizontal surface and 15 ml of L agar was poured on top; they were used immediately upon hardening. An aliquot (50 μl) of a 1:10^5 dilution of a saturated culture (approximately 2 × 10^9 cells per ml) was spread over half of the plate. A standard mutant was plated on the other half; standards used were as follows: for ampicillin and benzylpenicillin at 37°C and for 6-aminopenicillanic acid at 30°C, Thr-71 → Leu; for 6-aminopenicillanic acid at 37°C, Thr-71 → Cys; for ampicillin and benzylpenicillin at 30°C, Thr-71 → Glu.

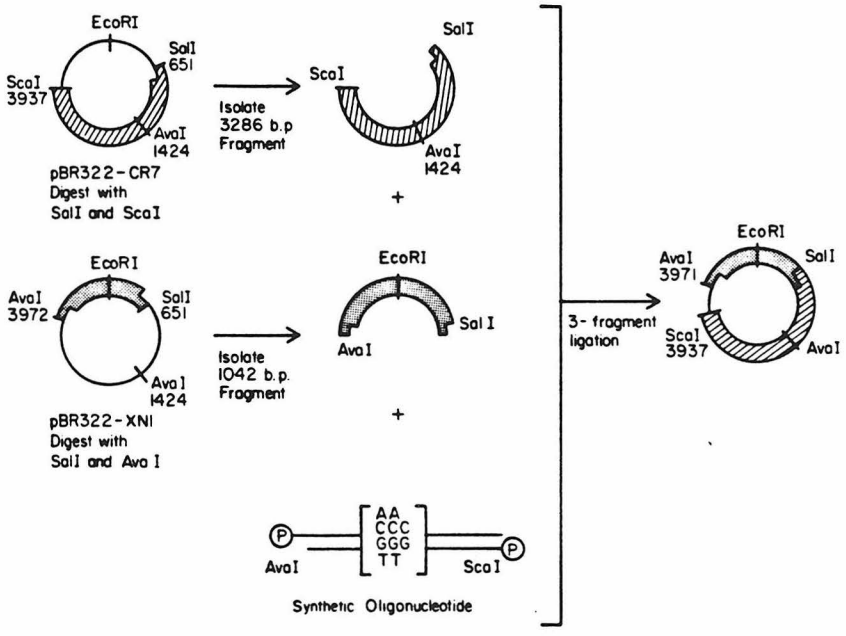


FIG. 1. Design of three-fragment ligation for inserting the mixture of oligonucleotides. bp, Base pair.

Protein for electrophoretic transfer blots was prepared by pelleting 1×10^9 *E. coli* LS1 cells from a late-logarithmic-phase culture containing tetracycline at 15 mg/liter in LB medium. These cells were resuspended in 100 μl of 10% (vol/vol) glycerol/5% (vol/vol) 2-mercaptoethanol/3% sodium dodecyl sulfate/62.5 mM Tris·HCl, pH 6.8, and incubated 10 min at 95°C. An aliquot (20 μl) was loaded onto a 15-cm 12% polyacrylamide gel with a 2-cm 4% stacking gel. After electrophoresis, protein blots were prepared according to the Bio-Rad procedure (32); the transfer was at 50 V for 5 hr. Visualization was according to the Vectastain procedure (33) using rabbit antibody to β-lactamase.

RESULTS

Introduction of the *Ava* I 3972 and *Sca* I 3937 sites and removal of the *Sca* I 3845 site were verified by restriction mapping of the plasmids. The absence of other changes in the β-lactamase gene was verified by sequencing one of the mutants from the *Pst* I site through the residues of the structural gene encoding the protein N terminus.

After ligation of the mixture of oligonucleotides into the plasmid, transformation of competent *E. coli* LS-1 cells gave 423 colonies that were resistant to tetracycline. From these 423 colonies, 108 were picked and inoculated onto tetracycline plates and were also tested for resistance to ampicillin (10, 50, 100, and 500 mg/liter), benzylpenicillin (100 mg/liter), 6-aminopenicillanic acid (10, 25, 50, and 100 mg/liter), cephalothin (25, 50, and 100 mg/liter), and cephalexin (25 and 50 mg/liter). Ampicillin and cephalothin plates were incubated at 30°C and 37°C; all others were incubated only at 37°C. This phenotype screening indicated that a change in antibiotic specificity and reduced stability, particularly at 37°C, were characteristics of some of the mutants.

Plasmids derived from 105 of the colonies were sequenced from *Ava* I (3972) through *Sca* I (3937); Table 1 lists the codon frequencies in this collection of colonies. The sequences were matched to the phenotypes determined above. Mutants with Tyr, Trp, Asp, Lys, or Arg at position 71 gave no resistance to any of the five antibiotics. A mutant with Phe was resistant to ampicillin at 100 mg/liter at 30°C but sensitive even to low levels of ampicillin at 37°C. Mutants with Gly, Gln, or Glu produced resistance up to 100 mg of ampicillin or benzylpenicillin per liter at 37°C. Mutants with Ala, Val, Leu, Ile, Pro, His, Cys, Ser, Thr, or Asn gave resistance to ampicillin at >500 mg/liter and benzylpenicillin at >100 mg/liter. Only mutants with His, Cys, Ser, or Thr were resistant to 6-aminopenicillanic acid at >100 mg/liter, those with Gly, Ala, Val, Leu, Ile, Pro, Glu, or Asn were resistant to 25 mg/liter, and those with Met or Gln (although resistant to ampicillin) showed no resistance to 6-amino-penicillanic acid. Only the wild-type enzyme (Thr-71) conferred resistance to cephalothin and cephalexin at 37°C. A mutant, Thr-71 → Ser, was resistant to cephalothin at 50 mg/liter at 30°C, but none of the other mutants grew in the presence of this antibiotic even at 30°C.

Table 1. Distribution of codons for residue 71

Codon	No.	Codon	No.	Codon	No.	Codon	No.
AAC	2	CAC	3	GAC	2	TAC	8
AAG	6	CAG	3	GAG	3	TAG	2
ACC	4	CCC	1	GCC	5	TCC	5
ACG	2	CCG	2	GCG	4	TCG	3
AGC	3	CGC	0	GGC	2	TGC	4
AGG	4	CGG	2	GGG	1	TGG	5
ATC	6	CTC	1	GTC	0	TTC	1
ATG	7	CTG	2	GTG	2	TTG	6

In addition to these 101 point mutations, there were three deletions at the *Sca* I site, six insertions at the *Ava* I site, one deletion at the *Ava* I site, four insertions in the codon for residue 71, and one insertion in the synthetic segment.

More accurate values for the level of antibiotic resistance provided to cells by the mutant proteins were determined by spreading cells onto an agar plate with a continuous concentration gradient of antibiotic. The maximal concentration at which colonies were established could then be observed. These phenotypic characteristics are summarized in Table 2. The values were independent of dilutions of cells higher by a factor of 2 or lower by a factor of 10 than those used in the experiments summarized in Table 2. Particularly noteworthy are the resistances at 37°C of cells producing the mutants Thr-71 → Ile and Thr-71 → His. In addition, the differences in levels of resistance at 30°C versus those at 37°C are especially dramatic.

Protein boiled in sodium dodecyl sulfate after extraction from cultures of the various mutants grown at 30°C and 37°C was electrophoresed in denaturing sodium dodecyl sulfate/polyacrylamide gels and then transferred to nitrocellulose. Fig. 2 shows antibody stains of these blots. Protein of the same size as β-lactamase is present for all mutants, but the quantity varies substantially for different mutants; this variation is greatly reduced at 30°C. We believe that the protein responsible for the band slightly above β-lactamase is pre-β-lactamase because it represents a protein of appropriate size and because this band is absent for the Thr-71 → amber mutant. The blots suggest that the mutant proteins are processed normally but that mature mutant proteins are degraded, probably by *E. coli* proteases.

DISCUSSION

Importantly, for site-saturation studies, one should use mixtures of synthetic oligonucleotides having all specified codons at nearly equal frequencies; if codons for certain amino acids are present at low frequencies, the corresponding mutants will be underrepresented, difficult to find, and

Table 2. Maximal level of resistance of strains with mutations at β-lactamase residue 71

Amino acid	Maximal antibiotic concentration (mg/liter) at which colonies grew					
	Ampicillin		Benzyl-penicillin		6-Amino-penicillin	
	30°C	37°C	30°C	37°C	30°C	37°C
Gly	>500	Trace	>500	50	35	20
Ala	>500	75	>500	100	Trace	Trace
Val	>500	150	>500	175	75	Trace
Leu	>500	100	>500	125	80	Trace
Ile	>500	>500	>500	>500	90	40
Met	>500	125	>500	150	30	0
Pro	>500	200	>500	350	50	40
Phe	60	0	0	0	0	0
Trp	0	0	0	0	0	0
Tyr	0	0	0	0	0	0
His	>500	>500	>500	>500	>250	>250
Cys	>500	>500	>500	>500	>250	100
Ser	>500	>500	>500	>500	>250	>250
Thr	>500	>500	>500	>500	>250	>250
Asn	270	Trace	210	Trace	55	Trace
Gln	130	20	70	Trace	55	0
Asp	0	0	0	0	0	0
Glu	200	Trace	70	Trace	Trace	Trace
Lys	0	0	0	0	0	0
Arg	0	0	0	0	0	0
Stop	0	0	0	0	0	0

A > indicates that this value was the highest level of antibiotic tested.

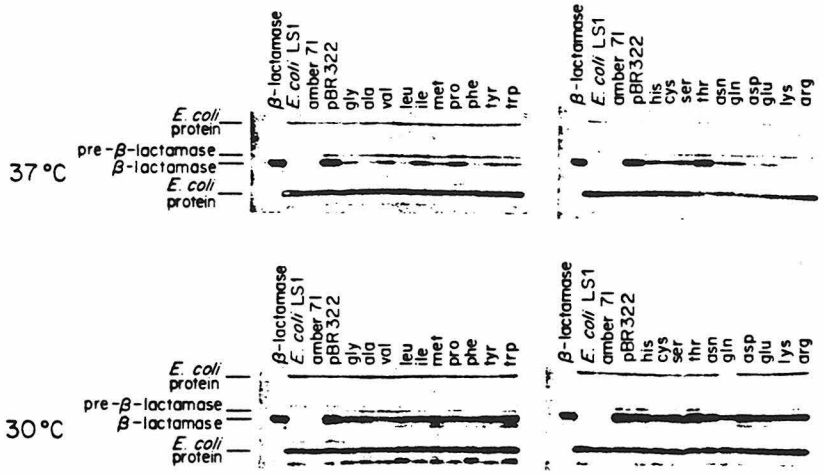


FIG. 2. Blots of β-lactamases with the 19 amino acid substitutions at residue 71. The blots were treated with rabbit anti-β-lactamase antibody and visualized by using the horseradish peroxidase assay (see ref. 33).

may go untested in the phenotype screening, particularly when multiple sites are involved. Nearly equimolar mixtures of oligonucleotides can be obtained by careful synthesis (S. Horvath, personal communication) and then used in cassette mutagenesis to generate the various mutations with known frequencies. In a Poisson distribution the probability of finding a particular codon at least once among 101 colonies is 96%. As shown in Table 1, 94% (30/32) of the codons were observed one or more times. The absence of two codons, therefore, was probably the consequence of a small sample size.

In sequencing plasmids from 105 individual colonies we found none in which the 35-base-pair segment contained unexpected base substitutions. However, 14 frameshift mutants resulting from single-base deletions or additions were observed (see Table 2). The 9 additions and deletions at restriction sites may occur during processing of the synthetic fragment or during ligation; similar deletions have been observed, for example, in ligation of a synthetic fragment into the *Pst* I site of the gene for subtilisin (14). The other 5 mutations that occur within the synthetic fragment probably result from imperfections in the DNA synthesis. None of the colonies containing plasmids with these frameshift mutations show any resistance to penam or cephem antibiotics.

The antibiotic resistances of colonies containing the 19 mutant β-lactamases were initially determined by picking colonies and patching them onto agar plates with various levels of antibiotics. More accurate values were measured by spreading mutant colonies onto agar plates having a linear concentration gradient of antibiotics (see Table 2). This latter method gives consistently lower values for antibiotic resistance, as in this case individual cells must establish colonies, whereas when colonies are picked and applied in patches many cells are initially present at a single site, allowing them to cooperate in inactivating sufficient antibiotic to allow growth. Both sets of data are, however, useful, as colonies with mutants having low lactamase activity would be classified as inactive if only plating were used. In Table 2, colonies that grew when picked, but not when plated, are designated as having trace activity.

The antibiotic resistance of colonies containing mutant plasmids reflects many factors: plasmid replication, transcription, stability and ease of translation of the mutant mRNAs, stability of the pre-β-lactamases in the cytoplasm, possible differences in the rate of processing and secretion, stability of the proteins in the periplasm, and, finally, intrinsic enzymatic activity. With regard to the steps from transcription to translation, the codon used for a given amino acid at residue 71 did not alter observed antibiotic resistance; in all cases, a given amino acid mutant confers the same level of antibiotic resistance regardless of which codon was used (see Table 1).

Mutations may affect processing and secretion of proteins (34) and thereby alter the relative concentrations of mutant β-lactamases in the periplasm. Although mutants such as Thr-71 → Ser and Ser-70 → Thr and the double mutant Ser-70 → Thr, Thr-71 → Ser seem to be processed and secreted normally in *Salmonella typhimurium* (12), these mutations are unlikely to affect these processes as drastically as mutations such as Thr-71 → Trp, Tyr, Lys; even in these cases, however, abundant processed protein is apparent in cells growing at 30°C (see Fig. 2).

Resistance to thermal denaturation and proteolysis can greatly influence the *in vivo* concentration of mutant proteins (12, 35). Analysis by protein blots (Fig. 2) for cell lysates of colonies growing at 30°C and 37°C show significant differences in the concentrations of mutants. A band representing the precursor protein shows similar intensities for the mutants and wild type; the presence of normal amounts of precursor proteins suggests that processing and secretion have not been drastically altered for the mutants. However, the amounts of protein in the periplasm vary greatly for the mutants as contrasted to wild type, particularly for cells growing at 37°C, indicating that the mutants are more susceptible to proteolysis than the wild-type enzyme.

A surprising number of mutants with amino acid substitutions at Thr-71 display lactamase activity; indeed, of the 19 substitutions at residue 71 only 5 (Trp, Tyr, Lys, Arg, or Asp) give proteins that confer no observable resistance to ampicillin or benzylpenicillin. Enzymes with Gln, Glu, Asn, or Phe at residue 71 provide low-level resistance, but enzymes with all other amino acids at this site have high activity toward benzylpenicillin and ampicillin. Some preliminary conclusions about structure–function relationships can be drawn from these data. Mutants having amino acids with charged or very large side chains at residue 71 have at best very low activity, otherwise both polar and nonpolar side chains seem compatible with appreciable activity. When 6-aminopenicillanic acid is used to select for activity, mutants with amino acids having nonpolar side chains show decreased activity, whereas mutants with amino acids having polar side chains retain activity toward this antibiotic. All mutants show a dramatically decreased ability to confer resistance toward the

cephem antibiotics cephalothin and cephalexin. At 30°C, wild-type β-lactamase (Thr-71) confers resistance to at least 100 mg/liter while the mutant Thr-71 → Ser confers resistance only to low levels of the cephem antibiotics; at 37°C only wild-type enzyme provides observable resistance. Though we have yet to characterize fully the inherent catalytic activity of all the mutants at residue 71, some representative data for benzylpenicillin as substrate are, respectively, k_{cat} (sec^{-1}), K_m (μM), k_{cat}/K_m ($M^{-1} \cdot sec^{-1}$), and relative values for k_{cat}/K_m as follows: wild-type enzyme, 2000, 26, 7.7×10^7, and 1; Thr-71 → Ser, 300, 21, 1.6×10^7, and 0.21 (12); Thr-71 → Ile, 1530, 350, 4.4×10^7, and 0.57 (Y. H. Chang and J.H.R., unpublished results).

Because many mutants at residue 71, including those having amino acids with polar and nonpolar side chains, show appreciable activity toward penams, we conclude that Thr at 71 is not essential for binding and catalysis, though it may perform some contributing role in these functions. However, residue 71 does play a very important role in the structural stability of the protein, as observed in the increased sensitivity to proteolysis of all 19 mutants. The importance of branching at the β-carbon of Thr with a polar hydroxyl group and nonpolar methyl group is demonstrated, for example, by the increasing antibiotic resistance of cells growing at 37°C provided by the series of enzymes with Gly < Ala < Leu < Val. The increased thermal and proteolytic sensitivity relative to wild type of the mutant Thr-71 → Ser (12), lacking only the methyl group of Thr, further supports this view. The relatively high catalytic activities of Thr-71 → Cys and Thr-71 → Ser demonstrate the importance of the polar group. The reduced activity toward 6-aminopenicillanic acid of enzymes having amino acids at residue 71 with nonpolar side chains suggests that the side chains of residue 71 influence the hydrophobic or hydrophilic characteristics of the pocket that accommodates the substituent attached to the β-lactam ring of the substrates. This possibility is further emphasized by the lower activity of enzymes with Glu or Gln at residue 71 toward benzylpenicillin (neutral, nonpolar substituent) than toward ampicillin (positively charged, polar substituent).

All mutant enzymes confer on cells at best very low resistance toward cephalothin and cephalexin; this likely results from the inherently lower activity of β-lactamase against cephem antibiotics (36) and accords with results observed for mutations in other regions of the enzyme (37). This behavior also demonstrates the importance of the conserved residues at the catalytic site in providing an enzyme with high levels of catalytic activity and stability.

We thank Suzanne Horvath and Marilyn Tomich for synthesis of the oligonucleotides. We thank Gloria Dalbadie-McFarland for her help on the oligonucleotide-directed mutagenesis, Mingjim Wu for his laboratory assistance, and James Neitzel for providing the anti-β-lactamase antibody. We also thank Gloria Dalbadie-McFarland, James Neitzel, Steven Carroll, Yie-Hwa Chang, and William Healey for many helpful discussions and suggestions. This work was supported by National Institutes of Health Grant GM 16424. This is contribution no. 7282 from the Division of Chemistry and Chemical Engineering, California Institute of Technology.

1. Dalbadie-McFarland, G. & Richards, J. H. (1983) *Annu. Rep. Med. Chem.* **18**, 237–245.
2. Ackers, G. K. & Smith, F. R. (1985) *Annu. Rev. Biochem.* **54**, 597–629.
3. Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R. & Rutter, W. J. (1985) *Science* **228**, 291–297.
4. Villafranca, J. E., Howell, E. E., Voet, D. H., Strobel, M. S., Ogden, R. C., Abelson, J. N. & Kraut, J. (1983) *Science* **222**, 782–788.
5. Robey, E. A. & Schachman, H. K. (1984) *J. Biol. Chem.* **259**, 11180–11183.
6. Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M. & Smith, M. (1982) *Nature (London)* **299**, 756–758.
7. Wilkinson, A. J., Fersht, A. R., Blow, D. M., Carter, P. & Winter, G. (1984) *Nature (London)* **307**, 187–188.
8. Straus, D., Raines, R., Kawashima, E., Knowles, J. R. & Gilbert, W. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2272–2276.
9. Petsko, G. A., Davenport, R. C., Jr., Frankel, D. & Raibhandary, U. L. (1984) *Biochem. Soc. Trans.* **12**, 229–232.
10. Perry, L. J. & Wetzel, R. (1984) *Science* **226**, 555–557.
11. Dalbadie-McFarland, G., Cohen, L. W., Riggs, A. D., Morin, C., Itakura, K. & Richards, J. H. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6409–6413.
12. Dalbadie-McFarland, G., Neitzel, J. & Richards, J. H. (1986) *Biochemistry*, in press.
13. Seeburg, P. H., Colby, W. W., Capon, D. J., Goeddel, D. V. & Levinson, A. D. (1984) *Nature (London)* **312**, 71–75.
14. Wells, J. A., Vasser, M. & Powers, D. B. (1985) *Gene* **34**, 315–323.
15. Matthew, M. & Hedges, R. W. (1976) *J. Bacteriol.* **125**, 713–718.
16. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene* **2**, 95–113.
17. So, M., Gill, R. & Falkow, S. (1975) *Mol. Gen. Genet.* **142**, 239–249.
18. Datta, N. & Kontomichalou, P. (1965) *Nature (London)* **208**, 239–241.
19. Ambler, R. P. (1980) *Philos. Trans. R. Soc. London Ser. B* **289**, 321–331.
20. Fisher, J., Belasco, J. G., Khosla, S. & Knowles, J. R. (1980) *Biochemistry* **19**, 2895–2901.
21. Knott-Hunziker, V., Waley, S. G., Orlek, B. S. & Sammes, P. G. (1979) *FEBS Lett.* **99**, 59–61.
22. Beaucage, S. L. & Caruthers, M. H. (1981) *Tetrahedron Lett.* **22**, 1859–1862.
23. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–321.
24. Messing, J. (1983) *Methods Enzymol.* **101**, 20–89.
25. Boyer, H. W. & Roulland-Dussoix, D. (1969) *J. Mol. Biol.* **41**, 459–472.
26. Ish-Horowicz, D. & Burke, J. F. (1981) *Nucleic Acids Res.* **9**, 2989–2998.
27. Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557–580.
28. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
29. Norris, K., Norris, F., Christiansen, L. & Fiil, N. (1983) *Nucleic Acids Res.* **11**, 5103–5112.
30. Zoller, M. J. & Smith, M. (1983) *Methods Enzymol.* **100**, 468–500.
31. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
32. Anonymous (1984) *Trans-Blot Cell Operating Instructions* (Bio-Rad Laboratories, Richmond, CA).
33. Anonymous (1984) *Immunodetection of Antigens on Nitrocellulose Using Vectastain ABC Kit* (Vector Laboratories, Burlingame, CA).
34. Kadonaga, J. T., Gautier, A. E., Straus, D. R., Charles, A. D., Edge, M. D. & Knowles, J. R. (1984) *J. Biol. Chem.* **259**, 2149–2154.
35. Hecht, M. H., Sturtevant, J. M. & Sauer, R. T. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5685–5689.
36. Sykes, R. B. & Smith, J. T. (1979) *Beta-Lactamases* (Academic, New York), pp. 369–401.
37. Barany, F. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4202–4206.

PROPOSITION I

SITE SATURATION STUDIES OF THE DNA BINDING HELIX OF CAP AND ITS OPERATOR SEQUENCE

Background

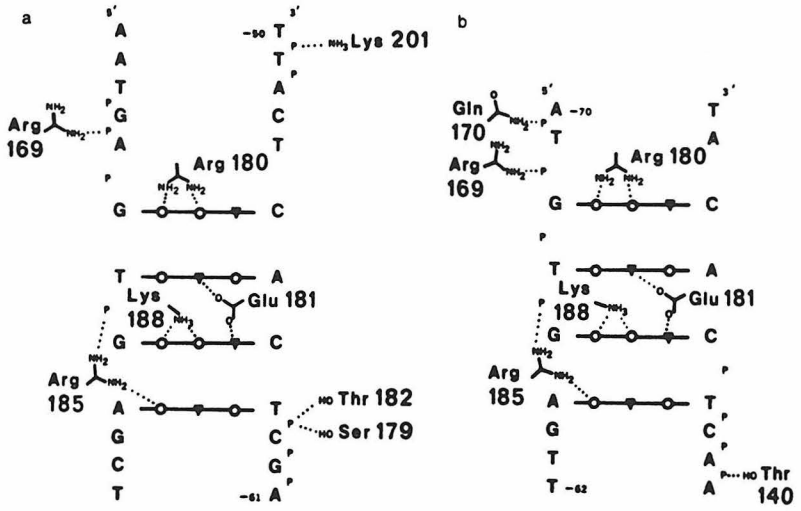
By interacting with specific base sequences in the DNA helix, proteins can bind to the operator regions of genes and exert the positive and/or negative control necessary for proper gene expression in cells. Recently, x-ray crystallography has provided the three-dimensional structures for several DNA binding proteins including: the catabolite activator protein (1), the lambda cro repressor (2), a proteolytic fragment of the lambda CI repressor (3), a proteolytic fragment of the phage 434 repressor (4), and the tryptophan repressor (5). From among the very dissimilar structures of these proteins, a common feature has emerged that may explain how proteins can recognize and bind to specific base sequences in the DNA helix. Each identical subunit of these dimeric proteins contains a helix-turn-helix structure in which one of these α-helices protrudes from the surface of each subunit. In the protein, these protruding helices are 34 A apart and could align themselves with two successive major grooves on one face of the DNA double helix. Amino acid side chains extending from these helices could interact with specific arrangements of chemical groups in the major groove of the DNA helix and thereby provide the specificity these proteins require for binding to their operator sequences. This model has recently been verified by a solution for the crystal structure of phage 434 repressor bound to its operator sequence (4).

Due to the helical twist of DNA, the helix-turn-helix

structure will allow only three to five amino acid side chains to contact no more than three to five base pairs in the major groove of the DNA helix. Although limited to three to five sets of interactions for each subunit, DNA binding proteins can distinguish between very similar sequences and bind very tightly to the proper sequence. Because such functions must be accomplished by these limited interactions and because only four nucleotide bases and twenty amino acids occur naturally, one may expect to find a "code" for the interactions of proteins and DNA; this "code" would define what amino acid side chains could bind to which specific base pairs. Several possible bonding schemes have been proposed by which certain amino acid side chains could contact particular bases in the major groove of the DNA helix (6). For example, a glutamine or asparagine could simultaneously form hydrogen bonds with the N6 and N7 positions of adenine or interact with an adenine N6 or cytosine N4 while at the same time bind a thymine O4 or purine N7. Discrete bonding schemes have also been proposed for serine, arginine, and lysine (6).

The catabolite activator protein (CAP) binds to operator regions of more than 20 E. coli genes when cyclic-AMP is present and stimulates their transcription (7). CAP can also repress transcription of some genes (8,9). The structure of the CAP-cAMP complex has been solved at 2.9 A resolution (1). Specific interactions have been proposed for binding of amino acid side chains of the F helix with specific base

groups in the major groove of the B-DNA helix (10). These interactions are shown below (from reference 10).



Studies of mutant CAP proteins and altered CAP binding sites support this model (11). For example, changing base pairs 57 or 66 from GC and CG to AT and TA respectively requires a compensating change of Glu 181 to Lys, Leu, or Val in CAP (11).

I propose to use site saturation (12) of CAP and its operator sequence to investigate the recognition code for the binding of this protein to specific base sequences.

Proposal

Effective site saturation studies require a selectable phenotype such that active mutants can be isolated from among large numbers of inactive mutants. ß-Lactamase exhibits this type of selectable phenotype; the presence of active ß-lactamase provides resistance to penam and cephem antibiotics such that cells producing this enzyme can grow in the presence of

these antibiotics while all others perish (13). This activity allows one to isolate cells producing active mutants of this enzyme even when present only once among 10^6 or more inactive mutants.

For the experiments proposed here, the ß-lactamase gene will be placed under control of the lac promoter, which requires CAP to initiate transcription (14), such that this antibiotic selection can be used to isolate mutants of CAP that retain their ability to activate transcription. A synthetic insert containing the lac operon will be introduced into the Ssp I restriction site that immediately precedes the ß-lactamase gene in pBR322. The lac operon is 123 base pairs in length (15), but 17 base pairs at the end of this segment will be omitted to maintain proper spacing between the promoter and the beginning of the structural gene. The following oligonucleotides will be synthesized, annealed, and extended with the large (Klenow) fragment of DNA polymerase I.

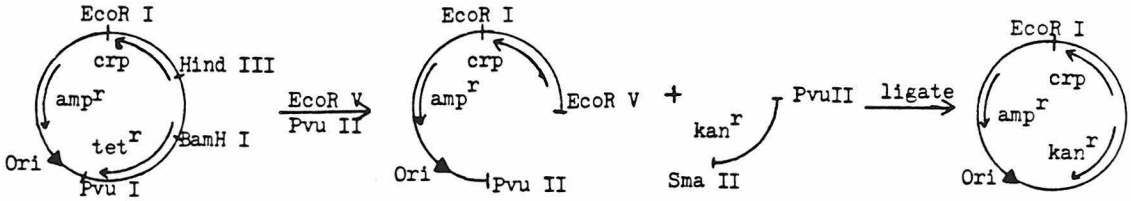
```
GCCCGAGGGCAATTAATGTGAGTTAGCTCACTCATTAGCTCGGGCAGGCTTTACAC
                                    CCGAAATGTGAAATAGCAAGGCCGAGCATACAACACACCTTAACACTCGCGTATTGTT
     Ava I                               Ava I
```

The resulting fragment will be ligated into the Ssp I site of pBR322. This construction will probably interfere with the ß-lactamase promoter already present - preventing the normal constitutive transcription (16); if not, this promoter region can be removed.

If this construction is successful, E. coli cells con-

taining the plasmid should grow in the presence of ampicillin only when cAMP is added to the media. The promoter may be slightly "leaky" such that some ß-lactamase is produced; if low levels of resistance are observed, one could use either higher levels of ampicillin or more potent antibiotics, such as cephalexin. Note that two different Ava I sites have been included in the synthetic fragment; these sites will facilitate mutagenesis of the operator sequence.

The gene for CAP (called the crp gene) has been cloned and expressed in pBR322 (17). The resulting plasmid (called pHA7) confers ampicillin but not tetracycline resistance to cells harboring the plasmid. Since the ampicillin gene will be placed under control of the lac operon, a constitutive selectable marker must be introduced into the plasmid. The following construction is proposed:



The portion of the gene coding for the DNA binding domain of CAP has been synthesized to contain several convenient restriction sites (T. Steitz, unpublished); this portion of the CAP gene can be easily incorporated into the cloned, expressed gene. By combining this plasmid with plasmid pCKP described above, all the necessary elements for site saturation studies of both CAP and the operator sequence will be present in a single plasmid called pCRS.

The first experiment will be site saturation of Arg 185. By finding a mutant that continues to initiate transcription, we can eliminate residue 185 and base pairs -58 and -65 from future studies. We could then concentrate on saturating the three amino acid residues and three base pairs involved in the sequence specific binding of CAP. If Arg 185 cannot be replaced by any other amino acids, this residue will be left unchanged.

I propose to simultaneously saturate residues 180, 181, and 188 from CAP and base pairs -55, -56, -57, -66, -67 and -68 from the operator DNA.

The following mixture of sequences will be incorporated into Mlu I and Xho I sites in the CAP gene:

CGCGTCAGGAAATTGGTCAGATTGTCGGCTTGTTCTXXXXXXACCGTGGGACGCATTCTGXXXATGC
AGTCCTTTAACCAGTCTAACAGCCGAACAAGAXXXXXXTGGCACCCTGCGTAAGACXXXTACGAGCT

The xxx sites will consist of codons for all twenty amino acids. Phosphoramidite derivatives of twenty nucleotide trimers will be used to incorporate these sequences in a single synthesis. This mixture will include 20^3 or 8000 possible amino acid sequences.

Since 8000 possible amino acid sequences are present, we must limit the number of operator sequences to keep this system workable. Therefore, 64 different oligonucleotides must be synthesized separately. The sequences will be:

GCACTGCCCGAGGCAATTAATxxxAGTCGACT

These 64 sequences will contain different base sequences at the site marked xxx. Notice that the 3' ends of this sequence are self-complementary. For each individual sequence, the strands will be annealed and extended using the Klenow fragment of DNA polymerase I. This procedure will make all recognition sequences palindromic. By cleaving this fragment with Ava I restriction endonuclease, protruding 5' ends will be produced. All 64 individual fragments can now be combined in equal amounts and ligated into a mixture of plasmids containing all possible amino acid sequences at residues 180, 181, and 188. The resulting plasmid mixture will contain 512,000 unique sequences.

This mixture of plasmids will be used to transform E. coli and the resulting cells will be grown in the presence of cAMP and ampicillin. CAP mutants that can still initiate transcription will be resistant to ampicillin. The mutations in the protein and operator regions of plasmids from amp^r colonies can then be identified by DNA sequencing. Since 512,000 possible combinations exist, approximately 5×10^7 kanamycin resistant cells (kan^r is the constitutive selectable marker on pCRS) should be selected for ampicillin resistance; most of these cells should be ampicillin sensitive.

Several problems may exist in these experiments. For example, particular CAP mutants may bind at sites present in the cell genome and, therefore, could be lethal to these cells. Similarily, recognition of specific altered operator

sequences by regulatory proteins already present in the cell could interfere with the results of the experiment. Also, some CAP mutants may bind to the operator region as desired, but fail to activate transcription such that these cells will be ampicillin sensitive.

Even though all possible combinations of CAP and operator sequences may not be tested because of the problems listed above, emphasis should be placed on the altered protein and operator sequences that are obtained from these experiments. Even if only a few sets of interactions are found, the implications will be very exciting.

References

1. McKay, D.B. & Steitz, T.A. (1981) Nature (London) 290, 744-749.

2. Anderson, W.F.,Ohlendorf, D.H., Takeda, Y. & Matthews, B.W. (1981) Nature (London) 290, 754-758.

3. Pabo, C.O. & Lewis, M. (1982) Nature (London) 298, 443-447.

4. Anderson, J.E., Patshne, M. & Harrison, S.C. (1985) Nature (London) 316, 596-601.

5. Shevitz, R.W., Otwinowski, Z., Joachimiak, A., Lawson, C.L. & Sigler, P.B. (1985) Nature (London) 317, 782-786.

6. Pabo, C.O., Jordan, S.R. & Frankel, A.D. (1983) J. Biomol. Sterodyn. 1, 1039-1049.

7. Pastan, I. & Adhya S. (1976) Bacteriol. Rev. 40, 527-551.

8. Aiba, H. (1983) Cell 32, 141-149.

9. Movva, R.N., Green, P., Nakamura, K. & Inouye, M. (1981) FEBS Lett. 128, 186-190.

10. Weber, I.T. & Steitz, T.A. (1984) Proc. Natl. Acad. Sci. USA 81, 3973-3977

11. Ebright, R.H., Cossart, P., Gicquel-Sanzey, B. & Beckwith, J. (1984) Nature (London) 311, 232-235.

12. Schultz, S.C. & Richards, J.J. (1986) Proc. Natl. Acad. Sci. USA 83, 1588-1592.

13. Datta, N. & Kontomichalou, P. (1965) Nature (London) 208, 239-241.

14. de Crombrugghe, B., Chen, B., Anderson, W., Nisley, P., Gottesman, M., Pastan, I. & Perlman, R. (1971) Nat. New. Biol. 231, 139-145.

15. Dickson, R.C., Abelson, J., Barnes, W.M. & Reznikoff, W.S. (1975) Science 187, 27-35.

16. Brosius, J., Cate, R.L. & Perlmutter, A.P. (1982) J. Biol. Chem. 257, 9205-9210.

17. McKay, D.B. & Fried, M.G. (1980) J. Mol. Biol. 139, 95-96.

PROPOSITION II

EFFECT OF SER --> THR MUTATIONS ON THE THERMAL STABILITY OF AN AMINOCYCLITOL-3'-PHOSPHOTRANSFERASE

Background

Proteins exhibit a large variety of thermal stabilities; while some undergo irreversible denaturation below 50°C, others can maintain normal activity even above 90°C. Recent studies of mutant proteins containing single amino acid substitutions have shown that very minor changes in protein structure can have major effects on the stability of a protein (1). For example, the absence of a single methyl group in the Thr 71 --> Ser mutant of ß-lactamase dramatically reduces the stability of the protein toward heat, pH, and proteolysis (2). Many examples exist for destabilization of proteins by single amino acid substitutions, but only a few exist for stabilization of proteins by specific amino acid substitutions. These examples include mutations that introduce a disulfide bond (3), an ionic interaction (4,5), a hydrophobic interaction, (6) and an aromatic ring interaction (7). To stabilize a protein, very specific interactions must be generated in the three-dimensional structure of a protein; consequently, designing such mutations is difficult.

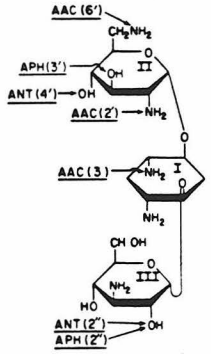
Many interactions contribute to the stability of a protein including hydrogen bonds, ionic interactions, disulfide bonds, and hydrophobic interactions. By analyzing the amino acid sequences of ferredoxins, glucose 6-phosphate dehydrogenases, and lactate dehydrogenases from several mesophilic and thermophilic bacteria, several amino acid substitutions that increase the thermal stabilities of proteins have been identified (8). The substitutions observed

most frequently were Gly --> Ala, Ser --> Ala, Ser --> Thr, Lys --> Arg, and Asp --> Glu. The aliphatic indexes of proteins from thermophilic bacteria are consistently higher than those from mesophilic bacteria (9). These results indicate that hydrophobic residues play an important role in providing thermal stability to proteins. Since we have observed that the absence of a single methyl group in the Thr 71 --> Ser mutant of ß-lactamase results in a dramatic reduction in stability (2), I propose to study the opposite possibility: stabilization of proteins by substitution of Thr for Ser.

In choosing a protein for such studies, two factors must be considered. First, temperatures ranging from 30°C to >60°C must be utilized; therefore, the protein of interest must be expressed in both mesophilic and thermophilic bacteria. Second, a selectable phenotype will be necessary for investigating large numbers of mutations. The aminocyclitol-3'-phospotransferase from the transposable element Tn5, which confers kanamycin resistance to several bacteria (10), fulfills both of these requirements. By selecting cells containing random mutants of a related protein (an aminoglycoside nucleotidyltransferase) for resistance to kanamycin in the themophile B. stearothermophilus at >60°C, two groups independently isolated mutants that were more stable than the wild type enzyme (4,5); both groups identified these mutations as Asp 80 --> Tyr and Thr 130 --> Lys.

Several different enzymes confer kanamycin resistance to

bacteria; these have been classified according to how they modify kanamycin (acetyltransferase, nucleotidyltransferase, and phosphotransferase) and where they modify kanamycin (6'-amino, 3'-hydroxy, 4'-hydroxy, 2'-amino, 3'-amino, or 2" hydroxy group) (11). The structure of kanamycin is shown below (from reference 11).



Of the three types of resistance proteins, the aminoglycoside (or aminocyclitol) 3'-phosphotransferases (APH's) are best characterized.

I propose to study the effect of Ser --> Thr mutations on the thermal stability of the aminoglycoside phosphotransferase encoded by the transposable element Tn5.

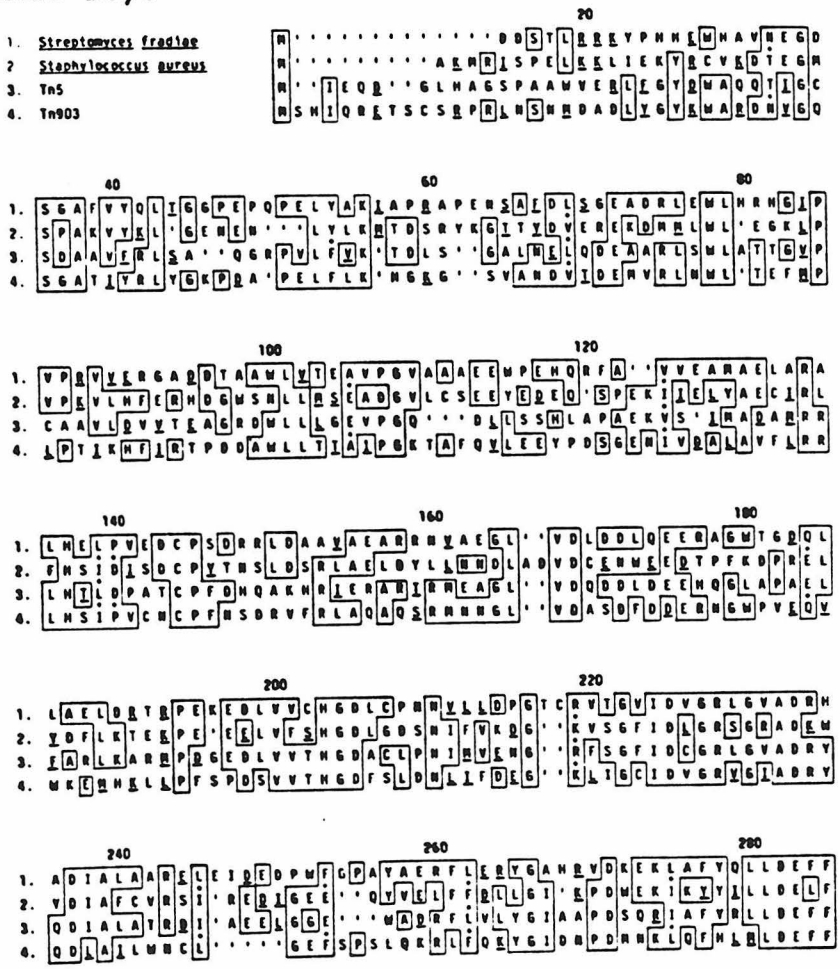
Proposal

The gene for the aminocyclitol-3'-phosphotransferase from Tn5 has been cloned into the HindIII - BamHI site of pBR322; this plasmid is called pGR417 (10). To generate a plasmid that will replicate in both E. coli and B. stearothermophilus, a Cla I fragment (approximately 2 Kb long) from plasmid pBST 1, which is a plasmid that occurs naturally in B. stearothemophilus (5), will be cloned into the Nar I sites of pGR417; this plasmid will be called pTTS. Similar constructions containing the aminoglycoside nucleotidyltrans-

ferase have been successfully propagated in both E. coli and B. stearothermophilus (5).

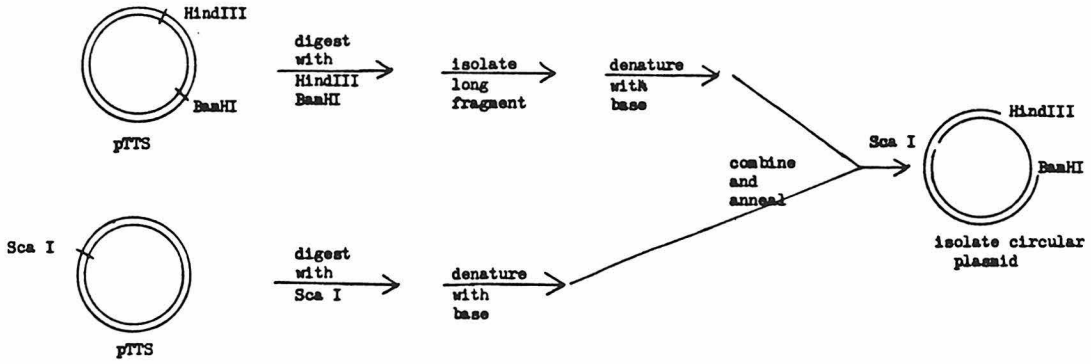
The plasmid described above could be used to transform B. stearothermophilus by using previously reported procedures (5). By growing transformed cells at various temperatures, the maximum temperature at which the APH protein confers kanamycin resistance could be determined; temperatures above this value will be referred to as nonpermissive temperatures.

The Tn5 encoded APH has 10 serine residues (10). Of these 10 residues, only the Ser at residue 36 is conserved among the 4 related APH enzymes from Tn5, S. aureus, S. fradiae, and Tn903 (12) as shown in the figure below (from reference 10).



This residue, therefore, will be mutated separately to determine whether it is essential for activity. At the other 9 residues, a variety of amino acids are present in the four APH proteins. The following experiments are proposed for studying the effect of changing these 9 serine residues to threonine.

By utilizing unique HindIII and BamHI restriction sites in the pTTS plasmid construction described above, a gapped duplex can be generated as shown below; this gap encompasses the entire kan^r gene.



A method for producing such gapped duplexes has been reported previously (13).

Various Ser --> Thr mutations can be generated simultaneously by oligonucleotide-directed mutagenesis (14,15) of this gapped duplex. Seven oligonucleotides will be synthesized to introduce the nine mutations. All seven of these these oligonucleotides will be annealed simultaneously to the gapped duplex, extended with the Klenow fragment of DNA polymerase I, and ligated together with T4 DNA ligase. If the efficiency of mutagenesis is approximately 3% per site, 27% of the resulting plasmids will introduce single site

mutations and 0.07% will introduce double mutations. This estimate assumes that the frequencies of mutation at each site will be additive, which is unclear. Therefore, large numbers of colonies must be screened in the experiments described below. Alternate methods for introducing the mutations more efficiently should also be investigated.

The mixture of plasmids resulting from oligonucleotide-directed mutagenesis will then be used in two experiments: (i) direct screening for increased thermal stability and (ii) random mutagenesis to search for second-site mutants with enhanced thermal stability.

For the first experiment, previously described procedures (5) will be used to transform B. stearothermophilus with the mixture of mutagenized plasmids. The transformation mixture will be plated onto agar plates containing kanamycin and incubated slightly above the permissive temperature. Plasmid would then be isolated from colonies that grow under these conditions. DNA sequencing of these plamids will be used to identify any Ser --> Thr mutations that enhanced the thermal stability of APH.

For the second experiment, the mixture of mutagenized plasmids will be used to transform E. coli and these cells will be grown in broth at 30°C. Plasmids prepared from this mixture will be used to generate circular gapped duplexes as described above. By using dNTP's containing a small proportion of α-thionucleotides and the Klenow fragment of DNA polymerase I, high frequencies of random point mutations

would be generated in the kan^r gene (16); as a control, the same experiment will be performed on gapped heteroduplexes with the wild type sequence. These plasmids would be used to transform B. stearothermophilus; the transformation mixtures will be plated on agar containing kanamycin and incubated at temperatures well above the permissive range. If the control experiment produces as many colonies as the reversion experiment, the cells should be incubated at higher nonpermisive temperatures. If equal numbers of colonies are still present, mutations from the control experiment could be identified and removed from the Ser --> Thr experiment by screening the colonies with synthetic oligonucleotides. If this approach becomes unworkable, individual Ser --> Thr mutations could be isolated from E. coli at 30°C and used in reversion experiments.

The experiments described here are designed to test two possible mechanisms for increasing the thermal stability of proteins. The first possibility, stabilization of proteins by substitution of one or more Ser residues by Thr, would seem unlikely; however, such substitutions are observed when comparing proteins from mesophilic and thermophilic bacteria (8,9). The second set of experiments are designed to test for second site mutations that might complement the initial Ser --> Thr mutation. Several problems may exist in these experiments since such large numbers of mutations will be studied; however, the ability to select for resistance at several nonpermissive temperature will allow one to select the

most stable proteins from among 10^6 or more less stable mutants. Methods such as these that utilize phenotypic selections in thermophiles should have many general applications in studying factors that enhance the thermal stability of proteins.

References

1. Yatami, K., Ogasahara, K. & Sugino, Y. (1985) Adv. Biophys. 20, 13-29.
2. Dalbadie-McFarland, G., Neitzel, J.J., & Richards, J.H. - (1986) Biochemistry 25, 332-338.
3. Perry, L.J. & Wetzel, R. (1984) Science 226, 555-557.
4. Matsumura, M. & Aiba, S. (1985) J. Biol. Chem. 260, 15298-15303.
5. Liao, H., McKenzie, T. & Hageman, R. (1986) Proc. Natl. Acad. Sci. USA 83, 576-580.
6. Yutant, K., Ogasahara, K., Aoki, K., Kakuno, R. & Sugino, Y. (1984) J. Biol. Chem. 259, 14076-14081.
7. Hecht, M.H., Sturtevant, J.M. & Sauer, R.T. (1984) Proc. Natl. Acad. Sci. USA 81, 5685-5689.
8. Argos, P., Rossman, M.G., Grau, U.M., Zuber, H., Frank, G. & Tratschin, J.D. (1979) Biochemistry 18, 5698-5703.
9. Ikai, A. (1980) J. Biochem. 88, 1895-1900.
10. Beck, E., Ludwig, G., Auerswald, E., Reiss, B. & Schaller, H. (1982) Gene 19, 327-336.
11. Davies, J. & Smith, D.I. (1978) Ann. Rev. Microbiol. 32, 469-518.
12. Gray, G.S. & Fitch, W.M. (1983) Mol. Biol. Evol. 1, 57-66.
13. Oka, A., Sugimoto, K. Sasaki, H. & Takanami, M. (1982) Gene 19, 59-63.
14. Dalbadie-McFarland, G. & Richards, J.H. (1983) Ann. Rev. Med. Chem. 18, 237-245.

15. Smith, M. (1985) Ann. Rev. Genet. 19, 423-462.

16. Kunkel, T.A., Eckstein, F., Mildvan, A.S., Koplitz, R.M. & Loeb, L.A. (1981) Proc. Natl. Acad. Sci. USA 78, 6734-6738.

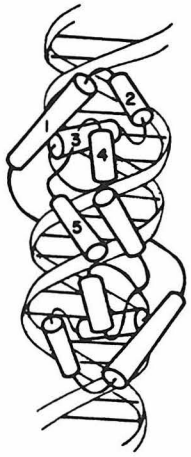
PROPOSITION III

PROBING THE FLEXIBILITY OF THE N-TERMINAL ARMS AND α-α TURN REGIONS OF LAMBDA REPRESSOR HEADPIECE BY NMR

Background

Lambda repressor binds to specific base sequences in two operator regions of lambda phage and subsequently represses transcription of lambda phage proteins while promoting its own transcription (1,2). This activity maintains the lysogenic state of lambda phage and prevents super-infection with other lambda phage DNA. The repressor protein binds as a dimer of identical subunits to an operator region that also has approximate two-fold symmetry (3,4). Each subunit of the dimeric protein contains two domains (4); by treating lambda repressor with papain, one can produce an N-terminal fragment composed of the first 92 amino acids of the protein (5,6). This fragment (commonly called the lambda "headpiece") binds operator DNA as a dimer and exerts both positive and negative control on flanking genes, as does the intact enzyme (5,6).

The crystal structure of the headpiece of lambda repressor shows that this domain contains five α-helices (7), two of which (α-2 and α-3) form a helix-turn-helix structure that apparently is able to contact specific bases in the major groove of the DNA helix (7). A proposed model for the binding of lambda repressor headpiece to DNA is shown below (from reference 7).



Similar helix-turn-helix structures exist in lambda cro repressor (8) and catabolite activator protein (9) as well as many other DNA binding proteins (10,11); this structure probably represents a general motif by which proteins can bind to specific sequences of DNA.

However, lambda repressor differs from cro repressor and CAP in several ways. The α-α turn between helix 2 and 3 contains four rather than three amino acids as in CAP and cro protein (7). The lambda repressor has N-teminal arms that apparently "wrap" around the DNA to make contacts on the opposite side of the helix (12). The biological activity of lambda repressor is also unique in that this protein exerts both positive and negative control from the same operator region (1,2). Interestingly, a mutation in the four residue α-α turn (Gly 43 --> Arg) results in a protein that can still exert negative control on flanking genes, but can no longer exert positive control (13).

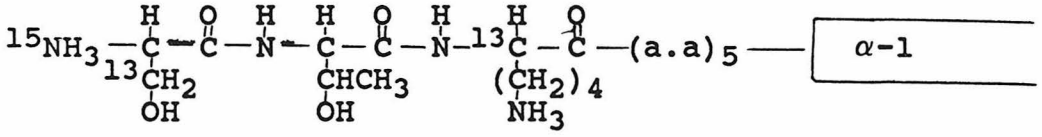
I propose to study the dynamics of the N-terminal arm and α-α turn of the lambda repressor headpiece. By incorporating ^{13}C and ^{15}N spin labels at specific positions in the headpiece of lambda repressor, NMR could be used study local motions in these regions of the protein.

Proposal

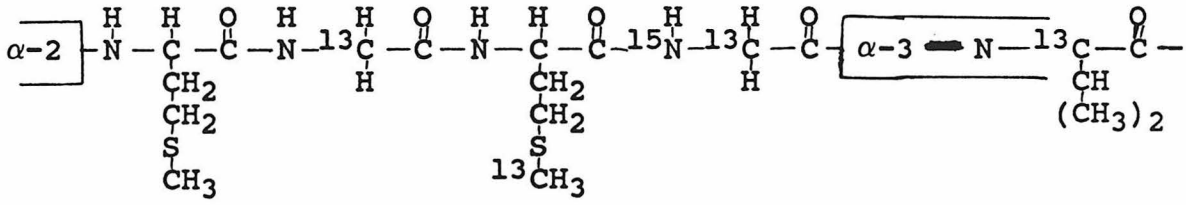
NMR provides a sensitive probe for local motions in macromolecules, especially when spin labels can be placed at specific positions. Many examples of such studies exist

(14,15).

By chemically synthesizing a peptide, spin labels could be incorporated into very specific positions in the sequence. Modern methods for peptide synthesis can be used to synthesize peptides composed of 70 or more amino acids (16); however, this length may may vary from 50 to 150 amino acids depending on the sequence (16). I propose to incorporate the following spin labels into the headpiece of lambda repressor (see figure below) by synthesizing a peptide 92 residues long. Although synthesis of the entire headpiece is desirable, a peptide containing the first 75 amino acids may also be sufficient for these experiments if the 92-mer cannot be synthesized.



Ser 1 Thr 2 Lys 3



Met 40 Gly 41 Met 42 Gly 43 Val 47

The isotopically labelled amino acids required for synthesis of this molecule are commercially available. Identification of the ^{13}C resonances of the peptide should be straitforward. Although minor changes in the chemical shifts of these

nuclei may occur as a result of the conformation of the peptide, reference shifts determined for these residues in short, flexible peptides are listed below (17,18).

Residue	Enriched Nucleus	Shift (in ppm with respect to dioxane)
Ser 1	β-^{13}C	-5.4
Lys 3	α-^{13}C	-12.9
Gly 41	α-^{13}C	-23.95
Met 42	ϵ-^{13}C	-52.3
Gly 43	α-^{13}C	-23.9
Val 47	α-^{13}C	-6.9

The ^{15}N labels will assist in distinguishing between the resonances for Gly 41 and Gly 43 and for identifying the Ser 1 resonance. Proton coupling patterns should unambiguously identify the remaining Gly 41, Lys 3, and Met 42 resonances.

Relative relaxation rates will be determined for (i) free peptide in solution, (ii) peptide in the presence of non-operator DNA, and (iii) peptide in the presence on operator DNA. The following synthetic oligonucleotide segments will be used.

Non-operator DNA	CGTTATCCACAGGGCCTCTATATGC GCAATAGGTGACCCGGAGATATACG
Operator DNA	CGTTTACCTCTGGCGGTGATAATGG GCAAATGGAGACCGCCACTATTACC

The magnitude of the relaxation rates will depend partly on the conditions used in the measurements; therefore, these conditions should be as consistent as possible. In a study of myoglobin, T_1 values (MS) at 67.9 MHz of the carbon atoms

in a lysine residue that lies in an α-helix were as follows: α-^{13}C = 170 ± 8, ϵ-^{13}C = 734 ± 50 (14).

In the absence of operator DNA, the N-terminal residues of the lambda headpiece will probably rotate freely. The first three residues were not observed in the crystal structure of the lambda headpiece indicating that these residues are disordered in the crystal lattice (7). The relaxation rates for the resonances of these nuclei should reflect this free rotation. When the peptide is bound to operator DNA, however, the relaxation rates should decrease dramatically.

The α-α turn between helices 2 and 3 of lambda repressor is probably quite rigid. However, in solution, peptides frequently are more disordered than the intact protein (19). Changes in the mobility of this turn upon binding to non-operator and/or operator DNA should be reflected in the relaxation rates of the labeled nuclei of residues 40, 41, and 42. The relaxation rate of Val 47 will act as a control; this residue should be rigidly held by the α-3 helix and, therefore, the relaxation rate of this nucleus should not change significantly for the three conditions described above.

Several additional experiments could also be performed using the labeled peptide described above. Analysis of the chemical shifts may indicate the nature of the local environments in which the labelled nuclei exist for the various DNA-peptide complexes. NOE experiments with the ^{31}P nuclei of the phosphate backbone of operator and non-operator DNA

may help determine the proximity of the labeled residues to the DNA phosphates. Heteronuclear 2-D NMR or magnetization transfer experiments may allow one to utilize proton resonances of the labeled residues.

The primary purpose of these experiments, however, is to determine the relative mobilities of the N-terminal arm and the α-α turn of lambda repressor headpiece when the peptide is free in solution, nonspecifically bound to non-operator DNA, and specifically bound to operator DNA. The N-terminal arm is a unique feature of lambda repressor. These residues probably function only to stabilize the specific complex; however, the behavior of these residues in the nonspecific complex have not been investigated. Determining the mobility of the α-α turn between helices 2 and 3 in lambda repressor whether flexible or rigid, will have implications in the study of the behaviour of this structural unit in recognition of specific sequences in the DNA helix. The α-α turn of lambda repressor is especially interesting since the Gly 43 --> Ala mutant was found to exert normal negative control but not positive control on flanking genes.

Problems may exist in the experiments described above. For example, accurate values for relaxation rates can be difficult to obtain; several factors can introduce errors into individual experiments. Incorporating an α-^{13}C into Val 47 should be useful as an internal control for estimating such errors; Val 47, which lies in the α-3 helix and faces away from the DNA helix, should be rigid and, therefore, the

relaxation rates of this nucleus should be relatively constant for the various conditions.

Despite possible difficulties, the experiments described here should be very effective for determining whether the N-terminal and α-α turn regions of the DNA helix are flexible or rigid when the protein is (i) free in solution, (ii) nonspecifically bound to non-operator DNA, and (iii) specifically bound to operator DNA.

References

1. Hershey, A.D. & Dove, W. (1971) in The Bacteriophage Lambda (ed. Hershey, A.D., Cold Spring Harbor Laboratory, New York, 1971), 3-12.

2. Botstein, D. & Herskowitz, I. (1974) Nature (London) 251, 584-589.

3. Chadwich, P., Pirrotta, V., Steinberg, R., Hopkins, N. & Ptashne, M. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 283-294.

4. Ptashne, M. Jeffery, A., Johnson, A.D., Maurer, R., Meyers, B.J., Pabo, C.O., Roberts, T.M. & Sauer, R.T. (1980) Cell 19, 1-11.

5. Johnson, A.D., Poteete, A.R., Lauer, G., Sauer, R.T., Ackers, G.K. & Ptashne, M. (1981) Nature (London) 294, 217-223.

6. Pabo, C.O., Sauer, R.T., Sturtevant, J.M. & Ptashne, M. (1979) Proc. Natl. Acad. Sci. USA 76, 1608-1612.

7. Pabo, C.O. & Lewis, M. (1982) Nature (London) 298, 443-451.

8. Anderson, W.F., Ohlendorf, D.H., Takeda, Y. & Matthews, B.W. (1981) Nature (London) 290, 754-758.

9. McKay, D.B. & Steitz, T.A. (1981) Nature (London) 290, 744-749.

10. Sauer, R.T., Yocum, R.R., Doolittle, R.F., Lewis, M. & Pabo, C.O. (1982) Nature (London) 298, 447-451.

11. Schevitz, R.W., Otwinowski, Z., Joachimiak, A., Lawson, C. & Sigler, P.B. (1985) Nature (London) 317, 782-786.

12. Pabo, C.O., Krovatin, W., Jeffrey, A. & Sauer, R.T. (1982) Nature (London) 298, 441-447.

13. Guarente, L., Nye, J.S., Hochschild, A. & Ptashne, M. (1982) Proc. Natl. Acad. Sci. USA 79, 2236-2239.

14. Wittebort, R.J., Szabo, A., Gurd, F.R.N., J. Am. Chem. Soc. 102, 5723-5728.

15. Henery, G.D., Weiner, J.H. & Sykes, B.D. (1986) Biochemistry 25, 590-598.

16. Suzanne Horvath, personal communication.

17. Richarz, R. & Wüthrich, K. (1978) Biopolymers 17, 2133-2141.

18. Keim, P., Vigna, R.A., Marshall, R.C. & Gurd, F.R.N. (1973) J. Biol. Chem. 248, 6104-6113.

19. Walters, L. & Kaiser, E.T. (1985) J. Am. Chem. Soc. 107, 6422-6424.

PROPOSITION IV

EFFECTS OF FLANKING SEQUENCES ON FACILLITATED DIFFUSION OF EcoRI

Background

Although proteins that bind to specific base sequences on the DNA helix must search for their target sites in cells from among $>10^7$ non-target sites, such proteins can very rapidly locate their recognition sequences. A second order rate constant of 10^{10} $M^{-1}s^{-1}$ was observed for binding of lac repressor to a lac operator that was inserted into bacteriophage lambda DNA (1); yet, the second order rate constant for a one step diffusion-controlled interaction between the repressor and operator should be only 10^7 or 10^8 $M^{-1}s^{-1}$ (2). Apparently, DNA binding proteins can surpass the diffusion control limit by first binding to the DNA helix nonspecifically and then diffusing along the helix to search for its recognition site. This mechanism is commonly called facilitated diffusion. By utilizing such a mechanism, the monotony of the DNA helix is transformed from an obstacle into an asset for rapid location of specific sequences of base pairs from among many sequences.

Four mechanism have been proposed for facilitated diffusion (2). The first two involve positionally uncorrelated sites: (i) macroscopic dissociation and reassociation between positionally uncorrelated sites and (ii) transfer of the protein between different segments of the DNA strand. The second two mechanisms involve positionally correlated sites: (iii) microscopic dissociation and reassociation (hopping) and (iv) sliding along the contour length of the double helix. Presently, several studies suggest that facil-

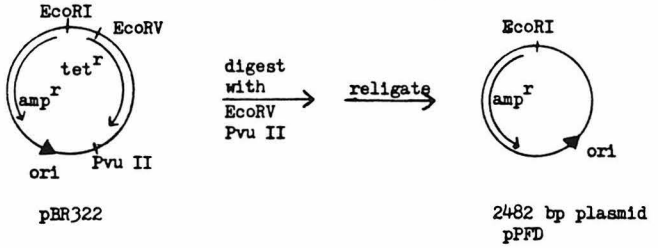
itated diffusion proceeds via one of the positionally correlated mechanisms (3,4), although a positionally uncorrelated mechanism may function in cooperation with a positionally correlated mechanism (5,6,7).

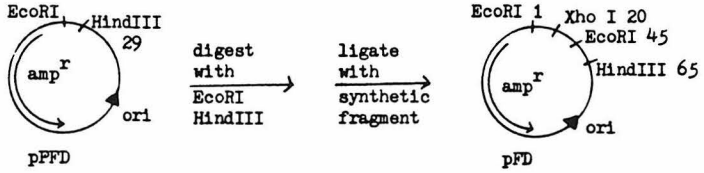
Several restriction enzymes locate their recognition sites by facilitated diffusion in vitro at low ionic strengths (4,8) such that the rate of cleavage increases as the lengths of DNA segments flanking the recognition site increases; the mean diffusion length is approximately 1000 base pairs in 1 mM $MgCl_2$ for EcoRI. However, this rate enhancement is very sensitive to the ionic strength of the solution; above ionic strengths of approximately 0.2, no rate enhancement occurs (4,8).

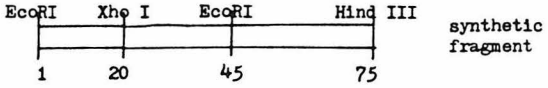
Implicit in the discussion of facilitated diffusion by a positionally correlated mechanism is a regularity in the structure of the double helix; for a DNA binding protein to move freely along the helix, many identical, closely spaced sites must exist. Therefore, facilitated diffusion by a positionally correlated mechanism may be affected by local distortions in the DNA helix. For some resriction enzymes, flanking sequences are known to affect the rate of cleavage, although such effects may or may not result from differences in facilitated diffusion. Several distortions are known to occur due to particular arrangements of bases in the DNA helix (9,10). I propose to study the effect of sequence induced variations in the double helix on facilitated diffusion for EcoRI.

Proposal

The proposed experiments will require an effective and general assay. Two approaches could be used. First, one EcoRI site could be placed at one end of a long double-stranded DNA segment; a variety of sequences could be placed in front of this site and the absolute rate of cleavage could be determined for various constructions. The second approach is to place an EcoRI site on each end of a long double-stranded DNA segment; a variety of sequences could then be placed in front of one of these sites and the relative rates of cleavage at the two sites could be determined. An assay that measures relative rates of cleavage for sites containing normal and altered flanking sequences would probably be more reliable than measuring and comparing absolute rates of cleavage for different fragments since the purity and quality of DNA and protein as well as minor variations in the conditions could cause significant errors in absolute rates. The following plasmid construction is proposed.







For determining the relative rates of cleavage for these two EcoRI sites, the following procedure is proposed. First, the plasmid will be digested with Xho I and radiolabeled using [α-^{32}P]dTTP and the large (Klenow) fragment of DNA polymerase I. This linear fragment will then be treated with low concentrations of EcoRI at low ionic strength (~ 0.1) and let react for various time periods. The resulting mixtures of fragments will be loaded onto a 20% polyacrylamide gel and electrophoresed. Cleavage of one site will produce a fragment 20 bases long and cleavage at the other site will produce a fragment 25 bases long. An audioradiogram of the gel will be obtained and the relative quantities of the two fragments will be determined using a densitometer. The relative quantities of the two fragments should reflect the relative rates of cleavage at the two sites. If the rates of cleavage at the two sites differ significantly in this construction, an alternate construction may be necessary or a new assay utilizing absolute rates of cleavage could be developed.

To test the effects of flanking sequences on facilitated diffusion, the following synthetic oligonucleotides could be incorporated into the Xho I - HindIII sites:

I TCGAG/ /ATATATATATATATATATATA
C/ /TATATATATATATATATATATTCGA

II TCGAG/ /GCGCGCGCGCGCGCGCGCGCA
C/ /CGCGCGCGCGCGCGCGCGCGTTCGA

III TCGAG/ /ATATATGCGCATATATGCGCATATAT
C/ /TATATACGCGTATATACGCGTATATATTCGA

IV TCGAG/ /GTGTGTGTGTGTGTGTGTGTA
C/ /CACACACACACACACACACATTCGA

V TCGAG/ /CAAAATGTCAAAAAATAGGCAAAAAATGCCAAAAATCA
C/ /GTTTTACAGTTTTTTATCCGTTTTTTACGGTTTTTAGTTCGA

VI TCGAG/ /GAATTTCGAAATTCCGAAATTCGGGAATTTA
C/ /CTTAAAGCTTTAAGGCTTTAAGCCCTTAAATTCGA

VII TCAG/ /GTGTGTGTGTGTGTGTACACACACACACACACACGTACGTA
C/ /GGTTGGTTGGTTGGTTAACCAACCAACCAACCTGCATGCATTCGA

The first six sequences could probably be cloned into the plasmid and grown in E. coli if desired. Alternatively, these fragments could be ligated into plasmids and the resulting DNA could be used directly; this latter approach is the only one that could be used for sequence VII.

Sequences I - III will study the effects of differences in melting and hydration of the minor groove. The $(AT)_n$ sequences are easily melted and their minor grooves are hydrated in a very orderly manner (10). The $(GC)_n$ sequences are not easily melted and their minor grooves are hydrated very poorly. Sequences III and IV are intermediate and could help distinguish whether a decrease in rate is due to differences in hydration or melting.

Sequence V should introduce a large bend into the DNA fragment, perhaps as much as 40^{o} to 80^{o} (9).

Sequence VI would introduce four sites that contain 5 out of 6 bases from the recognition sequence of EcoRI. The enzyme may stall at sequences similar to its recognition sequence.

Sequence VII will introduce a cruciform structure into

the helix.

By measuring the ratio of the labeled 20 and 25 base segments obtained from digestion with EcoRI at various time points, as discussed above, the preference for a particular site could be determined. The enzyme may simply stall at the perturbed region of the helix or it may dissociate from the helix. These two possibilities may be distinguishable by carefully measuring the absolute rate of cleavage at the nonperturbed site. The rate of cleavage at this site should be reduced if the enzyme stalls since the concentration of free EcoRI will be lowered; the magnitude of this reduction in rate will depend on how long the enzyme stalls. The rate of cleavage at the non-perturbed site will either remain the same or be enhanced if the perturbed region of the helix causes dissociation of the enzyme from the nonspecified complex; an enhanced rate would occur if this dissociation is greater than occurs after cleavage of an EcoRI site.

If the ratio of cleavage at the two sites does not change for the various DNA constructions, the results would indicate that either (i) facilitated diffusion is more positionally uncorrelated than previously expected (4,8) or (ii) the nature of the nonspecific complex is less dependent on the structure of the DNA helix than previously proposed (2,3,6).

The assay system presented here is quite general and could be used to test many different flanking sequences. Other restriction enzymes could also be tested.

References

1. Riggs, A.D., Suzuki, H. & Bourgeois, S. (1970) J. Mol. Biol. 48, 67-83.
2. Berg, O.G., Winter, R.B. & von Hippel, P.H. (1981) Biochemistry 20, 6929-6948.
3. Winter, R.B., Berg, O.G. & von Hippel (1981) Biochemistry 20, 6961-6977.
4. Terry, B.J., Jack, W.E. & Modrich, P. (1985) J. Biol. Chem. 260, 13130-13137.
5. Fried, M.G. & Crothers, D.M. (1984) J. Mol. Biol. 172, 263-282.
6. Winter, R.B. & von Hippel, P.H. (1981) Biochemistry 20, 6948-6960.
7. von Hippel, P.H. (1979) in Bacterial Regulation and Development (Goldberger, R.F., ed., Plenum Press, NY) 279-347.
8. Ehbrecht, H.-J., Pingoud, A., Urbanke, C., Maass, G. & Gualerzi, C. (1985) J. Biol. Chem. 260, 6160-6166.
9. Thomas, M. & Davis, R.W. (1975) J. Mol. Biol. 91, 315.
10. Forsblum, S., Rigler, R., Ehrenberg, M., Pettersson, U. & Philipson, L. (1976) Nucl. Acids Res. 3, 3255.
11. Nath, K. & Azzolina, B.A. (1981) in Gene Amplification and Analysis (ed. J.G. Chirikjian, 113 Elsevier North Holland, Inc., NY)
12. Wu, H.-M. & Crothers, D.M. (1984) Nature (London) 308, 509-513.

13. Dickerson, R.E. & Drew, H.R. (1981) J. Mol. Biol. 149, 761-786.

PROPOSITION V

INVESTIGATIONS OF *IN VIVO* AND *IN VITRO* MISMATCH REPAIR IN YEAST

Background

The existence of mechanisms by which cells can repair mismatched base pairs in DNA was originally proposed to explain gene conversion in genetic recombination (1). Since then, numerous studies have been directed toward elucidating the nature of mismatch repair (2,3). Several mutant strains of E. coli that exhibit deficiencies in mismatch repair have been isolated (3,4,5). These cells exhibit "mutative" phenotypes in which spontaneous mutations occur at much higher frequencies than in wild type cells. Such studies indicate that for cells to maintain their genetic integrity, they must correct mismatches that arise from imperfections in replication and other cell processes. However, for a repair system to be effective, strand discrimination is necessary such that mismatches will be converted back to the original base pair. In E. coli, this discrimination apparently arises as a result of methylation (2,6,7); the newly synthesized strand will not be methylated and, therefore, a mechanism that repairs only the unmethylated strand will exhibit the necessary discrimination. Recent results from cell free extracts of E. coli that are capable of mismatch repair strongly supports this scheme (8).

Although such studies are rapidly elucidating many aspects of prokaryotic mismatch repair, little is known about such processes in eukaryotes. Although genetic studies have strongly suggested that mismatch repair does exist in eukaryotic organisms (9), only recently has the first direct

analysis been reported (10). In this study, heteroduplex plasmids containing small insertions were used to transform yeast; analysis of plasmids derived from tranformed clones indicated that one site was repaired in 85% of the clones and the second site was repaired in greater than 97% of the clones.

I propose to develop an efficient assay to investigate mismatch repair in yeast in vivo and possibly in vitro.

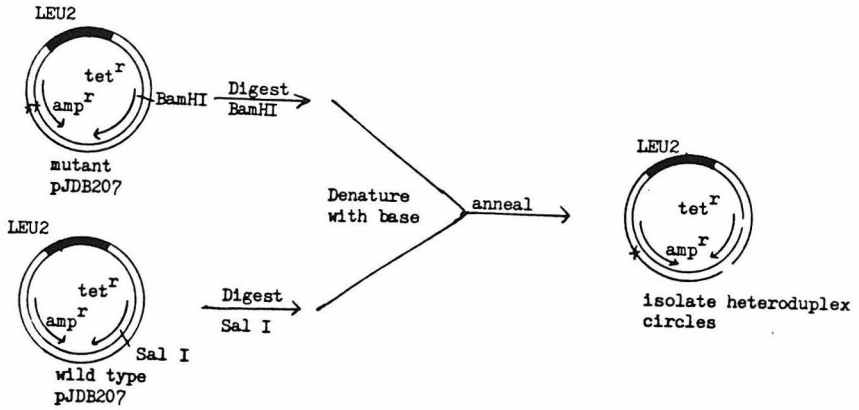
Proposal

I propose to use four different mutants of ß-lactamase to probe mismatch repair in yeast (S. cerevisiae). Two of the mutants are already available, Ser 70 --> Thr (AGC --> ACC) (11) and Lys 73 --> Arg (AAA --> AAG) (12), and two will be produced by standard techniques for oligonucleotide-directed mutagenesis (13), Met 1 --> Thr (ATG --> AGC) and Ile 279 --> Asn (ATC --> AAC); the first two mutants are inactive and the second two should also result in an amp^S phenotype.

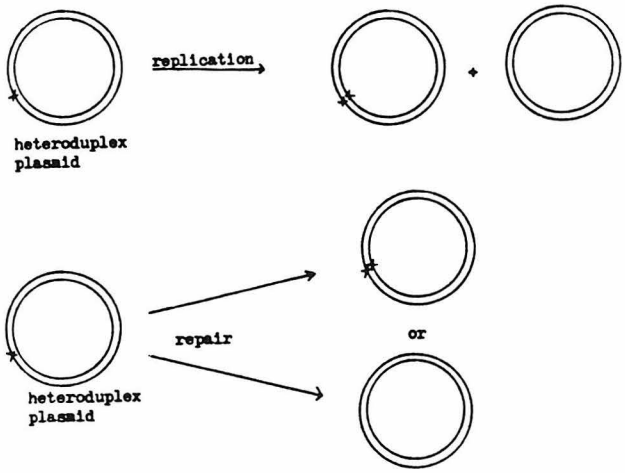
Each of these mutations will be incorporated into plasmid pJDB207 (14). This is 6.9 Kb plasmid that contains the entire pBR322 sequence as well as an origin of replication for yeast and a Leu2 gene (which is a selectable marker for $Leu2^-$ yeast strains). This plasmid will replicate in both E. coli and yeast. The mutant ß-lactamase genes discussed above will be introduced into Ssp I and Sal I sites of pJDB207.

Using these plasmids, heteroduplexes will be generated

as shown below:



The four heteroduplex plasmids resulting from the four mutants described above will then be used to transform S. cerevisiae. Plasmid derived from individual colonies will be isolated and used to transform E. coli. The transformed E. coli cells will be grown on plates containing tetracycline or ampicillin and the ratio of ampicillin:tetracycline resistant colonies will be determined. If mismatch repair did not occur, this ratio should be 1:2 for amp^r:tet^r colonies. If all of the mismatches were corrected, 50% of the plasmids derived from individual yeast colonies should give ratios of 0:1 and 50% should give ratios of 1:1. If mismatch repair occurs only part of the time, plasmids derived from different individual colonies will give all three ratios of 1:2, 0:1, and 1:1. Diagrams of these predictions are shown below.



Four different mutants are used since we might expect certain mismatches are apparently to be repaired more readily than others. The heteroduplexes described above will test 4 different transitions and 4 different transversions. Notice also that since both strand are used in generating the duplexes, strand discrimination will not interfer with the results of these experiments.

If the results from these experiments indicate that mismatch repair is occurring as one would expect, new heteroduplex plasmids will be generated as described previously. However, two different mutant plasmids will be used to generate the heteroduplex rather than one mutant and one wild type plamid; this will produce plasmids that contain two mismatches rather than one. Four constructions are proposed: (i) Ser 70 --> Thr & Lys 73 --> Arg, (ii) Met 1 --> Thr & Ser 71 --> Thr, (iii) Ser 70 --> Thr & Ile 279 --> Asn, and (iv) Met 1 --> Thr & Ile 279 --> Asn. These constructions will contain mismatches separated by 9, 209, 626, and 833 base pairs respectively. The resulting heteroduplex plasmids will be used in the same experiments described above.

These experiments should indicate whether the two mismatches are repaired independently or together. Proposed mechanisms for repair generally invoke production of a gap in the DNA helix which is then repaired (2); the size of the gap is uncertain and possibly quite variable. The experiments described here should indicate the size of the gap. If the

repair events are not independent (i.e., the gap covers both sites), no ampicillin resistant colonies will be obtained. If the events are entirely independent, the fraction of yeast colonies containing plasmids that produce 1:1 ratios of amp^r:tet^r phenotypes in E. coli will be the product of the fractions obtained for the two heteroduplexes containing the corresponding single mismatched base pairs; if repair occurs 100% of the time, 25% of the yeast colonies should contain plasmids that produce 1:1 ratios of tet^r:amp^r phenotypes in E. coli. If the observed fractions are not the product of the two fractions from the individual events, the events are either partially independent and partially non-independent or strand discrimination is occurring, although this seems unlikely since both strands were originally derived from the same E. coli strain.

For the heteroduplex constructions containing two mismatches as described above, if gaps are produced in the repair of heteroduplexes, one might expect repair of the Ser 70 --> Thr & Lys 73 --> Arg plasmids to be frequently non-independent events since they are only 9 base pairs apart. If the gap forms only on one side of the mismatched base, both independent and non-independent repair will be observed, possibly even when the mismatches are only 9 bases apart. For the Met 1 --> Thr & Ile 279 --> Asn constructions, the mismatches are much further apart (833 bp) and, therefore, may be repaired independently unless the gap that forms is very large.

If the repair events are sometimes independent for one of the heteroduplex constructions, this assay would be quite useful for determining the activity of an *in vitro* mismatch repair system. Ideally, one could simply treat plasmid with cell extract, re-isolate the plasmid, transform *E. coli*, and then selecting for ampicillin resistance; amp^r phenotypes would only be observed if mismatch repair had occurred. However, since replication may not occur in an *in vitro* system, unrepaired heteroduplexes could be repaired after incorporation into *E. coli* and thereby produce an ampr phenotype. Mutant strains of *E. coli* dificient in mismatch repair (5) could possibly be used to avoid such problems. This possibility can be easily tested by transforming these *E. coli* strains with the various heteroduplex plasmids.

Recently, a method was reported for preparing cell free extracts that were capable of recombination (15). I propose to use the same preparation to screen heteroduplexes for mismatch repair. An assay that utilizes heteroduplexes containing two mismatched base pairs would provide a very sensitive screen for mismatch repair. Such a system could eventually help in isolating and characterizing the components of mismatch repair in yeast.

Many other experiments are also possible by utilizing the assay system described here. For example, several DNA repair systems are induced by stress (16); one could measure mismatch repair in cells after various forms of stress such as heat shock and UV irradiation. Also, large numbers of

mutant strains of yeast have been isolated, some of which are deficient in DNA repair systems induced by stress (16); one could test for mismatch repair in these strains as well. Another interesting experiment would be to generate double mismatch heteroduplexes as described previously, except obtain one of the mutant plasmids from a dam$^-$ strain of E. coli (17); in this case one strand will not be methylated and one will. One could then determine if methylation can give rise to strand discrimination in yeast.

The experiments described here should be quite useful for studying mismatch repair systems in yeast. Although substantial amounts of genetic data exist on DNA repair, very little is known about the biochemistry of these processes. The studies discussed above could greatly assist in beginning such studies.

References

1. Holiday, R. (1964) Genet. Res. 5, 282-290.

2. Wagner, R., Jr., & Meselson, M. (1976) Proc. Natl. Acad. Sci. USA 73, 4135-4139.

3. Nevers, P. & Spatz, H. (1975) Mol. Gen. Genet. 139, 233-243.

4. Tiraby, J.G. & Fox, M.S. (1973) Proc. Natl. Acad. Sci. USA 70, 3541-3545.

5. Glickman, B.W. & Radman, M. (1980) Proc. Natl. Acad. Sci. USA 77, 1063-1067.

6. Glickman, B.W., van den Elsen, P. & Radman, M. (1978) Mol. Gen. Genet. 163, 307-312.

7. Marinus, M.G. & Moris, N.R. (1975) Mutat. Res. 28, 15-26.

8. Lu, A.-L., Welsh, K., Clark, S., Su, S.-S. & Modrich, P. (1984) Cold Spring Harbor Symp. Quant. Biol. 49, 589-596.

9. Fincham, J.R.S. & Holliday, R. (1970) Mol. Gen. Genet. 109, 309-322.

10. Bishop, D.K., Williamson, M.S., Fogel, S. & Kolodner, R.D. (1986) J. Cell. Biochem. Supplement, UCLA Symp. Mol. & Cell Bio. 10B, 208.

11. Dalbadie-McFarland, G., Neitzel, J.J. & Richards, J.H. (1986) Biochemistry 25, 332-338.

12. Neitzel, J., unpublished results.

13. Smith, M. (1985) Ann. Rev. Genet. 19, 423-462.

14. Broach, J.R. (1983) Met. Enzymol. 101, 307-325.

15. Symington, L.S., Morrison, P.T. & Kolodner, R. (1984) Cold Spring Harbor Symp. Quant. Biol. 49, 805-814.

16. Friedberg, E. (1985) DNA Repair (W.H. Freeman and Co., San Francisco).

17. Lacks, S. & Greenberg, B. (1977) J. Mol. Biol. 114, 153-168.