STRUCTURAL REQUIREMENTS FOR PROTEIN FUNCTION STUDIED BY IN VITRO MUTAGENESIS ON BETA-LACTAMASE

Thesis by

Gloria Dalbadie-McFarland

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1985

(Submitted September 24, 1984)

ACKNOWLEDGEMENT

My years as a Caltech graduate student have been not only challenging, but enjoyable. I would particularly like to express my appreciation to John H. Richards, my advisor. His encouragement and support, and the opportunities he provided to work with leading scientists both at Caltech and other institutions made it all possible.

Thank you, Jack.

Arthur D. Riggs at the City of Hope contributed a great deal of his time, concern, and facilities to my efforts and I am grateful.

I would like to thank David Botstein for his guidance in the genetic transfer of beta-lactamase mutants to P22 phage during my visit to his lab, and the opportunity to spend time with his very dynamic group.

Charlotte B. Clark provided excellent technical assistance.

The contributions of the Richards group, particularly Steve Schultz and Jim Neitzel, are greatly appreciated.

My work was supported by National Institutes of Health Grants GM16424, , GM07616-06, GM06716-05, GM25825, GM30395,

the Office of Naval Research Grant N00014-83-K-0487, the William Barton Jones Fellowship and the W. R. Grace Fellowship.

Finally, I thank my family for their support; my sister, Lu Dalbadie-Ross, my mother, Luciana Dalbadie, who always encouraged me to be all I could be, and my brother, Peter Dalbadie, who with great good cheer gave me his ongoing support, then his organizational skills to help edit this manuscript.

Wayne, my friend, what can I say.

STRUCTURAL REQUIREMENTS FOR PROTEIN FUNCTION STUDIED BY IN VITRO MUTAGENESIS ON BETA-LACTAMASE

ABSTRACT

The study of naturally occurring variants of proteins has been successfully used for a long time to assign roles to structural elements in a protein and to correlate functional requirements with the nature of these structural elements.

Advances in techniques for DNA synthesis and DNA sequencing, along with the development of recombinant DNA techniques now allow one to clone the gene for a protein and then to modify it at will. In this way, presumably any and all aminoacid substitutions can be engineered into any protein whose gene has been cloned, characterized and expressed.

Beta-lactamase has been used as a model system in which to study the feasibility of the approach, and it has been demonstrated that not only is it possible to introduce

specific predetermined changes in the structure of a protein, but that these mutants can in turn serve as substrates for further modifications. An inactive enzyme can be used to search for a broad range of structural requirements by imposing selective conditions that require a function for the survival of the host organism for the mutant protein.

Four variants of beta lactamase, one of which is catalytically active, have been obtained by site specific mutagenesis. In the inactive mutants the conserved active site sequence -ser70-thr71- was altred to either -thr70-ser71- , -thr70-thr-71- or -arg-70-thr71- ; a variant in which a disulfide bond was removed by mutating one of the only two cysteines in E. coli beta-lactamse to serine was found to be active. In addition, a revertant to activity has been obtained from one of the the inactive mutants (-thr70-ser71-). The revertant is different in aminoacid sequence (-ser70-ser71-) and in some of its properties from the wild type enzyme while still having catalytic activity. No revertants with aminoacid substitutions at a secondary site were found.

Both the catalytically active revertant and the mutant

lacking the disulfide bridge were found to have reduced thermal stability. The rate of secretion of three of these mutants (ser70-->thr, thr71-->ser and ser70-->thr/thr71-->ser) was compared to that of the wild type beta lactamase and no significant differences were found. The -arg70-ser71- mutant will be used to identify chemical mutagens and carcinogens that induce GC to TA transversions by collaborators P. L. Foster and D. Botstein.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Chapter II

LIST OF TABLES

NOMENCLATURE

Expressions of the form ser70-->thr are meant to indicate that a mutation which replaces serine with threonine at position 70 of a protein in which position 1 is the first N terminal aminoacid, has been introduced.

ABBREVIATIONS

DTT is dithiothreitol, SDS is sodium dodecyl sulfate, SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0 buffer and TE is 10 mM Tris-HCl/1 mM EDTA pH 8.0 buffer.

INTRODUCTION

IN VITRO MUTAGENESIS:
POWERFUL NEW TECHNIQUES FOR STUDYING
STRUCTURE-FUNCTION RELATIONSHIPS IN PROTEINS

2

Chapter 24. In Vitro Mutagenesis: Powerful New Techniques for Studying Structure-Function Relationships in Proteins

Gloria Dalbadie-McFarland and John H. Richards
California Institute of Technology, Pasadena, California 91125

Introduction - The linear sequence of amino acids uniquely determines the exquisite three-dimensional structures of proteins and their wide diversity of biological function--as catalysts, hormones, transport agents, cell surface receptors, structural elements, transducers of chemical energy into work. To be able to generate, at will, any sequence of amino acids will therefore allow rational study of the relation between protein structure and function. This objective has become possible only recently through development of reliable methods for chemical synthesis of oligodeoxynucleotides and DNA sequence analysis; these advances, along with rapidly developing cloning technology, have made possible a large variety of genetic manipulations that alter the structures of genes and of gene products. Not only do these techniques now allow one to generate essentially any variant of a protein to study how three dimensional structure and biochemical function depend on amino acid sequence, they further ensure that, once the DNA of the structural gene has been prepared, the protein can be produced ever after in any quantities desired.

In vitro mutagenesis of a cloned structural gene of a protein can introduce deletions, insertions and single or multiple amino acid substitutions either randomly within a pre-selected region or at a single, specific site. In addition to these mutagenic techniques, the in vitro synthesis of entire structural genes has been accomplished.[1-4] Together, these techniques now allow the relatively rapid construction of a protein with any predetermined amino acid sequence. In short, one can specify a particular structure and determine its function. Analogous approaches have, over the years, greatly enriched our resources of useful pharmaceuticals. The procedures characteristic of this approach are those of site-directed, specific mutagenesis.

For structure-function studies of proteins, a second powerfully complementary approach should also be emphasized. This approach takes advantage of the ability of biological systems to produce a very large number of random structural variants; these can then be screened for those variants that have a particular function. In this case, one specifies a particular function and then determines which structures, of many millions that can be easily tested, have that function. The procedures useful in this approach are those of random mutagenesis, perhaps directed toward a particular region of the protein.

This review will focus particularly on those aspects of mutagenesis that have a direct bearing on the study of protein function. These techniques have also been used to address a large variety of genetic problems as well as to probe structure-function relationships; the more general subject has been reviewed extensively.[5-8]

ANNUAL REPORTS IN MEDICINAL CHEMISTRY—18

ISBN 0-12-040518-0

Amino Acid Substitutions - The roles of particular amino acid residues in catalysis, ligand binding, folding and in determining other properties of a protein can be investigated by the study of the effects on these properties of both specific and non-specific substitutions. Mutant proteins can be generated by methods that fall into four major categories: (a) deliberate, site-specific substitutions, (b) segment directed, random mutagenesis, (c) non-directed, random mutagenesis, and (d) suppression of nonsense codons generated by an appropriate mutagenic procedure.

a. Site Specific Mutagenesis - Unique, predetermined base changes at any desired specific site in a structural gene, leading to any desired amino acid substitution in the corresponding protein, can be readily accomplished by oligonucleotide-directed mutagenesis; this is the most specific and generally applicable form of site-directed mutagenesis. First developed in the single stranded phage $\phi\chi$ 174,[9,10] this method has since been used to produce several point mutations in this phage,[11-15] as well as in M13[16] and genes cloned into M13.[17-19] The technique has been extended also to double-stranded circular DNAs.[20-22]

As outlined in Figure 1, this method involves priming in vitro DNA replication with a chemically synthesized oligodeoxynucleotide that has been designated to have a sequence that is largely complementary to the "wild-type" DNA template in the region where the mutation(s) is(are) to be introduced. The synthetic nucleotide contains, however, one, or more,

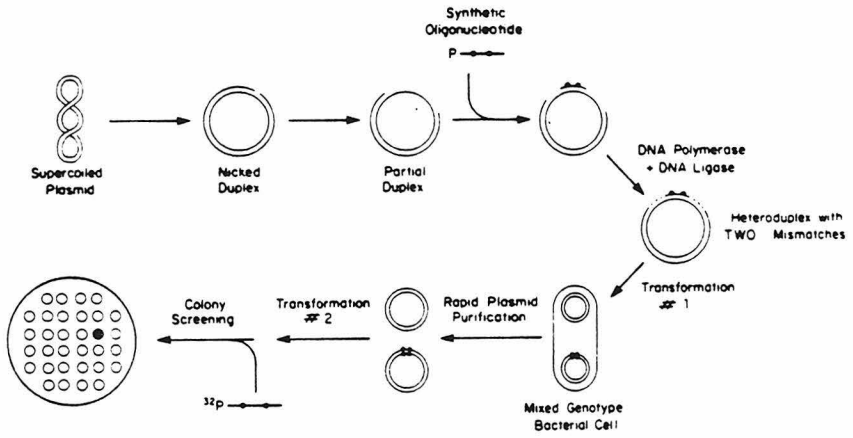


FIGURE 1: Overall scheme for oligonucleotide-directed mutagenesis of double stranded circular plasmid DNA. Supercoiled plasmid circles are nicked in one strand and rendered partially single stranded by treatment with exonuclease. The gapped circles are hybridized with a homologous oligodeoxynucleotide carrying, by design, some mismatches. In vitro DNA synthesis, primed in part by the oligodeoxynucleotide, leads to heteroduplex plasmid circles. Molecular cloning and in vivo DNA replication generates homoduplexes, some of which have the DNA sequence of the primer oligodeoxynucleotide. Colony screening, with the same oligodeoxynucleotide labeled with ^{32}P as a hybridization probe, allows identification of the desired mutant colony regardless of its phenotype.

bases that are not complementary to the wild-type template; the mutation originates from these mismatched bases. Annealing of the synthetic oligonucleotide to the wild-type circular DNA, which must of course be at least

partially single stranded, can be accomplished under conditions of sufficiently low strigency that a heteroduplex forms, in spite of some mismatching of bases. The site-specificity of this duplex formation is, nevertheless, exceedingly high. For example, in a 15 base oligodeoxynucleotide there will normally be 13-14 complementary base pairings and 1-2 mismatches at the site where the mutation is being introduced. Binding of the synthetic oligonucleotide to the template DNA anywhere else than at the desired site will generally involve no more than 9-10 pairings by design. The heteroduplex region thus formed now serves as an initiation site for in vitro DNA synthesis that occurs in the presence of the four deoxynucleotide triphosphates and is catalyzed by the Klenow fragment of DNA polymerase I (which requires a double stranded region to begin DNA replication). After ligation, the heteroduplex is introduced into an appropriate bacterial host and multiplies as a result of normal semiconservative, in vivo replication into two types of homoduplexes; the mutant homoduplexes have a sequence that reflects the synthetic, mutagenic oligonucleotide; the other homoduplexes have the original, wild-type sequence.

An essential feature of any procedure for producing specific mutants is the necessity of identifying mutant colonies in a background of wild-type colonies. To be generally applicable, the procedure should allow identification of mutants at the level of DNA and not rely on some phenotypic difference between mutant and wild-type that may be difficult to predict and to assay. In this regard, oligonucleotide mutagenesis has the distinct advantage that mutants can be found by screening bacterial colonies with the same synthetic oligonucleotide, now labeled with ^{32}P, that was originally used to introduce the mutation.[19-22] In this case, the hybridization is carried out under conditions of high stringency such that even a single base mismatch between oligonucleotide probe and template can be readily distinguished from perfectly matching hybridization. This allows clear discrimination between wild-type and mutant colonies of bacteria randomly spread on a Petri dish.[21,23]

We have recently demonstrated that these procedures for oligonucleotide directed mutagenesis applied to the structural gene of an expressed protein can be used to produce mutant proteins with amino acid substitutions introduced deliberately and precisely, and we have shown that these sequence changes can be generated with a negligible background of non-specific mutations.[20] Both single and double base changes, leading to single and double amino acid substitutions, were created in the structural gene for β-lactamase in the double stranded circular plasmid pBR322. Specifically, synthetic 15 and 16 base oligomers, respectively, were used to mutate the wild-type sequence, 5'-AGC ACT-3' (-ser70 thr71-) to 5'-AC̆C TC̆T-3'; (-thr ser-)(a double base, and double amino acid mutation) and to 5'-AC̆C ACT-3': (thr thr)(a single base, single amino acid mutation).[21] Both of these mutants are catalytically inactive. (The serine residue at position 70 is known to be involved in catalysis[24-29] and all lactamases of this family have a conserved ser-thr dyad at residues 70-71.[29]) In these cases mutants were recovered at the level of several mutants per thousand colonies. Though seemingly a low yield, this represents a few mutant colonies on each Petri dish and oligonucleotide screening at the DNA level is sufficiently rapid and reliable that at least several mutants can be easily recovered from each successful mutagenesis. (We have also obtained the fourth possible combination of serine-threonine residues in this region, (-ser70-ser71-); this protein shows a significant level of catalytic activity and was obtained by spontaneous mutation as discussed later.)

β-Lactamase has recently proved a popular subject for studies of mutagenesis. In an elegant use of oligonucleotide directed mutagenesis, a primer was used to create a ser→arg substitution at the second position in the leader sequence.[30] In this "piggyback" procedure, the original template had an ochre codon TAA at residue 4; the oligonucleotide primer corrected this sequence to CAA (glu), thereby allowing protein expression, and also introduced the mutation of ser→arg at the nearby site; at the same time a new Eco RI site was introduced in this region which facilitated screening for mutants. Another mutant of β-lactamase, thiolactamase, with a cysteine residue at position 70 was prepared by a specific reconstruction of the DNA sequence.[31] This approach depended on the location of a restriction site in this region of the genome. In fact, the ser 70→cys mutant has lost this restriction site, which was useful in identifying mutants. The thiolactamase shows catalytic activity.

As noted earlier, single-strand templates serve very effectively as subjects for oligonucleotide directed mutagenesis, and generally give a higher yield of mutants than do double-stranded plasmids. This approach has been used in creating a substitution mutant, with cysteine in place of serine, in the ATP binding site of tyrosyl tRNA synthetase.[19] This substitution was achieved by mutating the structural gene attached to M13 and was accompanied by an unusually efficient protein synthesis allowing easy isolation of the mutant protein. The major effect of the ser→cys substitution was manifest in Km for binding ATP which was raised from 0.9 to 4.1 mM; Vmax was reduced from 1.4 to 0.9 sec^{-1}.

Oligonucleotide directed mutagenesis has also been used to study[22] the role of positively charged amino acids in the amino-terminal region of the outer membrane prolipoprotein of E. coli. Substitutions that reduce or eliminate the positive character of the amino terminus of the signal peptide substantially alter the processing and secretion of the prolipoprotein.

b. Segment Directed Mutagenesis - The techniques so far described focus on the creation of a mutant protein with a predetermined structure; such approaches are useful to learn the properties and functions of a protein with a prespecified structure. An alternate approach is to generate, by nonspecific mutagenesis, a large number of structural variants, to select those with particular properties and thus to determine the structures of those variants that manifest the desired properties. Such random mutagenesis can be allowed to take place anywhere in the structural gene of interest or, to focus more closely on the role of a particular domain of the protein, can be restricted to a region, such as the leader sequence or a loop of a catalytic domain.

Several methods are available to restrict the action of a non-specific mutagen to a particular segment of DNA. In general they depend on the in vitro generation of a single stranded gap or loop in the desired region of a double stranded DNA molecule, followed by a mutagenesis technique that operates only on single stranded DNA.

i. Deletions Target a Region of DNA for Mutagenesis - To target a particular region of DNA for subsequent mutagenesis, deletion mutants have proved useful; they are relatively easy to obtain and have been extensively used in diverse genetic studies. In deletion loop mutagenesis[32,33] the wild-type DNA and the deletion mutant DNA, both cloned in a bacterial plasmid in the same manner, are used to form circular heteroduplexes that contain a looped-out single stranded region of DNA from the wild-type

strand; this looped-out single stranded region corresponds to the segment that was absent in the deletion mutant (see Fig. 2a).

A gapped heteroduplex can also be constructed by annealing two linear molecules derived from the same plasmid if the shorter of the two is open at the site of the deletion.[34] Again, the deletion defines the region of DNA to be mutagenized (see Fig. 2b).

ii. Specific Nicks Target Regions of DNA for Mutagenesis - In a related approach, single strand gaps can be introduced into a double stranded DNA molecule (see Fig. 2c). In one procedure, the site is specified by a "nick"[35,36] introduced at a site recognized by a specific restriction endonuclease.[37] An alternate approach used a single strand of a suitable restriction fragment in the presence of Rec A protein to generate a displacement loop that is then susceptible to the action of S1 (single strand specific) nuclease[38] (see Fig. 2d). The "nick" introduced by these procedures can now be converted to short, single stranded gaps by limited

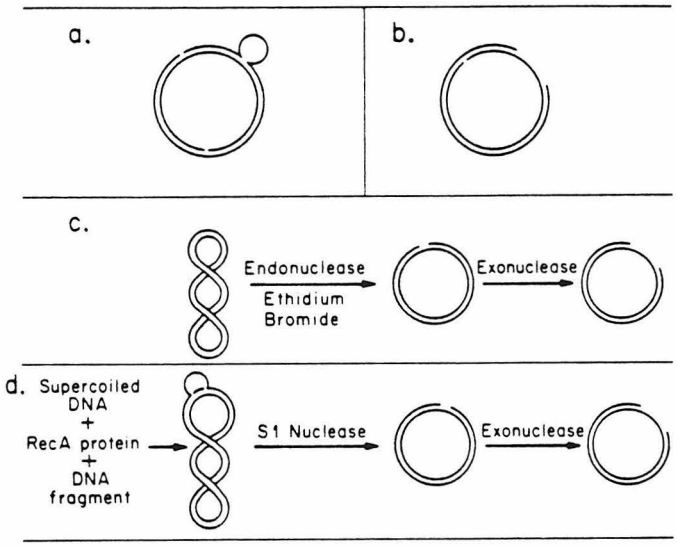


FIGURE 2: Specific segments of DNA can be made single stranded in various ways. These single stranded regions then can be subjected to the action of a single strand specific mutagen (a,b,c,d) or used to introduce mutations during in vitro enzymatic DNA synthesis (b,c,d).

a) Deletion loop

b) Deletion gap

c) Nicking at a specific site followed by gap expansion.

d) Nicking in a specific region followed by gap expansion.

digestion with an exonuclease;[38] alternately, specifically nicked molecules can serve as substrates for nick translation.[39]

After formation of single stranded gaps by the types of methods just discussed, a variety of techniques are available to mutagenize the single stranded regions. For example treatment with sodium bisulfite selectively converts cytosine to uracil residues in the single stranded DNA but leaves intact cytosine residues in the double stranded DNA. This procedure generates single or multiple G-C to A-T transitions.

Procedures are also available for misincorporation of non-complementary nucleotides during enzymatic repair of the single stranded region.[40,41] For example, an error prone DNA polymerase has been used[41] to incorporate mismatched nucleotides at a specific site on the $\phi\chi174$ genome. The site on the single strand phage template was specified by a restriction fragment used to prime *in vitro* DNA synthesis. When the DNA sequence in the region of interest is known, as in this case, base substitutions can be introduced at a particular site; the method can also be used to generate a "library" of mutations directed to a given region of the DNA template.

One can also achieve misincorporation simply by omitting one of the four deoxynucleotide triphosphates normally present during enzymatic gap repair; under certain conditions, misincorporation of the three remaining nucleotides occurs at roughly equal frequency. Use of excess DNA ligase traps the misincorporated base.[40]

α-Thionucleotides[42,43] have been shown to be incorporated into DNA in lieu of the normal nucleotides by DNA polymerase I of *E. coli*.[44] These thionucleotide analogues prevent normal proofreading during *in vitro* DNA synthesis with both *E. coli* DNA polymerase I and T4 DNA polymerase;[45] they also prevent digestion by *E. coli* exonuclease III.[46] These analogues have also been used for misincorporation of non-complementary, excision resistant bases into circular DNA molecules that contained single stranded gaps in specific regions.

Both types of misincorporation reactions are reported to be efficient. The nucleotide omission method, however, produces substitutions that are not predictable *a priori* and may introduce multiple base substitutions when the single stranded segment of DNA being "misrepaired" is longer than a few bases.[40]

Again, β-lactamase has served as the subject of many of these approaches. Two different single stranded fragments of the DNA of the β-lactamase gene were used to direct the formation of D-loop structures in the presence of Rec A protein;[38] subsequent digestion with S1 nuclease produced single stranded nicks distributed throughout the length of the single stranded D-loops. The nicks were then enlarged to gaps and these then mutagenized with bisulfite and repaired enzymatically. Mutants were selected by screening for loss of ampicillin resistance; virtually all such mutants have base substitutions in the regions spanned by either one or the other of the two single stranded fragments used to direct mutagenesis; they were a 369 base pair Taq I fragment that includes the initial 135 base pairs (45-N-terminal amino acids) of the β-lactamase structural gene and a 128 base pair PvuI-PstI fragment from the center of the gene.

c. Random Mutagenesis - The mutagenic strategies discussed above can be usefully complemented by approaches involving random mutagenesis. For example, the construction of a catalytically inactive mutant enzyme by site-specific mutagenesis provides a unique opportunity to study the structural requirements for reactivity.[21] Such an inactive mutant, produced for example by a change in one of the residues at the active site that is essential for catalytic activity, can serve as a subject for testing what changes, if any, will lead to a restoration of activity. Such changes could either be at the active site itself or might, more interestingly, occur elsewhere in the protein in a fashion that compensates for the original, inactivating mutation. Such compensatory sequence changes at a secondary site, should they be observed, would provide infor-

mation on residues that can play a key role in catalysis or in protein folding and may be distant, in terms of primary sequence, from the site of the original mutation. (The mutations at positions 211 and 175 of the α subunit of tryptophan synthetase constitute a well-known example of second site reversion: the substitution of glycine 211 by glutamate abolishes catalytic activity, but activity is restored by a tyrosine to cysteine mutation at position 175.)[47]

Such an approach has been applied to the double mutant of β-lactamase (ser 70→thr; thr 71→ser).[21] The normal level of misincorporation of bases, in the absence of techniques to enhance mutagenesis, in the replication of plasmid DNA in E. coli is of the order of $1/10^6$ at any locus. Accordingly, one might expect to recover catalytically active revertants of an inactive double mutant that have changes at a single amino acid residue; changes in two or more sites are very much less likely; thus phenotypic screening for catalytically active revertants is unlikely to be overwhelmed by reversions to the original wild-type enzyme. When applied to the inactive double mutant of β-lactamase, this procedure gave rise to an active revertant (thr71→ser) that had recovered the normal serine residue at 70, apparently essential for catalytic activity, but retained the mutant serine (in place of threonine) at residue 71.

Random point mutations, both spontaneous and induced by mutagenic agents can be useful in identifying sites essential for function, and also for the isolation of large numbers of nonsense mutations as in the case of the lac repressor protein.[48] Random mutagenesis is also well suited to the generation of proteins with altered characteristics when selective pressures are applied to mutagenized populations of cells; this approach produced a mutant β-lactamase with an increased affinity for cephalosporins.[49]

d. Suppression of Nonsense Mutations - Nonsense mutations can be obtained in large numbers at many different sites or at a specific site or region by adaptation for this purpose of any one of a number of mutagenic strategies, including any of those described above. Proteins with altered amino acid sequences can then be obtained when the mutant gene is expressed in an appropriate suppressor strain.[50] In this way a single amber mutant can be used to introduce different amino acid substitutions by using different suppressor strains. As an example of this approach, many mutants of the lac repressor protein have been produced[48-50] and characterized[50,51] yielding important insights into the nature of protein-DNA interactions.

Major Restructuring - Today's ability to manipulate the DNA of structural genes allows one not only to create proteins that are relatively minor variants of the parent; it also allows one to create almost totally redesigned structures. For example, to establish the relationship between structural domains and the contribution of these domains to function, various segments in the structural gene can be progressively deleted. Such an approach has been used[52] to delete from the structural gene for alanyl tRNA synthetase segments that code for the carboxyl terminal region of the protein.

Another approach to the elimination of C terminal regions involved the use of chain terminating mutations;[53,54] the objective in this case was to study the role of the carboxyterminus in the processing and secretion of β-lactamase into the periplasmic space. Drastic alteration in the sequence in a region of a protein, without altering the length of the polypeptide chain, has also been accomplished in this system by the introduction of double frameshift mutations.[55]

Other major restructurings involve use of recombinant techniques to produce fused or hybrid proteins so that, for example, the products of exogenous genes can be expressed, and even sometimes secreted, by bacterial cells. Examples abound, but two involving β-lactamase may serve to illustrate the strategy. In one, the gene for rat proinsulin was fused to the gene for the leader sequence of β-lactamase with consequent transport of the proinsulin into the periplasmic space of E. coli.[56,57] In a second example, the gene for β-lactamase, devoid of its own leader sequence, was fused to the leader sequence of α-amylase from B. amyloliquefasiens;[58] in consequence, β-lactamase from gram negative E. coli was efficiently secreted by gram positive B. subtilis.

Conclusion - As we hope this review demonstrates, many techniques have recently been developed that allow a variety of quite novel approaches to studies of the relationship between the linear sequence and the function of proteins. Much of the discussion has focused on procedures for manipulating the DNA to achieve proteins with altered function. Largely because it is still early days, few specific cases have been discussed in which insights have been gained into the functional roles of particular structural domains, or even single residues of proteins. However, toward solution of these problems, the new methodologies allow, for the first time, a truly systematic approach; their application should now proceed rapidly and allow one to address such intriguing questions as: What are the essential structural features for a protein to function effectively as a catalyst, hormone, or receptor? How are the various structural aspects of an antibody related to its ability to distinguish self from non-self and thereby, for example, to initiate the destruction of infectious agents such as bacteria and viruses, or the altered cells that form tumors? What regions of a protein located on the surface of a cell carry the information essential to the interactions between cells that play a central role in the development from a single cell (the fertilized egg) of a complex organism such as man? From these studies will come striking new insights into the mechanisms by which proteins successfully carry out their myriad functions as well as the ability to design proteins with specific, novel, and useful properties.

References

1. R. Crea, A. Kraszewski, T. Hirose and K. Itakura, Proc.Natl.Acad.Sci.USA, 75, 5765 (1978).
2. K. Itakura, T. Hirose, R. Crea and A.D. Riggs, Science, 198, 1056 (1977).
3. H.G. Khorana, Science, 203, 614 (1979).
4. M.D. Edge, A.R. Greene, G.R. Heathcliffe, P.A. Meacock, W. Schuch, D.B. Scanlon, T.C. Atkinson, C.R. Newton and A.F. Markham, Nature, 292, 756 (1981).
5. D. Shortle, D. DiMaio and D. Nathans, Ann.Rev.Genet., 15, 265 (1981).
6. C. Weissman, S. Nagata, T. Taniguichi, H. Weber and F. Meyer in "Genetic Engineering, Principles and Methods," Vol. 1, J.K. Setlow and A. Hollander, Eds., Plenum, New York, N.Y., 1979, p. 133.
7. D. Shortle, J. Pipas, S. Lazarowitz, D. DiMaio and D. Nathans in "Genetic Engineering, Principles and Methods," Vol. 1, J.K. Setlow and A. Hollander, Eds., Plenum, New York N.Y., 1979, p. 73.
8. M. Smith and S. Gillam in "Genetic Engineering, Principles and Methods," Vol. 3, J.K. Setlow and A. Hollander, Eds., Plenum, New York, N.Y., 1979, p. 1.
9. A. Razin, T. Hirose, K. Itakura and A.D. Riggs, Proc.Natl.Acad.Sci.USA, 75, 2170 (1978).
10. C.A. Hutchison, S. Phillips, M.H. Edgell, S. Gillam, P. Jahnke and M. Smith, J.Biol. Chem., 253, 6551 (1978).
11. O.S. Bhanot, S.A. Khan and R.W. Chambers, J.Biol.Chem., 254, 12684 (1979).
12. S. Gillam, R.C. Astell and M. Smith, Gene, 12, 129 (1980).
13. S. Gillam, P. Jahnke, C. Astell, S. Phillips, C.A. Hutchinson and M. Smith, Nucleic Acids Res., 6, 2973 (1979).
14. S. Gillam and M. Smith, Gene, 8, 81 (1979).
15. S. Gillam and M. Smith, Gene, 8, 99 (1979).

16. G.F.M. Simons, G.H. Veeneman, R.N.H. Konings, J.H. van Boom and J.G.G. Schoenmakers, Nucleic Acids Res., 10, 821 (1982).
17. I. Kudo, M. Leineweber and U.L. RajBhandary, Proc.Natl.Acad.Sci.USA, 78, 4753 (1981).
18. C. Montell, E.F. Fisher, M.H. Caruthers and A.J. Berk, Nature, 295, 380 (1978).
19. G. Winter, A.R. Fersht, A.J. Wilkinson, M. Zoller and M. Smith, Nature, 299, 756 (1982).
20. R.B. Wallace, P.F. Johnson, S. Tanaka, M. Schoeld, K. Itakura and J. Abelson, Science, 209, 1396 (1980).
21. G. Dalbadie-McFarland, L.W. Cohen, A.D. Riggs, C. Morin, K. Itakura and J.H. Richards, Proc.Natl.Acad.Sci.USA, 79, 6409 (1982).
22. S. Inouye, X. Soberon, T. Franceschini, K. Nakamura, K. Itakura and M. Inouye, Proc. Natl.Acad.Sci.USA, 79, 3438 (1982).
23. R.B. Wallace, J. Shaffer, R.F. Murphy, J. Bonner, T. Hirose and K. Itakura, Nucleic Acids Res., 6, 3543 (1979).
24. R.F. Pratt and M.J. Loosemore, Proc.Natl.Acad.Sci.USA, 75, 4145 (1978).
25. V. Knott-Hunziker, S.G. Waley, B.S. Orlek and P.G. Sammes, FEBS Lett., 99, 59 (1979).
26. V. Knott-Hunziker, B.S. Orlek, P.G. Sammes and S.G. Waley, Biochem.J., 147, 365 (1979).
27. M.J. Loosemore, S.A. Cohen and R.F. Pratt, Biochemistry, 19, 3990 (1980).
28. S.A. Cohen and R.F. Pratt, Biochemistry, 19, 3995 (1980).
29. R.P. Ambler, Phil.Trans.R.Soc.London Ser.B, 289, 321 (1980).
30. A.D. Charles, A.E. Gautier, M.D. Edge and J.R. Knowles, JBC, 257, 7930 (1982).
31. I.S. Sigal, B.G. Harwood and R. Arentzen, Proc.Natl.Acad.Sci.USA, 79, 7157 (1982).
32. D. Kalderon, B.A. Oostra, B.K. Ely and E.A. Smith, Nucleic Acids Res., 10, 5161 (1982).
33. K.W.C. Peden and D. Nathans, Proc.Natl.Acad.Sci.USA, 79, 7214 (1982).
34. A. Oka, K. Sugimoto, H. Sasaki and M. Takanami, Gene, 19, 59 (1982).
35. L. Greenfield, L. Simpson and D. Kaplan, Biochimica et Biophysica Acta, 407, 365 (1975).
36. R.C. Parker, R.M. Watson and J. Vinograd, Proc.Natl.Acad.Sci.USA, 74, 851 (1977).
37. D. Shortle and D. Nathans, Proc.Natl.Acad.Sci.USA, 75, 2170 (1978).
38. D. Shortle, D. Koshland, G.M. Weinstock and D. Botstein, Proc.Natl.Acad.Sci.USA, 79, 1588 (1982).
39. W. Muller, H. Weber, F. Meyer and C. Weissmann, J.Mol.Biol., 124, 343 (1978).
40. D. Shortle, P. Grisafi, S.J. Benkovic and D. Botstein, Proc.Natl.Acad.Sci.USA, 79, 1588 (1982).
41. R.A. Zakour and L.A. Loeb, Nature, 295, 708 (1982).
42. F. Eckstein and R.S. Goody, Biochemistry, 15 1685 (1976).
43. F.R. Bryant and S.J. Benkovic, Biochemistry, 18, 2825 (1979).
44. H-P Vosberg and F. Eckstein, Biochemistry, 16, 3633 (1977).
45. T.A. Kunkel, F. Eckstein, A.S. Mildvan, R.M. Koplitz and L.A. Loeb, Proc.Natl.Sci.USA, 78, 6734 (1981).
46. S.D. Putney, S.J. Benkovic and P.R. Schimmel, Proc.Natl.Acad.Sci.USA, 78, 7350 (1981).
47. D.R. Helinski and C. Yanofsky, J.Biol.Chem., 238, 1043 (1963).
48. J.H. Miller, D.G. Ganem, P. Lu and A. Schmitz, J.Mol.Biol., 109, 275 (1977).
49. A. Hall and J.R. Knowles, Nature, 264, 803 (1976).
50. J.H. Miller, C. Coulondre, M. Hofer, V. Schmeissner, H. Sommer, A. Schmitz and P. Lu, J.Mol.Biol., 131, 191 (1979).
51. K. Arndt, H. Nick, F. Boschelli, P. Lu and J. Sadler, J.Mol.Biol. (in press).
52. M. Jasin and P. Schimmel, Cold Spring Harbor meeting on In Vitro Mutagenesis, May 12-16, 1982.
53. D. Koshland and D. Botstein, Cell, 20, 749 (1980).
54. D. Koshland and D. Botstein, Cill, 30, 893 (1982).
55. D. Koshland, R.T. Sauer and D. Botstein, Cell, 30, 903 (1982).
56. K. Talmadge, S. Stahl and W. Gilbert, Proc.Natl.Acad.Sci.USA, 77, 3369 (1980).
57. K. Talmadge, J. Kaufman and W. Gilbert, Proc.Natl.Acad.Sci.USA, 77, 3988 (1980).
58. I. Palva, M. Sarvas, P. Lehtovaara, M. Sibakov and L. Kaarianen, Proc.Natl.Acad.Sci. USA, 79, 5582 (1982).

Since the publication of this review article, the application of *in vitro* mutagenesis techniques to problems in enzymology has aquired momentum as researchers interested in the chemistry and biochemistry of proteins and peptides have realized the great potential benefits of applying the experimental tools of molecular biology to the study of diverse problems in protein biochemistry.

The nature of the problems being investigated by site specific mutagenesis alone, ranges from studying the role of specific residues in catalysis, as in the case of trypsin and chymotrypsin (59), dihydrofolate reductase (60), tyrosyl tRNA synthetase (19, 61), and beta lactamase (62, 63, and this work), to the role of disulfide bridges in protein stability (60, 64, and this work), and the structural requirements for functional leader sequences in the secretion of proteins such as beta lactamase (65) and the outer membrane lipoprotein of E. coli (66). Until very recently these problems would have been approached by chemically modifying specific residues, by determination of various physical and chemical properties and the application of these same methods to the study of naturally occurring variants or variants obtained by random mutagenesis. What the methods of specific *in vitro* mutagenesis add to the former approach are the necessary tools to carry out a systematic study of protein variants, introducing an element of control which maximizes the effectivenes of physi-

cal and chemical methods by allowing one to precisely substitute functional groups within a protein in a rational manner.

The increasing efficiency with which long segments of DNA can be synthesized will make the construction of genes, rather than the cloning of natural genes for proteins of interest, the method of choice in many cases, since synthetic genes can be designed with restriction sites and structural features that facilitate the introduction of sequence changes as well as the expression of the gene products. This, as well as the many other applications of synthetic DNA in molecular biology, is another good example of one discipline, chemical synthesis in this case, providing a systematic solution to technical problems in another.

REFERENCES

59. Craik, C. S., Largeman, C., Fletcher, T., Roczniak, S., Barr, P., Sletterick, R. and Rutter, W., Submitted to Nature (1984).

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

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

62. Sigal, I. S, Harwood, B. G. and Arentzen, R., Proc. Natl. Acad. Sci. USA 79, 7157-7160 (1982).

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

64. Perry, L. J. and Wetzel, R., Science, in press (1984).

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

CHAPTER I

OLIGONUCLEOTIDE-DIRECTED MUTAGENESIS AS A GENERAL AND POWERFUL METHOD FOR STUDIES OF PROTEIN FUNCTION

Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 6409–6413, November 1982
Biochemistry

Oligonucleotide-directed mutagenesis as a general and powerful method for studies of protein function

(β-lactamase/plasmid pBR322/enzyme mechanisms/colony screening/protein structure–function)

G. DALBADIE-MCFARLAND*, L. W. COHEN†, A. D. RIGGS‡, C. MORIN‡§, K. ITAKURA‡, AND J. H. RICHARDS*

*Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, †Seaver Laboratory of Biology, Pomona College, Claremont, California 91711, and ‡City of Hope Research Institute, Duarte, California 91010

Communicated by Norman Davidson, June 25, 1982

ABSTRACT **We have used oligonucleotide-directed mutagenesis to make a specific change in the β-lactamase (EC 3.5.2.6) (ampicillin resistance) gene of the plasmid pBR322. Evidence suggests that the active site for this enzyme may include a serine-threonine dyad (residues 70 and 71). By priming *in vitro* DNA synthesis with a chemically synthesized 16-base oligodeoxyribonucleotide, we have inverted the Ser-Thr dyad to Thr-Ser and thereby generated a mutant with an ampicillin-sensitive phenotype. This "double-mismatch" method is relatively simple and also very general because detection of mutants is at the level of DNA and involves only colony hybridization. Accordingly, the procedure can be applied to any DNA sequence and does not depend on the phenotype of the mutant.**

Rational study of the influence of amino acid sequence on the three-dimensional structure and function of a protein would benefit greatly from the ability to change specific residues. When the gene for the protein has been cloned and is expressed in an appropriate vector, site-directed mutagenesis (1, 2) provides a method for creating new proteins whose structural and functional characteristics may be of practical value or may offer significant mechanistic insights.

Oligonucleotide-directed mutagenesis, the most specific form of directed mutagenesis, has been used to produce specific base changes in the single-stranded phage φX174 (3, 4). The method (Fig. 1) involves priming *in vitro* DNA synthesis with a chemically synthesized oligodeoxynucleotide that carries at least a single base mismatch with the complementary strand of the "wild-type" DNA. Because DNA polymerases require a double-stranded segment for initiation of DNA replication, the synthetic oligodeoxynucleotide primes DNA synthesis and is, itself, incorporated into the resulting heteroduplex molecule. After molecular cloning, semiconservative *in vivo* replication of this heteroduplex gives rise to homoduplexes whose sequences are either that of the original wild-type DNA or that of the synthetic oligodeoxynucleotide. With single-stranded circular DNAs, priming with oligodeoxynucleotides has been used to cause transitions, transversions, and deletions, in some cases very efficiently (5–11). Wallace *et al.* (12, 13) have extended the technique to double-stranded circular DNAs, and even though the frequency of directed mutagenesis may be lower, screening of colonies with the ^{32}P-labeled synthetic oligodeoxynucleotide allows easy detection of the desired mutant (12, 13).

The particular virtue of oligonucleotide-directed mutagenesis for structure–function studies lies in allowing one to produce mutant proteins with very specific changes in particular resi-

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.

dues, which, for example, may be directly involved in catalysis or in determining substrate specificity. In this way, one can test ideas about the roles of particular amino acid side chains. Other procedures are valuable for producing random or less specific changes as, for example, in the creation of a mutant of β-lactamase (EC 3.5.2.6) with enhanced activity toward cephalosporins (14). Interesting mutants of β-lactamase have also been generated by segment-directed mutagenesis (15, 16) or by misrepair of gaps placed at restriction sites (17).

β-Lactamase provides a particularly favorable system for developing the technique of oligonucleotide-directed mutagenesis and for demonstrating the general feasibility of this approach to the study of protein function. The gene for β-lactamase is contained in the common double-stranded cloning vehicle plasmid pBR322, which has a total of only 4,362 base pairs and which contains a second selectable marker for resistance to tetracycline (18). The complete nucleotide sequence of the β-lactamase gene has been determined (19), and when pBR322 is grown in *Escherichia coli*, the enzyme is expressed and secreted into the periplasmic space, from which it can be readily isolated (20).

β-Lactamase catalyzes the hydrolysis of the amide bond of the lactam ring of penicillins and related antibiotics; the catalytic pathway includes an acyl-enzyme intermediate (20). A serine residue probably participates in catalysis (21–26) and is part of a conserved Ser-Thr dyad (26). [These residues have been numbered Ser-70 and Thr-71 (26).] Presumably the hydroxyl group of Ser-70 adds nucleophilically to the carbonyl group of the β-lactam ring in a mechanism somewhat analogous to that of serine proteases. The role of Thr-71 is less clear, but it seems essential for catalytic activity because a mutant, probably with isoleucine at this position, shows no catalytic activity (26).

In this work, whose main focus was on the procedures for generating and identifying site-specific mutants, we used a double-mismatch method to invert the order of the dyad from Ser-Thr to Thr-Ser (i.e., a double mutation: Ser-70 → Thr and Thr-71 → Ser). The mutant bacteria show no β-lactamase activity, which suggests a functional defect in the mutant enzyme.

MATERIALS AND METHODS

DNAs and Transformation. Plasmid pBR322 (18) was grown in *E. coli* strain LS1 (*pro, leu, thi, rpsL20, hsdR, hsdM, ara-14, galK2, xyl-5, mtlA, supE44*), a derivative of HB101 (27). Plasmid DNA was prepared according to the method of Birnboim and Doly (28) and supercoils were further purified by chromatography on acridine yellow ED gel (Boehringer Mannheim) as

Abbreviations: NaCl/EDTA/Tris, 0.15 M NaCl/1 mM EDTA/0.03 M Tris·HCl, pH 8.0; NaCl/Cit, 0.15 M NaCl/0.015 M sodium citrate, pH 7.0.

§ Present address: Laboratoire de Chimie Organique, Institut Pasteur, Paris, France.

described by Vincent and Goldstein (29) or, alternatively, by ethidium bromide centrifugation followed by gel filtration (30). It probably is important that the plasmid DNA be free of oligodeoxynucleotides that often contaminate DNA preparations. For secondary transformations, where *in vitro* DNA synthesis is not involved, the rapid plasmid DNA isolation procedure of Ish-Horowicz and Burke (31) is convenient. Competent cells were prepared and transformations were done according to Hanahan (32). Selection was for resistance to tetracycline at 20 μg/ml.

The 16-base oligodeoxynucleotide 5′-A-A-T-G-A-T-G-A-C-C-T-C-T-T-T-T-3′ was synthesized from dinucleotides by the triester method on a solid-phase support (33).

Enzymes. *E. coli* exonuclease III, *E. coli* DNA polymerase 1 (large fragment), phage T4 ligase, and the restriction enzyme *Hpa* II were obtained from Bethesda Research Laboratories. T4 polynucleotide kinase was obtained from Boehringer Mannheim.

Preparation of Nicked DNA. Fifteen micrograms of pure supercoiled pBR322 DNA and 9.0 μg of ethidium bromide in 150 μl of 20 mM Tris·HCl, pH 7.4/1 mM dithiothreitol/7 mM $MgCl_2$ were treated in the dark at room temperature with 7.5 units of *Hpa* II for a time estimated to produce about 75% nicked circles and 25% linear molecules (about 1 hr). The time and concentrations should be determined empirically on a smaller scale with the same reagents. The reaction was assayed by electrophoresis on 1.2% agarose in 40 mM Tris acetate, pH 7.8/10 mM EDTA (conditions that separate linear, circular, and supercoiled DNA molecules). We used *Hpa* II for nicking because 52 *Hpa* II sites are in pBR322 (one of them is near the Ser-70 region). Moreover, *Hpa* II is appropriately inhibited by ethidium bromide, in contrast to some other restriction enzymes (e.g., *Pst* I). DNase 1 could also have been used (13), because the precise location of the nick probably is not critical.

The reaction was stopped by adding 0.1 vol of 0.1 M EDTA and the solution was extracted twice with phenol, once with chloroform, and then twice with diethyl ether. These extractions removed both protein and ethidium bromide. The DNA was precipitated by adding 0.1 vol of 3 M ammonium acetate and 2 vol of ethanol. After centrifugation, the pellet was washed

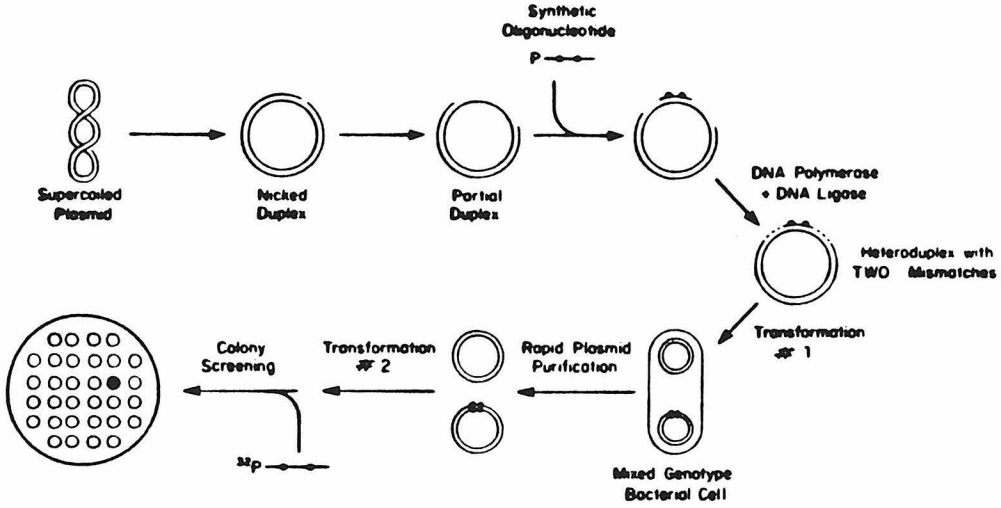


FIG. 1. Overall scheme for oligonucleotide-directed mutagenesis of double-stranded circular plasmid DNA. Supercoiled plasmid circles are nicked in one strand and rendered partially single-stranded by treatment with exonuclease. The gapped circles are hybridized with a homologous oligodeoxynucleotide carrying, by design, some mismatches. *In vitro* DNA synthesis, primed in part by the oligodeoxynucleotide, leads to heteroduplex plasmid circles. Molecular cloning and *in vivo* DNA replication generates homoduplexes, some of which have the DNA sequence of the primer oligodeoxynucleotide. Colony screening, with the same oligodeoxynucleotide labeled with ^{32}P as a hybridization probe, allows identification of the desired mutant colony regardless of its phenotype.

with 70% (vol/vol) ethanol, dried, and dissolved in 15 μl of sterile water.

Exonuclease Treatment, Heteroduplex Formation, and Transformation. About 1 μg (0.4 pmol) of the nicked DNA was treated for 15 min at room temperature with 6 units of *E. coli* exonuclease III in 50 mM Tris·HCl, pH 8.0/5 mM $MgCl_2$/10 mM 2-mercaptoethanol (total volume, 3 μl). The enzyme was inactivated by heating to 60°C in the presence of chloroform, which was then removed by evaporation under reduced pressure. The nicked and exonuclease-treated DNA was heated to 65°C for 5 min in 0.5 M NaCl in the presence of 80 ng of the synthetic oligonucleotide (18 pmol) in a 5-μl volume and allowed to stand at room temperature for 1 hr.

Buffer and salt concentrations were adjusted to 100 mM NaCl, 20 mM Tris·HCl at pH 7.5, 10 mM dithiothreitol, 10 mM $MgCl_2$, each of the four deoxynucleoside triphosphates at about 400 μM, and 500 μM ATP. Sixteen units of *E. coli* DNA polymerase I (large fragment) and 40 units of T4 DNA ligase were added to yield a final reaction volume of 50 μl. Four microliters of this reaction mixture was used directly to transform competent *E. coli* LS1. A second cycle of transformation was usually done after a quick plasmid preparation (31) from an overnight 10-ml broth culture containing tetracycline at 20 μg/ml and inoculated with the original transformation mix.

Colony Screening by Primer-Probe Hybridization. Gergen *et al.* (34) described a microtiter dish colony collection procedure and a procedure for making filter paper replicas of the colony collection. By following this procedure, a colony collection was made, transferred to agar with a 96-prong replicator, and grown for 18 hr. The colonies were "lifted" onto Whatman 541 filter paper for amplification on 250/μg/ml chloramphenicol plates for 24 hr, followed by alkali lysis and neutralization (34). The filters prepared in this way were then prehybridized for 3 hr at 65°C in 6× NaCl/EDTA/Tris, 0.5% Nonidet P-40, and denatured, sonicated salmon sperm DNA at 100 μg/ml (1× NaCl/EDTA/Tris is 0.15 M NaCl/1 mM EDTA/0.03 M Tris·HCl, pH 8.0). Hybridization with ^{32}P-labeled oligonucleotide was done at room temperature for 16–18 hr in 6× NaCl/EDTA/Tris, 0.5% Nonidet P-40, 250 μg of yeast tRNA per ml, and 1.6 ng of ^{32}P-labeled oligodeoxynucleotide per ml (at

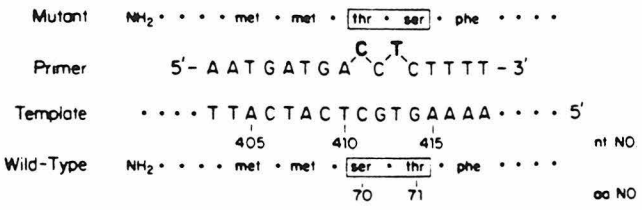


FIG. 2. Nucleotide sequence of the synthetic oligodeoxynucleotide used as a specific mutagen and of the complementary wild-type strand. The amino acids encoded by both mutant and wild-type sequences are shown. nt NO is the nucleotide number in pBR322 (19). aa NO is the amino acid number in β-lactamase (26).

least 10^7 cpm/μg). The filters were washed three times in 25–50 ml per filter of 6× NaCl/Cit (1× NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0) for 5 min at room temperature and autoradiographed. The temperature of the hybridization and washing is important and will vary with the nucleotide sequence and length. We have followed empirical rules developed by Suggs *et al.* (35) for oligodeoxynucleotide hybridizations on Southern blots and colony screens. One estimates the apparent melting point (t_d) to be the sum of 2°C per A·T base pair and 4°C for each G·C base pair when the hybridizations and washes are done in 6× NaCl/EDTA/Tris or 6× NaCl/Cit. We do the hybridizations at least 5°C below the calculated t_d and do a series of washes and autoradiography, first at 5°C and then at higher temperatures to ensure passing through the temperature that best distinguishes mutant from wild type.

RESULTS

Fig. 1 illustrates our use of a 16-base synthetic oligodeoxynucleotide with two deliberate base pair mismatches to prime *in*

vitro DNA synthesis on pBR322 DNA containing single-stranded regions that had been generated by exonuclease treatment of nicked, double-stranded circles. The two mismatches in the primer d(A-A-T-G-A-T-G-A-C-C-T-C-T-T-T-T-T) at the bases denoted (*) will accomplish the double mutation Ser-70 → Thr and Thr-71 → Ser (Fig. 2). For generality, the method is designed for screening at the DNA level rather than depending on the phenotype of the mutant protein. Therefore, we have used colony hybridization with a ^{32}P-labeled oligodeoxynucleotide to select mutants. Conveniently, the same oligodeoxynucleotide used as a primer in the *in vitro* synthesis can be labeled with ^{32}P and also used as a hybridization probe. Those colonies containing the mutant gene will match perfectly with the synthetic probe and therefore hybridize more strongly than will wild-type colonies, which have two base pair mismatches.

Fig. 3 shows the result of a colony screen; a mutant can be clearly identified. In this experiment, colonies were picked and distributed into microtiter dishes to provide an ordered array and a permanent library (34). After growth, a filter replica of the colonies was made and used for colony screening (13, 34).

The microtiter dish technique allowed convenient screening of a few thousand colonies and was used to isolate the first mutant (22G9). However, when it became apparent that the frequency of the desired mutation was only about one in a thousand, we began to use a random colony screening procedure, which allows the rapid evaluation of tens of thousands of colonies. Fig. 4 shows the results of such a screen: the bottom shows

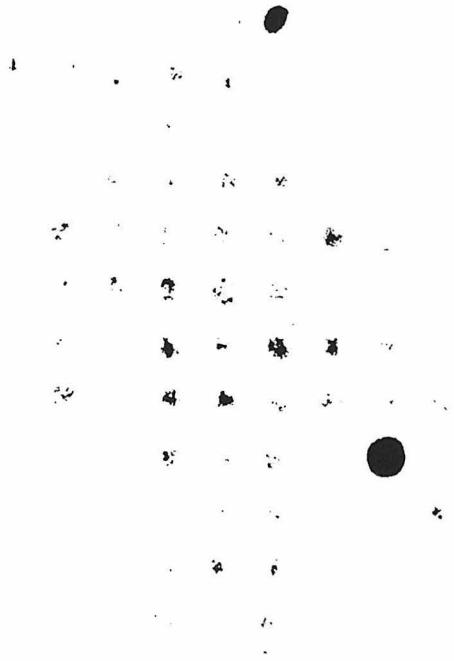


FIG. 3. Autoradiogram showing the result of a colony screen. Colonies from a second transformation (see Fig. 1) were picked and grown in a microtiter dish to provide a convenient permanent library. Filter paper replicas of the library, prepared as described in text and ref. 34, were hybridized with the ^{32}P-labeled oligodeoxynucleotide, washed, and autoradiographed. The darkest colony, designated 22G9, proved to be the desired mutant.

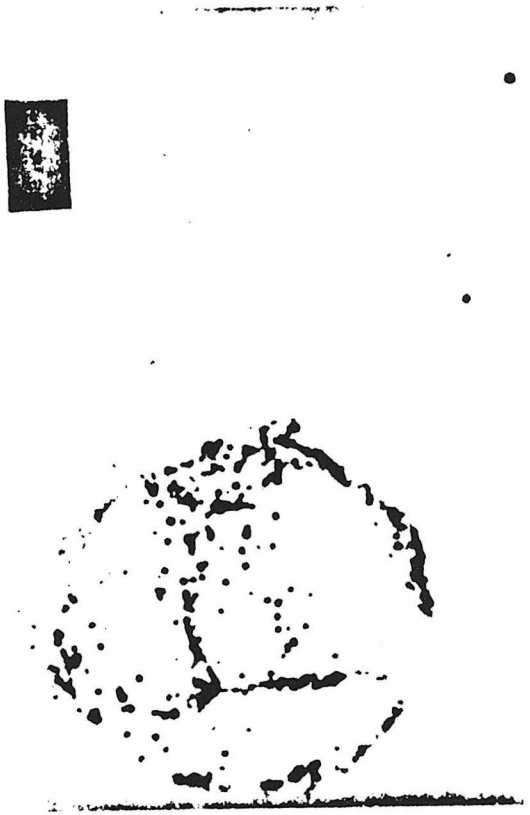


FIG. 4. Autoradiogram showing the results of a colony screen of randomly spread colonies. The upper circle shows cells from a second transformation spread randomly on the surface of 15-cm L agar plates. About 2,000 tetracycline-sensitive colonies were expected from preliminary experiments. After overnight growth, when the colonies were about 1 mm in diameter, they were transferred to 541 filter papers and then treated as described in the text and ref. 34. The lower circle is a control of randomly spread cells containing a perfect match (mutant 22G9), grown and treated as described above.

a control with random colonies of a known perfect match (mutant 22G9), the top shows a similar screening of about 1,800 colonies resulting from an independent *in vitro* mutagenesis experiment. Wild-type colonies do give a weak signal, but mutants, two of which can be seen, stand out clearly. The background and positive signal is somewhat variable, and, for this reason, oligodeoxynucleotides of at least 15 bases probably should be used for this type of screening, even though shorter oligodeoxynucleotides may serve for the priming step and for Southern blotting (7, 8). A 19-base probe gives a very strong and reliable signal (13). Differences of two bases between mutant and wild-type greatly increase the ability to identify mutants unambiguously, even though, in most cases, single-base mismatches probably can be identified (13).

Even with the double-mismatch method used in this work, false positives occasionally were observed, especially when unordered colonies were screened. Random dark spots, probably from [^{32}P]ATP or [^{32}P]orthophosphate, also occur and often cannot be clearly distinguished from mutant colonies. The number of these spots can be reduced significantly by filtering the solution for hybridization just before use; nevertheless, a second screening of plasmid DNA from a putative mutant is essential. Fortunately, this can be accomplished easily by a rapid preparation of DNA (31) followed by agarose gel electrophoresis and hybridization (36). The standard Southern transfer to nitrocellulose (36) can be used, but we have found it even more convenient to simply dry the agarose gel and use the resulting agarose membrane directly for hybridization with the ^{32}P-labeled oligodeoxynucleotide (S. G. S. Tsao, C. F. Brunk, and R. Pearlman, personal communication). Fig. 5 illustrates the unambiguous discrimination between wild-type and mutant colonies. This analysis also discriminates single base mismatches (ref. 13 and unpublished data) and should show differences between single- and double-base mismatches, possibly because any such mutants were not picked at the earlier colony screening step

(the faint colony in Fig. 3 has not yet been analyzed).

To show that the expected two-base change has actually been made, the sequence of the DNA in the region of the active site (Ser-70 and Thr-71) has been determined by the Maxam–Gilbert technique (37) and the sequence in one strand has been confirmed by the dideoxy method (38) after subcloning in M13mp7.

Although mutant colonies can be identified by colony hybridization after the first transformation with heteroduplex plasmid DNA, these colonies probably contain a mixture of wild-type and mutant plasmids because the colonies grow in ampicillin. After a rapid preparation of plasmid from these "mixed" colonies and a second transformation, 30–50% of the resulting colonies hybridize strongly to the ^{32}P-labeled oligodeoxynucleotide, all of these are sensitive to ampicillin.

We have obtained mutants in three independent repetitions of the procedure. All show similar characteristics of hybridization and ampicillin sensitivity. Several thousand apparently wild-type colonies also were screened and all of these were resistant to ampicillin. Thus, the procedure does not produce, at a detectable frequency, random mutants that have functionally inactive β-lactamase. A fortuitous connection between the demonstrated DNA sequence changes in the mutants and the ampicillin-sensitive phenotype seems remote.

DISCUSSION

The procedure described here reproducibly yields mutants that hybridize much more strongly than wild-type colonies with the synthetic oligodeoxynucleotide probe that is used first as the primer for *in vitro* synthesis of heteroduplex plasmid and then as a probe for identifying mutants. The plasmid DNA of one mutant has been characterized in the region of the mutation and shown to have the desired sequence.

All of the mutants that show strong hybridization with the synthetic probe are also sensitive to ampicillin. The inversion

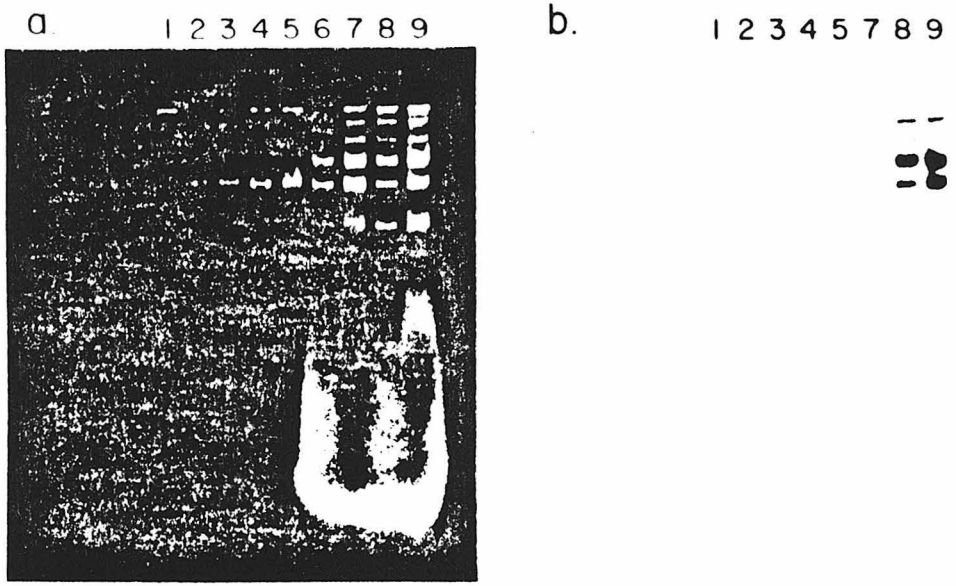


FIG. 5. Hybridization of ^{32}P-labeled oligodeoxynucleotide with gel-purified plasmid DNAs. Plasmid DNAs, prepared from 7-ml overnight unamplified cultures as described by Ish-Horowicz and Burke (31), were each digested with *Eco*RI in a reaction volume of 65 μl. Aliquots were electrophoresed in 1.2% agarose. The gel was stained with ethidium bromide and photographed. Then it was treated for 20 min at room temperature with 0.15 M NaCl and 0.5 M NaOH and neutralized in 0.15 M NaCl and 0.5 M Tris·HCl, pH 7.4. The gel was dried in a gel dryer (Bio-Rad) and the resulting membrane was hybridized overnight at room temperature in 6× NaCl/EDTA/Tris, 0.5% Nonidet P-40, 250 μg of tRNA per ml, and 1.6 ng of ^{32}P-labeled oligodeoxynucleotide per ml. No prehybridization is necessary. The gel membrane was then washed three times for 10 min each at 4°C in 6× NaCl/Cit followed by three washes for 5 min each at 37°C in 6× NaCl/Cit. (*a*) Ethidium bromide-stained gel. Lane 1, 0.4 μg of pBR322 standard; lanes 2–5, 0.1, 0.2, 0.5, and 1.0 μg of *Eco*RI-cut pBR322; lanes 6 and 7, 10 and 30 μl of *Eco*RI digest of clone 38G3, respectively; lanes 8 and 9, 10 and 30 μl of *Eco*RI digest of clone 22G9, respectively. (*b*) Autoradiogram of the dried gel after hybridization.

of Ser-Thr to Thr-Ser was expected to affect catalytic activity because, for example, the substitution of the primary alcohol of Ser-70 by the secondary alcohol of Thr might sterically hinder nucleophilic attack by the hydroxyl group; the amino acid changes also might cause conformational rearrangement within the active site. Particularly interesting in this respect will be studies on the ability of the mutant protein to bind penicillins and related substrates, as distinct from its ability to act as a catalyst. Also of interest will be the effects of the single changes Ser-70 → Thr and Thr-71 → Ser. These mutants might have arisen by partial repair during the mutagenesis procedure, but so far we have isolated no such single-base-change mutants. Although preliminary studies on cell extracts suggest that the mutant protein is catalytically inactive, other reasons for the absence of ampicillin resistance, such as more rapid proteolytic degradation of the mutant protein, remain a formal possibility.

The present yield of mutants, about one per thousand, should be raised significantly by further improvements. However, a major point is that improvements are not required because of the rapidity and reliability of the screening procedure (13). The screening procedure depends only on DNA sequence and not on any phenotypic difference between mutant and wild type. This feature allows easy detection of mutants even with phenotypically silent changes. For example, even though the Ser-70 → Thr and Thr-71 → Ser mutant of β-lactamase created in this work is functionally inactive, one could as easily have identified a mutant that retains normal activity, has acquired enhanced catalytic effectiveness, or displays altered specificities for various substrates. The use of double-base mismatches to facilitate identification of mutants, even when one of these base changes causes no amino acid change, should be emphasized. The degeneracy of the genetic code makes the double-mismatch approach applicable even for single amino acid changes. These features of the method, together with the ability to create specific mutations at any region of the genome, should encourage its widespread application.

The generation of a specifically restructured mutant protein generally will require six steps: (*i*) cloning of the relevant gene in an appropriate vector, (*ii*) expression of the cloned gene, (*iii*) determination of the DNA sequence of the gene, (*iv*) chemical synthesis of an oligodeoxynucleotide, (*v*) oligonucleotide-directed *in vitro* mutagenesis, and (*vi*) identification of mutant colonies. This work has focused on steps *v* and *vi*, which, together with present and rapidly improving techniques for accomplishing steps *i*–*iv*, should allow the restructuring of proteins to become a common undertaking and thereby increase our understanding of the relationship between primary structure and biological function. Moreover, such techniques also will allow the synthesis of new proteins with useful properties not hitherto available.

Note Added in Proof. Using this technique, we have generated a one-base mutant, Ser-70 → Thr [d(. . .A-T-G-A-T-G-A-Ċ-C-A-C-T. . .)], which is catalytically inactive. We have also obtained Thr-71 → Ser [d(. . .A-T-G-A-T-G-A-G-C-Ṫ-C-T. . .)], which has a low level of activity.

1. Shortle, D. & Nathans, D. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2170–2174.
2. Mueller, W., Weber, H., Meyer, F. & Weissmann, C. (1978) *J. Mol. Biol.* **124**, 343–358.
3. Razin, A., Hirose, T., Itakura, K. & Riggs, A. D. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4268–4270.
4. Hutchison, C. A., Phillips, S., Edgell, M. H., Gillam, S., Jahnke, P. & Smith, M. (1978) *J. Biol. Chem.* **253**, 6551–6560.
5. Gillam, S., Astell, R. C. & Smith, M. (1980) *Gene* **12**, 129–137.
6. Gillam, S., Jahnke, P., Astell, C., Phillips, S., Hutchison, C. A. & Smith, M. (1979) *Nucleic Acids Res.* **6**, 2973–2985.
7. Gillam, S. & Smith, M. (1979) *Gene* **8**, 81–97.
8. Gillam, S. & Smith, M. (1979) *Gene* **8**, 99–106.
9. Montell, C., Fisher, E. F., Caruthers, M. H. & Berk, A. J. (1982) *Nature (London)* **295**, 380–384.
10. Kudo, I., Leineweber, M. & RajBhandary, U. L. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4753–4757.
11. Simons, G. F. M., Veeneman, G. H., Konings, R. N. H., van Boom, J. H. & Schoenmakers, J. G. G. (1982) *Nucleic Acids Res.* **10**, 821–832.
12. Wallace, R. B., Johnson, P. F., Tanaka, S., Schöld, M., Itakura, K. & Ableson, J. (1980) *Science* **209**, 1396–1400.
13. Wallace, R. B., Schöld, M., Johnson, M. J., Dembek, P. & Itakura, K. (1981) *Nucleic Acids Res.* **9**, 3647–3656.
14. Hall, A. & Knowles, J. R. (1976) *Nature (London)* **264**, 803–804.
15. Shortle, D., Koshland, D., Weinstock, G. M. & Botstein, D. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5375–5379.
16. Koshland, D. & Botstein, D. (1980) *Cell* **20**, 749–760.
17. Shortle, D., Grisafi, P., Benkovic, S. J. & Botstein, D. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1588–1592.
18. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L. & Boyer, H. W. (1977) *Gene* **2**, 95–113.
19. Sutcliffe, J. G. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3737–3741.
20. Fisher, J., Belasco, J. G., Khosla, S. & Knowles, J. R. (1980) *Biochemistry* **19**, 2895–2901.
21. Pratt, R. F. & Loosemore, M. J. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4145–4149.
22. Knott-Hunziker, V., Waley, S. G., Orlek, B. S. & Sammes, P. G. (1979) *FEBS Lett.* **99**, 59–61.
23. Knott-Hunziker, V., Orlek, B. S., Sammes, P. G. & Waley, S. G. (1979) *Biochem. J.* **147**, 365–367.
24. Loosemore, M. J., Cohen, S. A. & Pratt, R. F. (1980) *Biochemistry* **19**, 3990–3995.
25. Cohen, S. A. & Pratt, R. F. (1980) *Biochemistry* **19**, 3995–4003.
26. Ambler, R. P. (1980) *Phil. Trans. R. Soc. London Ser. B* **289**, 321–331.
27. Boyer, H. W. & Roulland-Dussoix, D. A. (1969) *J. Mol. Biol.* **41**, 459–472.
28. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523.
29. Vincent, S. W. & Goldstein, E. S. (1981) *Anal. Biochem.* **110**, 123–127.
30. Norgard, M. V. (1981) *Anal. Biochem.* **113**, 34–42.
31. Ish-Horowicz, D. & Burke, J. D. (1981) *Nucleic Acids Res.* **9**, 2989–2998.
32. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

33. Ito, H., Ike, Y. & Itakura, K. (1982) *Nucleic Acids Res.* **10**, 1755–1769.
34. Gergen, J. P., Stern, R. H. & Wensink, P. C. (1979) *Nucleic Acids Res.* **7**, 2115–2136.
35. Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K. & Wallace, R. B. (1981) in *Developmental Biology Using Purified Genes*, eds. Brown, D. D. & Fox, C. F. (Academic, New York), pp. 683–693.
36. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
37. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–559.
38. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

We acknowledge many helpful discussions with Norman Davidson, R. Bruce Wallace, and John Rossi. This work was supported by National Institutes of Health Grants GM16424, GM25525, and GM30395. K.I. and A.D.R. are members of the Cancer Research Center (CA16434) at the City of Hope Research Institute. C.M.'s stay was made possible by a grant from the North Atlantic Treaty Organization.

CHAPTER II

SITE SPECIFIC MUTAGENESIS METHODS

SPECIFIC SINGLE BASE SUBSTITUTIONS GENERATE THREE VARIANTS OF BETA-LACTAMASE: ser70-->thr, ser70-->arg and cys77-->ser

Specific Single Base Substitutions Generate Three Variants of Beta-Lactamase: ser70-->thr, ser70-->arg and cys77-->ser

Two additional site specific mutations have been obtained at the active site of beta-lactamase by oligonucleotide directed mutagenesis in the double stranded plasmid pBR322 using exonucleased nicked DNA as a template, and a third one at another site using an M13 vector as a template. The mutations were designed to introduce threonine and arginine at position 70 and to remove one of the only two cysteines in the enzyme (cys 77), replacing it with serine, thereby removing the single disulfide bridge. In this Chapter, the construction of these three mutations will be described and the reasons for introducing the mutation ser 70-->arg will be discussed, but the changes ser70-->thr and cys 77-->ser will be discussed in detail in Chapter IV. Table I summarizes the changes that I have obtained in the active site region of beta-lactamase, and the oligonucleotides used to introduce them.

Ser70-->thr

The construction of the sequences -thr70-ser71- (Chapter I),-ser70-ser71-(Chapter III), and -thr70-thr71- was designed to shed light on the catalytic mechanism of beta-

lactamase and this will be discussed in Chapter IV.

Ser70-->arg

The sequence change to arg 70 was designed for other purposes, although based on existing knowledge of the catalytic machanism of the enzyme. (This is a project done in collaboration with Patricia L. Foster and David Botstein.)

It has been shown recently that some powerful carcinogens have as a common feature the ability to introduce G-C to T-A transversions (1,2). Whereas there are methods of identifying the mutagenic specificity of chemical mutagens by using the genetic system developed by Miller et. al. (3), it would be advantageous to have a simple direct assay to detect specific, induced base substitutions. Developing a tester strain for G-C to T-A transversions, for example, would allow the identification of other mutagens that induce this type of base change; it would also allow one to select for bacterial mutations that specifically affect the induction of G-C to T-A transversions, thereby helping understand the nature of the primary DNA lesion(s), and DNA repair events that lead to a particular type of base substitution. In principle tester strains could be constructed as described below, to assay for every kind of base change. Such an assay should have the following properties: (a) the induced mutations must give a dominant selectable phenotype, and (b) only the desired base change

TABLE I

MUTATIONS AT THE ACTIVE SITE

OF BETA LACTAMASE

TABLE I

MUTATION	aa sequence	DNA sequence	type and number of mismatches
none (ie. wild type sequences)	met met ser thr phe lys	A ATG ATG AGC ACT TTT	—
ser 70 thr 71 → thr 70 ser 71	met met thr ser phe lys	A ATG ATG ACC TCT TTT	two C/C and T/T
ser 70 → thr	met met thr thr phe lys	A ATG ATG ACC ACT TTT	one C/C
thr 71 → ser	met met ser ser phe lys	A ATG ATG AGC TCT TTT	N/A obtained by other methods
ser 70 → arg	met met arg thr phe lys	ATG ATG CGC ACT TTT	one C/T

must be able to produce the phenotype.

Organisms carrying the gene for beta-lactamase are resistant to penicillins and related antibiotics by virtue of the enzyme's catalytic activity which causes the hydro lysis of beta-lactams. Catalysis is brought about by a nucleophilic attack of the active site serine hydroxyl (4) group on the carbonyl carbon of the beta-lactam ring of penicillins and cephalosporins which results in the formation of an acyl-enzyme intermediate (5). The only aminoacid substitution at position 70 of beta lactamase known to produce an active enzyme is ser-->cys (6); with this in mind it seemed virtually certain that introducing arginine at that position would produce an inactive enzyme.

Serine is encoded by the codons TCN and $AG^{C}/_{T}$, and arginine by CGN and $AG^{A}/_{G}$. The serine at position 70 of beta-lactamase is encoded by AGC, which can be specifically mutated to CGC by a single base change. Bacteria carrying this mutant beta-lactamase gene could revert to ampicillin resistance by a single G-C to T-A transversion, which would generate a serine codon at position 70. The codons for cysteine are $TG^{T}/_{C}$ so a single C-G to T-A transition could give rise to a cysteine codon at position 70 of beta-lactamase; however, the cysteine enzyme which is known to be active is inhibited by PCMB (6) so it can be clearly distinguished from the serine enzyme that would be produced by a G-C to T-A transversion. In addition, the serine

enzyme is inhibited by borate, whereas the cysteine enzyme is not (7).

Figure II-1 summarizes the possible aminoacid substitutions that can occur by a single base change. None of the other aminoacid substitutions resulting from a single base change of any kind (leucine, proline, and histidine) would produce an active beta-lactamase. Therefore a strain carrying the beta-lactamase mutation ser->arg could be used to identify potential carcinogens capable of producing GC to TA transversions, by comparing the spontaneous rate of reversion to ampicillin resistance to the rate of reversion found in the presence of a given chemical.

Site specific mutagenesis is the ideal tool to develop this tester strain and I have obtained the mutation as described below; Patricia L. Foster is in the process of characterizing the strain and developing the assay for carcinogens.

Cys77-->ser

The E. coli beta lactamase has a single disulfide bond between cys77 and cys123 (8) which can only be reduced in conditions that also denature the protein. The mutation cys 77-->ser77 removes one of the two cysteines present in mature beta-lactamase, thus eliminating the possibility of forming this disulfide bond. Doing so enables one to study the effect of its absence on the secretion and processing

FIGURE II-1

Mutations that can arise by a single base substitution in the arginine codon CGC.

A SINGLE BASE CHANGE

IN THE arg CODON **CGC**

GIVES RISE TO:

*	**A**GC (**ser**)	C**A**C (his)	CG**A** (arg)
#	**T**GC (**cys**)	C**T**C (leu)	CG**T** (arg)
	GGC (gly)	C**C**C (pro)	CG**G** (arg)

* G-C to T-A transversion : Amp^R

\# G-C to A-T transition : low Amp^R, PCMB sensitive

all other single base changes : Amp^S

of the protein and the effect on the stability of the mature enzyme.

The two active site mutations (ser70-->arg and ser70-->thr) were constructed using nicked, exonucleased pBR322 as a template, whereas cys77-->ser was done on a fragment of the beta-lactamase gene subcloned into M13mp8 (9).

Oligonucleotide-directed mutagenesis : Methods

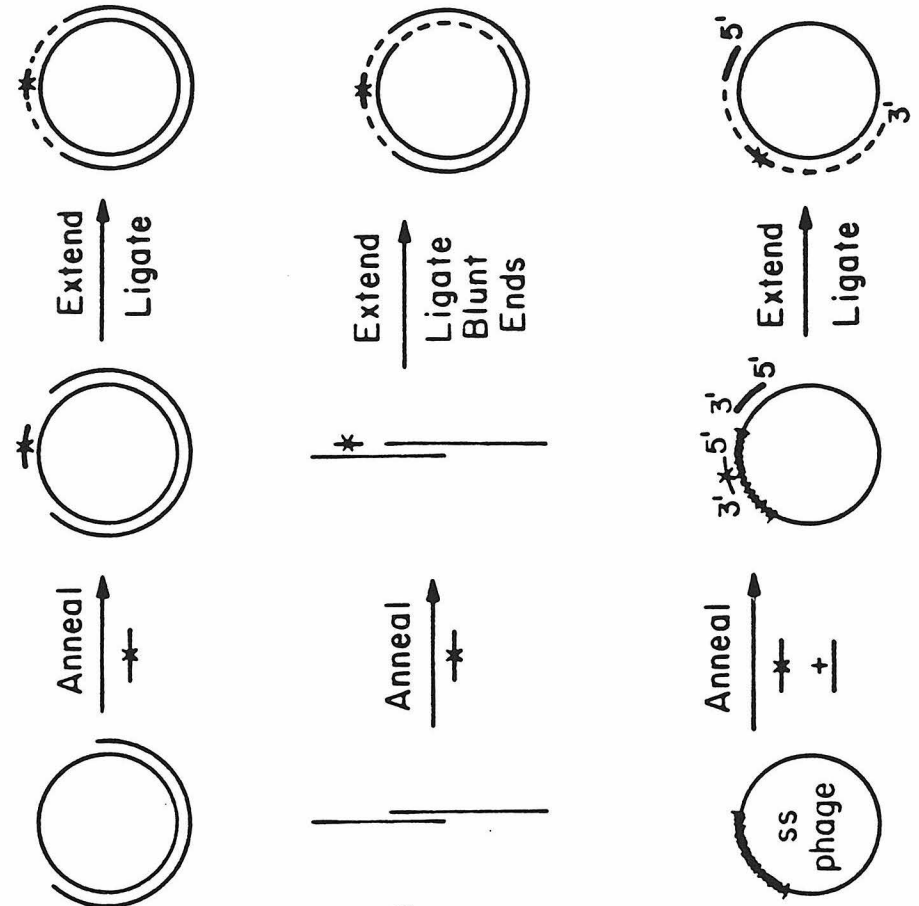
Site specific mutations can be obtained in genes cloned into either single or double stranded vectors by incorporating a synthetic oligonucleotide with the desired base changes. Several different strategies have been used successfully in both types of vector, and so the advantages and disadvantages of each method should be considered in relationship to each individual case. Figure II-2 illustrates the mutagenic strategies that apply in each case.

Oligonucleotide-directed mutagenesis can be used to obtain single (Eg., 10), double (11), and multiple (12) specific base substitutions, as well as deletions and insertions (13) of one or more bases. The following is a description of how to design and obtain a site-specific mutation by oligonucleotide-directed mutagenesis.

Sections (a) through (e) refer to mutagenesis on DNA cloned into a single or a double stranded vector; step (b)

FIGURE II-2

A schematic outline of different experimental strategies for oligonucleotide-directed mutagenesis using:
a) nicked and exonucleased plasmid DNA,
b) linearized and exonucleased plasmid DNA, or
c) single stranded phage DNA with an insert
as templates.


a.
Nick
Exonuclease
Anneal
Extend
Ligate
b.
Linearize
Exonuclease
Anneal
Extend
Ligate
Blunt
Ends
c.
ss
phage
Anneal
Extend
Ligate
3'
5'
5'
3'
5'
3'

is carried out by nature when a single stranded vector such as M13 is used. In the interest of clarity, M13 mutagenesis methods will be summarized at the end of this Methods section.

a. The oligonucleotide : design, synthesis and characterization.

b. Obtaining the template .

c. Repair reaction : enzymatic polymerization and ligation.

d. Transformation.

e. Screening.

a. **Oligonucleotide design, synthesis and characterization.**

1.Length.

The minimum oligonucleotide length required for reliable screening is about fifteen bases. Shorter oligomers (six or seven bases) would function adequately as primers (14) if only a single base substitution were being introduced, but the binding specificity would be much lower and the signal-to-noise ratio in screening would be poor. Fifteen base oligomers can be used for single and double base substitutions, insertions and deletions. However, the length should be increased to obtain more complex mutations. Deletions of several hundred base pairs have been obtained recently using oligomers 17, 21, and 25 bases long; the best results were obtained with the longer oligomers (John Rossi, personal communication, 1983).

2. Sequence.

The site and nature of the specific mutation to be introduced determine the sequence of the oligomer. However, some precautions are necessary to reduce the possibility of unintended mutations. A homology search on both strands of the DNA to be mutagenized is necessary to avoid oligomer sequences that would anneal with high specificity at sites other than the one chosen for the mutation. Sites on the vector are as important as sites in the cloned gene, therefore the choice of vector is not trivial. Whenever possible, a small vector of known sequence should be used.

If complete sequence information on the cloned gene to be mutagenized is not available, then it is preferable to subclone a smaller fragment whose sequence is known, for mutagenesis.

The sequence of the oligomer itself should be examined for self-complementary regions, which should always be avoided. If either of these problems (homology or self-complementary sequences) appear unavoidable, other vectors or other mutagenic strategies should be considered.

When the purpose of mutagenesis is to obtain mutant proteins, there are more alternatives than when the desired final product is a nucleic acid. The degeneracy of the genetic code can be used to advantage in designing mutations in proteins. For example, a single aminoacid substitution can be obtained using an oligonucleotide that in-

troduces two base substitutions. This makes screening for mutants easier, because the difference in oligomer hybridization to mutant as compared to wild-type colonies is greater. It is not necessary to do this, but it can be helpful.

3. Synthesis.

Rapid, solid-phase methods for synthesizing DNA oligomers using either the phosphotriester (15) or the phosphoramidite chemistry (16) have been recently developed ; either method is adequate, and both can and have been automated.

4. Characterization.

It is important to determine the purity of the oligonucleotide after synthesis. The major impurities found will be shorter oligomers, which occur to a varying extent depending on the methods of synthesis and purification. This can be checked easily by electrophoresis on 20% polyacrylamide, 40% urea gels. If impurities are present, they can be removed either by HPLC or by preparative gel electrophoresis.

If desired, the ability of the oligonucleotide to bind specifically at the selected site can be checked by hybridization to a restriction digest, or by priming in vitro DNA synthesis on the appropriate template and looking for a DNA strand of specific length after restriction. Alternatively, the oligonucleotide can be used to prime dideoxy sequencing

reactions to confirm that it is priming specifically at the desired site.

The sequence of the oligonucleotide can be checked by the Maxam and Gilbert method.

b. **Obtaining the template.**

In double stranded plasmids, the template for in vitro DNA synthesis can be obtained in either of two ways. In one case, the template is a partially single stranded circle obtained by nicking supercoiled DNA and treating it with an exonuclease (Figure II-2, a) and in another case, it is linearized and partially exonucleased plasmid DNA (Figure II-2, b).

Nicked circles are obtained by digesting supercoiled plasmid DNA with an endonuclease in the presence of ethidium bromide (17); under certain conditions, most of the DNA can be converted to nicked circles. The enzymatic nicking reaction is simple , but a series of pilot reactions should be done before attempting it on a preparative scale. The products of the reaction vary, depending on the relative concentrations of supercoiled DNA, enzyme, and ethidium bromide present initially. Temperature and length of incubation are also important. In addition, the optimal conditions for nicking are different when different restriction enzymes are used, and contaminants such as linear DNA can interfere by altering the effective ethidium bromide concentration.

The pilot reactions can be done using 1 to 2 units of a restriction endonuclease for each microgram of supercoiled DNA. The DNA concentration can be kept constant at about 30 to 40 micrograms/ml, and the ethidium bromide concentration varied between .01 and .10 mg/ml. The reactions should be stopped quickly after an hour at room temperature by extracting with phenol and chloroform.

The reaction products can be analyzed by gel electrophoresis on 1.2% agarose in 0.04 M Tris-acetate (pH 8.0) buffer; the supercoiled, linear and open circular forms of plasmid DNA are clearly resolved. It is possible to obtain more than 90% nicking, while eliminating virtually all of the supercoiled DNA.

After the optimal conditions are determined, the reaction can be carried out on a preparative scale. It is important that all the concentrations in the preparative reaction as well as the temperature and length of incubation are the same as in the pilot reaction. Phenol chloroform extraction (twice) stops the reaction and removes the ethidium bromide. The DNA is ready to use in the next reaction after precipitating and washing with ethanol. If several mutations are to be introduced in the same plasmid, it is convenient to prepare a large amount of nicked DNA at one time. Two hundred to three hundred micrograms is an ample supply.

In this case the template for in vitro DNA synthesis

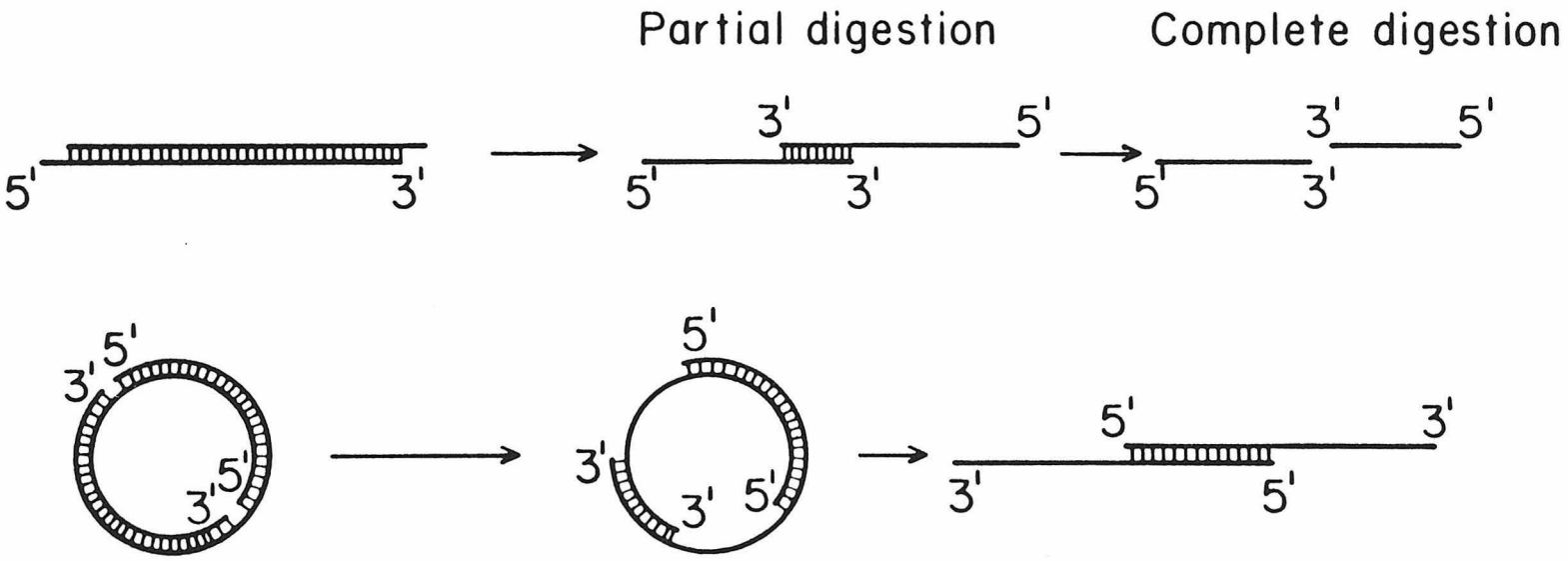
FIGURE II-3

Action of E. coli exonuclease III

a) on linear, double stranded DNA, and

b) on circular plasmid DNA with multiple single strand nicks.

Action of E.coli Exonuclease III



primed by the mutagenic oligomer can be obtained by digesting 5 micrograms of nicked plasmid for fifteen minutes at room temperature with 30 units of E. coli exonucleaseIII in a 50 microliter volume, in the appropriate buffer conditions. The reaction should be stopped by adding a small amount of chloroform and heating to 65°C for ten minutes. It is important to destroy the exonuclease activity to prevent further digestion of the template and digestion of the oligonucleotide in the following step of this procedure.

Presumably at this point, some fraction of the modified plasmid molecules now have single stranded regions which span the site of the desired mutation (Figure II-3). The extent of the exonuclease digestion has not been determined quantitatively in this work, and it is possible that some or all of the intact circular molecules are completely single stranded. However, I should emphasize that a partial digestion is preferable for two reasons:

(a) In order to prevent, or reduce in vivo repair of the mismatch (9, 18), it is necessary to extend the 3' end of the primer, and to ligate its 5' end to the 3' end of the repaired strand. This can be accomplished either by completely resynthesizing the strand primed by the oligonucleotide, on a completely single stranded circular template, or by "filling in" the gaps between the oligonucleotide primer and double stranded regions of the template.

The latter is easier to accomplish, since in actual practice it is sometimes problematic to extend a strand of DNA in vitro over a span of several thousand nucleotides.

(b) The plasmid nicked in the presence of ethidium bromide may contain more than one nick per open circular molecule. If one or more nicks are present on both strands, complete digestion with exonuclease III will render these molecules useless as templates for in vitro DNA synthesis. Some evidence for the presence of at least one nick in each strand comes from alkaline gel electrophoresis (19) of the nicked DNA: it appears that much more than one-half of the DNA runs as a single stranded linear monomer of the plasmid.

If the exonuclease treatment is only partial, then even molecules with multiple nicks could be repaired during enzymatic DNA synthesis as long as the circular nature of the molecule is preserved, thus improving the yield of molecules capable of serving as templates for the mutagenic oligomer.

In the second method of obtaining a template, the plasmid is linearized with a restriction enzyme that cuts the vector DNA near the cloned DNA (20); the distance between this restriction site and the site to be mutated should be less than one-half the entire length of the plasmid. Depending on the position of this restriction site relative to the cloned DNA, the mutagenic oligomer should

be complementary to either one or the other strand. After the DNA is linearized, it is partially digested with E. coli exonuclease III . The extent of digestion should be sufficient to expose the site of the intended mutation, but not so extensive as to eliminate a region of double stranded DNA that holds both strands together.

It is convenient to determine the best digestion conditions by treating aliquots of the linearized plasmid with varying amounts of exonuclease III. After stopping the enzymatic reaction by adding EDTA, the digests can be analyzed by gel electrophoresis on 1.2% agarose; if this gel is dried without heat or any other treatment, the samples that have been sufficiently digested can be identified by hybridization of the labeled mutagenic oligonucleotide to the resulting membrane. The oligonucleotide binds to single stranded regions of the DNA but not to regions that have remained double stranded. The remaining amount of the appropriate digest can then be used as template in the following reaction.

In the third case (Fig II-2 c), template for mutagenesis is obtained by subcloning an appropriate fragment of the gene of interest into a single stranded vector such as one of the M13 cloning vectors and preparing single stranded phage DNA by standard procedures (9).

c. Repair Reaction: polymerization and ligation.

Annealing of the primer and template is done by combining a fifty to one hundred fold molar excess of the synthetic oligonucleotide with 1 to 5 micrograms of exonucleased material from the previous reaction. They are heated to 65°C for one to five minutes, in a 5 microliter volume in either 10mM Tris-HCl pH 7.5, 0.5 M NaCl or 10mM Tris-HCl pH 7.5 0.2 M NaCl, 10 mM MgCl2 , and allowed to remain at room temperature for 45 minutes. Alternatively they can be incubated at 37°C for 45 minutes to avoid the possibility of denaturing the template DNA . Then the rest of the substrates for the repair reaction are added in an appropriate buffer such that the final concentrations are 0.5 mM each of the four deoxynucleotide triphosphates and ATP, 100 mM NaCl, 10 mM $MgCl_2$, 20 mM Tris-HCl pH 7.5, 10 mM DTT (dithiothreitol), and the final volume is 50 microliters. E. coli DNA polymerase (8 units, Klenow fragment, Boehringer-Mannheim) and T4 DNA ligase (400 units, as defined by the supplier, New England Biolabs) are added last. If the oligonucleotide used in this reaction is labeled with ^{32}P, its incorporation into the repaired plasmid can be monitored. After incubation at 25°C for 30 minutes and at 15°C for at least 4 hours, the repaired DNA can be purified by extraction with phenol and chloroform, followed by ethanol precipitation.An aliquot of the redissolved DNA is then used for transformation.

Specificity can be improved by carrying out the repair reaction at 37°C if the T_d is 42°C or higher; this prevents the oligonucleotide from binding at lower affinity sites, and it also helps reduce the formation of secondary structures on the template.

d. Transformation

When using M13 vectors, the standard transformation described by J. Messing (9) is appropriate.

If the template was prepared from a double stranded vector the procedure described by Hanahan (21) is the most appropriate because of its high efficiency at low DNA concentrations which allows one to screen as many independent transformants as necessary. Since the frequency at which mutations are introduced can be low in some cases , it is often necessary to screen several thousand independent transformants to have a high probability of finding a clone carrying the desired mutation. Frequencies range from 5% to 10% to as much as 40% for simple base substitutions on a single stranded phage template , or on a linearized double stranded plasmid template, but can be as low as one in one thousand when using exonucleased nicked circles as a template; for large deletions,it can be as low as one in ten thousand (John Rossi, personal communication, 1983).

For double stranded vectors, after transforming with the repaired DNA and ascertaining that a sufficient number of

independent transformants have been obtained, the transformation mixture can be used to inoculate a 5 ml broth culture. Plasmid DNA is prepared (22) and used in a second transformation in order to segregate the mutant and wild type aleles. The resulting transformants can be plated at an appropriate cell density for colony screening. It is possible to screen for mutants before segregating wild type and mutant plasmids, but the signal obtained when screening mixed genotype cells is not as clear, particularly for single base changes.

With M13 vectors one can plate the transformants and either make filter replicas for screening or pick individual plaques into broth cultures for dot blots, as described below.

e. Screening

When an adequate number of independent transformants have been obtained and plated at an appropriate cell density, filter replicas can be prepared either on Whatman 541 filter paper or on nitrocellulose as described in the literature (23) and in Chapter I. The synthetic oligonucleotide used to direct mutagenesis is now phosphorylated with gamma- ^{32}P-ATP and polynucleotide kinase, and used as a hybridization probe.

Whatman 541 filter replicas are first prehybridized for two hours in 6X SET, 0.5% NP-40 (Sigma), 100 ug/ml

deproteinized, denatured salmon sperm DNA, 1mM sodium orthophosphate, 1mM sodium pyrophosphate, 0.5mM ATP at 65°C for two hours. The filters are then hybridized for two hours at room temperature in 6X SET, 0.5% NP40, 250 ug/ml deproteinized yeast tRNA, 1mM sodium orthophosphate, 1mM sodium pyrophosphate, 0.5 mMATP, and 2 to 4 pmoles/ml of labeled oligonucleotide, filtered through a 0.2 micron filter immediately before using. Washing is done in 6X SSC, first at room temperature (3 times for 5 minutes each time) and an autoradiogram is made. Additional washes are done at higher temperatures, beginning at 5 degrees below the estimated T_d for the oligomer, taking autoradiograms after each wash.

Any positives found are confirmed by picking and re-screening at the colony hybridization level, and subsequently by hybridization to the purified plasmid or phage DNA, as well as by DNA sequencing.

Mutagenesis in M13 Vectors

An alternative to the mutagenesis method described for double-stranded DNA is to subclone the gene of interest or an apropriate fragment into an M13 vector. Single stranded template DNA is then obtained by the standard methods used to prepare template for sequencing with chain terminating inhibitors.

Annealing and in vitro DNA repair is done essentially as

described for double-stranded vectors. There is evidence (24, 25) that the complete synthesis of a second strand of DNA primed by an oligonucleotide, using the circular, single stranded viral template is not always efficient. This means that the 5' end of the mutagenic oligomer is not ligated, and any mismatched bases are more accessible to in vivo repair mechanisms, which results in few or no mutants.

This difficulty can be circumvented by using the universal M13 sequencing primer to "extend" the 5' end of the mutagenic oligomer (9). If both the mutagenic primer and the universal M13 primer are annealed to the phage template, extended by the action of DNA polymerase and ligated, the gap between the two is filled in, thus "sealing" the 5' end of the mutagenic oligomer and reducing the chance of in vivo mismatch repair in molecules that are incompletely synthesized in vitro.

Some authors advocate separating covalently closed circular molecules from totally and partially single stranded molecules by sucrose gradient centrifugation (24). However, this seems unnecessary because even if few mutations are introduced, many phage-infected colonies can be screened by hybridization to filter replicas. Furthermore, if the desired mutation entails an improbable molecular event (e.g. a very large deletion), it might be difficult to obtain enough ccc DNA to purify without doing unnecessary additional work.

The procedure for mutagenesis on M13 is described below.

Procedure:

The methods for growth of phage, and purification of phage DNA or RF DNA for sequencing, mutagenesis and cloning are described in detail by Joachim Messing in Methods in Enzymology, 101:20, 1983.

Oligonucleotide design and characterization is essentially as described in the previous section, although in searching for homologies between primer and template one need only consider the plus strand of the phage vector and the appropriate strand of the cloned DNA.

After subcloning an appropriate segment of the gene to be mutagenized into an M13 vector by standard methods, template DNA is prepared from the single stranded packaged phage by extraction with phenol as described by Messing (9).

Single stranded template DNA (0.5 pmoles) is combined with 20 pmoles of phosphorylated mutagenic oligomer, and 20 pmoles of the M13 universal sequencing primer, in 120 mM HEPES buffer pH 7.5, 6 mM NaCl, 8 mM $MgCl_2$, 10 mM DTT, in a 12 microliter volume. This solution is heated to 90°C for 5 minutes and allowed to cool slowly to room temperature. The synthesis of the second strand is

initiated by adding 3.5 microliters of a solution containing 7.5 nanomoles of each of the four deoxynucleotide triphosphates, and 4 units of the Klenow (large) fragment of E. coli DNA polymerase I (Boehringer-Mannheim). The reaction mixture is incubated at 15°C for 4 hours. The phage T7 gene 32 protein has been recently used to improve polymerase read-through and is reported to increase the yield of covalently closed double stranded circles dramatically (Charles Craik, personal communication, 1984).

After incubating at 15°C for 4 hours, 20 nanomoles of ATP and 400 units (as defined by the supplier, New England Biolabs) of T4 DNA ligase are added and the reaction is continued at 15°C for another 12 hours.

The resulting DNA is purified by extraction with phenol and chloroform, and precipitated with ethanol. The DNA is dried and redissolved in 1/10 TE buffer and used to transform an appropriate host such as E. coli JM103 or JM109.

The colonies carrying mutant phage are identified best by hybridizing the mutagenic oligonucleotide, now labeled with ^{32}P, to nitrocellulose (Shleicher & Schuell, BA85) filter replicas of the plated transformants. Prehybridization is done at room temperature for 30 minutes in a solution containing 0.2% SDS, 10X Denhardt's (22), 0.1 mg/ml sonicated and denatured salmon sperm DNA, 1 mM sodium

orthophosphate, 1mM sodium pyrophosphate, 0.5 mM ATP, and 6X SSC buffer.

Then the nitrocellulose filters are hybridized for at least 2hours at room temperature in a solution containing 0.2% SDS, 10X Denhardt's (22), 0.25 mg/ml yeast tRNA (Sigma, phenol extracted), 1 mM sodium orthophosphate, 1mM sodium pyrophosphate, 0.5 mM ATP, 6X SSC buffer and 1 to 2 ng/ml of the labeled oligonucleotide .

The filters are washed in 6X SSC, three times for five minutes, first at room temperature and then at 5 degrees below the estimated T_d, and at higher temperatures, until a discriminating temperature is found; autoradiograms should be taken after each wash.

Any positives found can be confirmed by dot blots. Several single plaques presumed to be positive are picked into 2 ml of broth each, and incubated for 4 to 6 hours with good aereation. The cells are centrifuged out and 2 microliters of each supernatant are spotted on dry, untreated nitrocellulose and baked for 90 minutes in a vacuum oven at 80°C. The same prehybridization and hybridization conditions described above should be used ; washing is also as above.

After positives are confirmed, the mutant phages should be plaque purified and characterized by DNA sequencing.

RESULTS and DISCUSSION

In Chapter I a double mutation was introduced using a sixteen base oligonucleotide that was complementary to the wild type template in the region of the active site of beta lactamase at 14 out of 16 bases. In that case two mismatches were present of necessity, but we argued that it would be advantageous in site specific mutagenesis of protein genes, to design mutagenic oligomers with double base mismatches even for single aminoacid mutations whenever the degeneracy of the genetic code allows it. Whereas single base substitutions had been previously obtained and identified by colony hybridization (10), and model studies (26) indicated that single base mismatches in short heteroduplexes could be detected by in situ hybridization, the relative merits of using a single versus a double base mismatch in screening for mutants had not been tested.

The two oligonucleotides used to introduce the ser70-->thr70 (a hexadecamer) and ser70-->arg (a pentadecamer) were designed to span the same sequence as the hexadecamer used for the double mutation in Chapter I; these two however, introduce a single base change each. A homology search revealed no competing sites on either strand of the template, pBR322 (27), that match either oligonucleotide at more than 11/15 bases; this is quite acceptable since an-

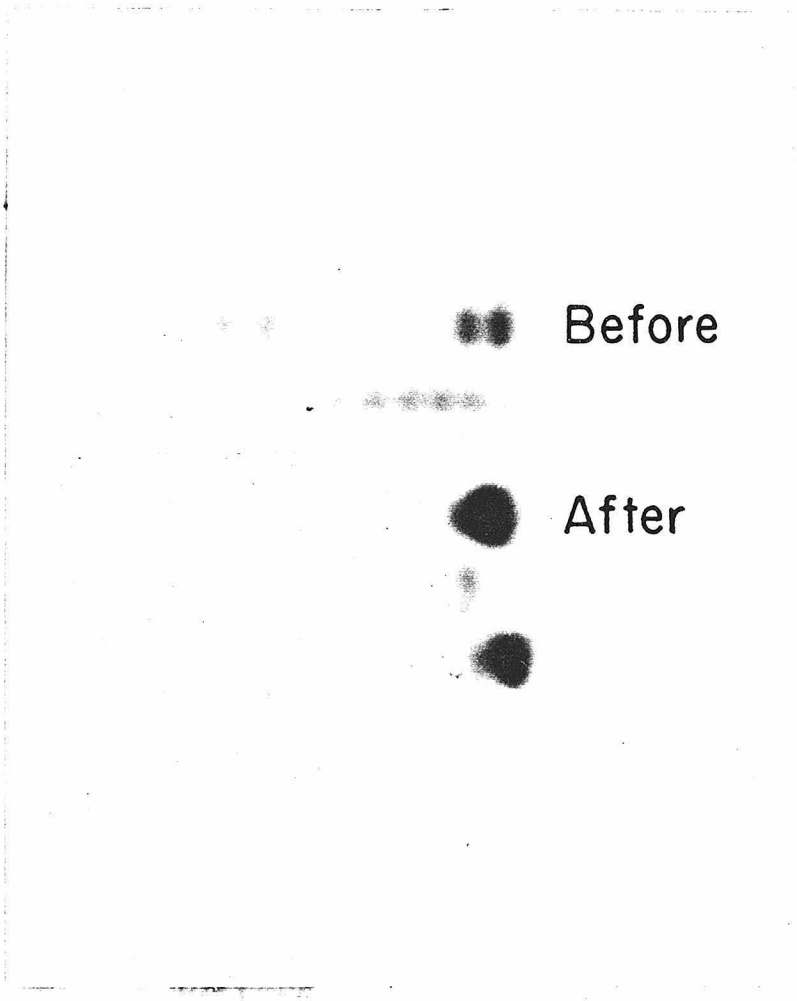
nealing of the mutagenic oligomers at the desired site will occur with 15/16 matches and 14/15 matches respectively. The oligonucleotides 5'ATGATGCGCACTTTT and 5'AATGATGACCACTTTT have virtually no self complementarity. The hexadecamer was synthesized by the phosphotriester method (15) and purified by HPLC in K. Itakura's laboratory at the City of Hope and the pentadecamer was synthesized by the phosphoramidite (16) method by Steve Schultz. The latter was purified by preparative gel electrophoresis through 20% polyacrylamide 40% urea. An autoradiogram of the 5'-^{32}P labeled oligonucleotide before and after purification is shown in Figure II-4; the purity of the HPLC purified oligonucleotides is comparable. No further characterization was done prior to mutagenesis.

Since the beta lactamase gene is carried by the small double stranded plasmid pBR322 (27) and there are no compelling reasons to use a single stranded vector, the same strategy for mutagenesis in double stranded DNA used in Chapter I and described in greater detail at the beginning of this chapter was used again.

Appropriate concentrations of supercoiled plasmid, ethidium bromide and restriction endonuclease HpaII were determined by trial reactions which were analyzed by gel electrophoresis on 1.2% agarose in 0.04 M Tris acetate buffer pH 8.0. Figure II-5 is a photograph of the stained gel. The conditions used in lane (d) were chosen for the

FIGURE II-4

Autoradiogram of a synthetic ^{32}P labeled oligodeoxynucleotide (a pentadecamer) before and after purification by preparative gel electrophoresis, analyzed on a 20% polyacrylamide, 40% urea gel.


Before
After

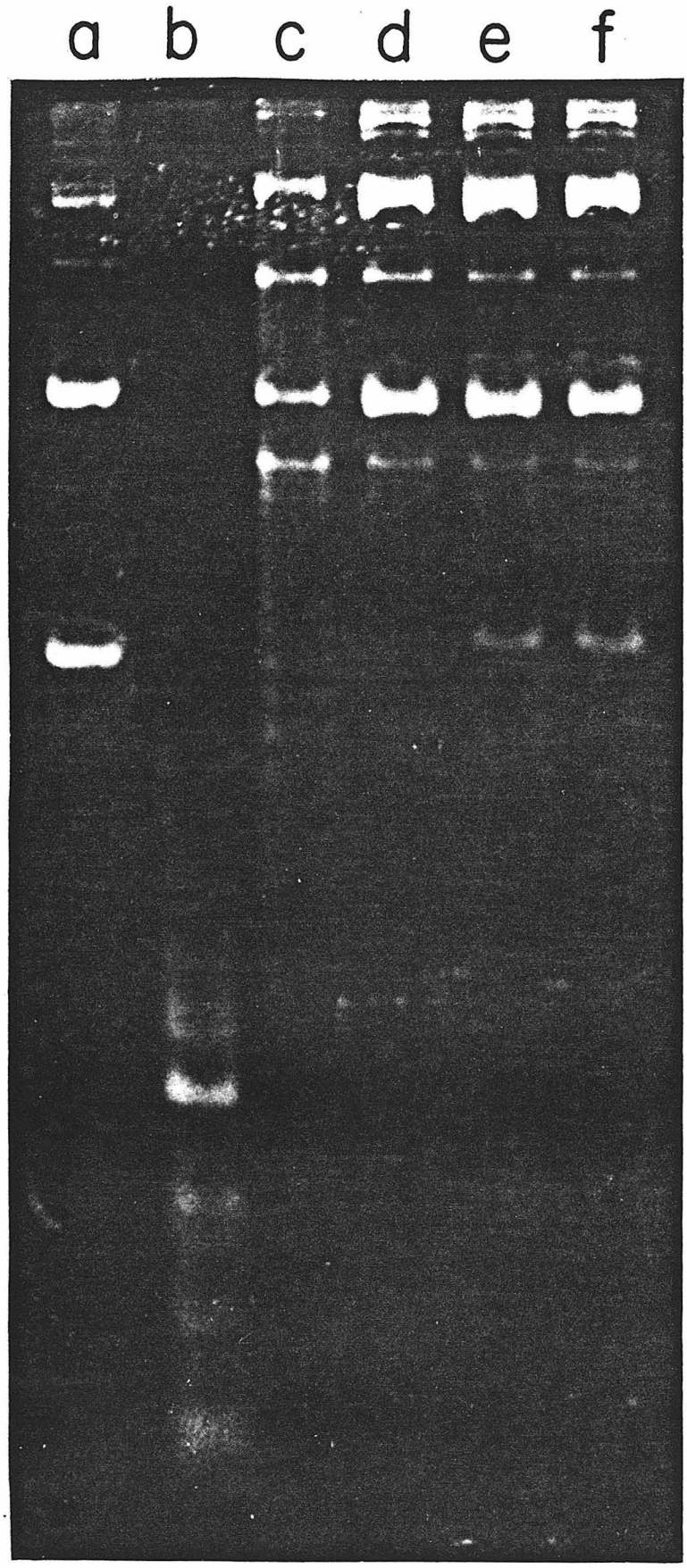
preparative reaction.

Treatment of this nicked DNA with E. coli exonuclease III results in material that runs as a smear on an agarose gel, and on prolonged incubation the amount of DNA diminishes until it is no longer detectable. This can be explained as follows: Initially as shown in Figure II-3, a population of open circular forms of the plasmid with varying amounts of single stranded regions are generated, and their mobilities vary depending on the size of the single stranded regions, so a smear, rather than a band is seen. On prolonged incubation, molecules with multiple nicks are digested extensively, giving rise to a very heterogenious population which would be dispersed throughout the length of the gel. Also, the presence of nonspecific nucleases in commercial preparations cannot be excluded, and they may be responsible for the digestion of any remaining intact single stranded circles.

Since it is cheap and easy to prepare a large supply of nicked plasmid, a simple way to make sure one has enough good template for a successful mutagenesis experiment is to use a large amount of nicked plasmid in each experiment. I have found that increasing the amount used to 5 micrograms is more than adequate. This approach has the further advantage of allowing for electrophoretic analysis after each step. The efficiency of the chemical synthesis methods currently in use contribute to make this feasible since the

FIGURE II-5

Gel electrophoretic analysis of pilot reactions used to determine the best conditions for nicking plasmid DNA. Lane (a) is a control containing 0.37 ug of supercoiled pBR322. Each reaction contained 0.37 ug of plasmid DNA, 20mM Tris-HCl pH 7.4, 7mM $MgCl_2$, 1mM DTT, and increasing amounts of ethidium bromide (0.1, 0.2, 0.4, 0.6, and 0.8 ug) in lanes (b) through (f). The reactions were started by adding 0.45 units of restriction endonuclease Hpa II and incubated at room temperature for one hour; the final reaction volume was 10 ul. They were stopped by adding 1/10 volume of 0.1 M EDTA and chilling on ice.


a
b
c
d
e
f

amount of oligonucleotide required per experiment is only a fraction of the amount routinely synthesized.

Therefore, for each of these two mutations, 5 ug of nicked pBR322 were digested with 30 untis of E. coli exonuclease III (BRL) for 15 minutes at room temperature. The reactions were stopped and a 50 fold molar excess of each primer was annealed (in separate reactions) as described in the methods section of this chapter. In both cases the repair reactions were done first at room temperature, and then at 15°C for 4 hours or overnight.

In the case of the ser70-->thr mutation approximately 4500 independent transformants were screened after the second transformation. Twenty-five colonies appeared to hybridize strongly after washing at 37°C. Plasmid was prepared from seven of these colonies and two of the plasmids retained the bound probe after washing in 6X SSC at 45°C.

One of these clones was streaked, colony purified and the mutation was confirmed by DNA sequencing (Maxam & Gilbert method).

In the case of the ser70-->arg mutation, at least 3000 to 4000 colonies were screened. Twenty-one appeared to be possible mutants, and one of them was confirmed by hybridization to the plasmid on a dry gel.

In this case, screening was done after the first transformation, so the colony identified carried a mixed

population of wild type and mutant plasmids; still, it was identified against a background of colonies carrying a plasmid that differs by a single base.

Plasmid DNA was prepared (22) from this clone and competent E. coli were transformed; 96 single colonies were tested for ampicillin and tetracycline resistance. 34/96 colonies grew on both ampicillin and tetracycline. 62 of the 96 did not grow on ampicillin. The same 62 bound the probe much better than the wild type control. A few of the ampicillin resistant colonies also bound the probe better than the control. When plasmid was made from these colonies it was apparent that it was a dimeric form which probably carried the mutant sequence in one site and the wild type sequence in the other.

This does not represent a problem, since if necessary the dimers could be cut, diluted and religated in order to transform again to obtain a pure strain.

Our collaborator Patricia L. Foster is in the process of characterizing the ser70-->arg mutation and developing the assay for induced G-C to T-A transversions. Her preliminary results indicate that indeed both aflatoxin B1, which has been shown to produce transversions (2) and the alkylating agent methyl methane sulfonate (MMS) induce a high rate of reversion to ampicillin resistance, whereas other mutagens such as nitrosoguanidine do not.

In producing these two mutations, it became immediately

obvious that a single mismatched base is sufficient to produce reliable hybridization data to distinguish wild type from mutant colonies even when mixed genotype colonies are screened; this agrees with a growing number of reports in the literature (eg., 10, 24, 28). The frequency with which mutants were found by this method ranged from 0.1% to about 0.5% of all the colonies screened.

The mutation cys77-->ser was obtained by mutagenesis on M13; the Eco R1 - Pst1 fragment of pBR322 containing the amino terminal part of the Amp^r gene, was subcloned into the phage vector M13mp8, and E. coli JM103 was used as a host (9). The sequence of the mutagenic oligonucleotide was 5'-CCACTTAGCAGAACT-3', and no competing sites that matched at more than 10 out of 15 bases were found. Figure II-6 shows a nitrocellulose filter replica where some positives are apparent. Figure II-7 shows a dot blot hybridization after a room temperature wash (A), and after washing at a discriminating temperature (B). This mutation was confirmed by DNA sequencing by the chain terminating method, using an oligonucleotide synthesized to sequence through this region and the active site region.

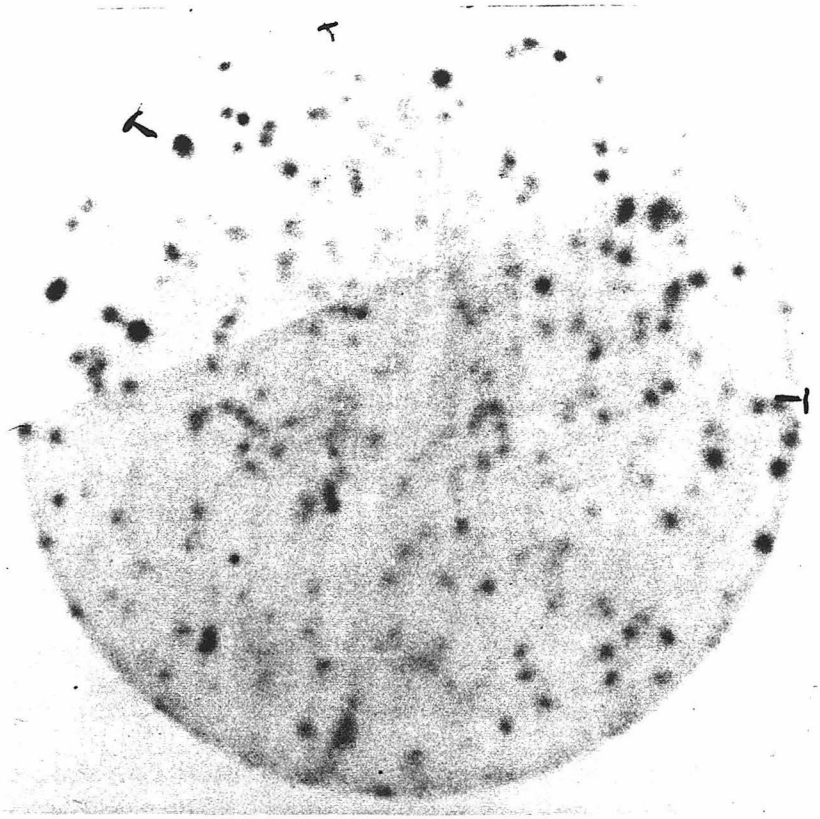
In this case the mutations were obtained at a frequency of 1% to 2% over background, with the added advantage of immediately confirming the mutation by sequencing on the M13 template.

The Eco R1 - Pst1 fragment now carrying the mutation

FIGURE II-6

Autoradiogram of a nitrocellulose filter replica of phage infected colonies washed (A) at room temperature, and (B), at 38°C in 6X SSC, after hybridization to the mutagenic oligodeoxynucleotide 5'-CCACTTAGCAGAACT. The hybridization conditions are described in the Methods section. The arrow indicates one colony that proved to have the desired mutation.

A



B

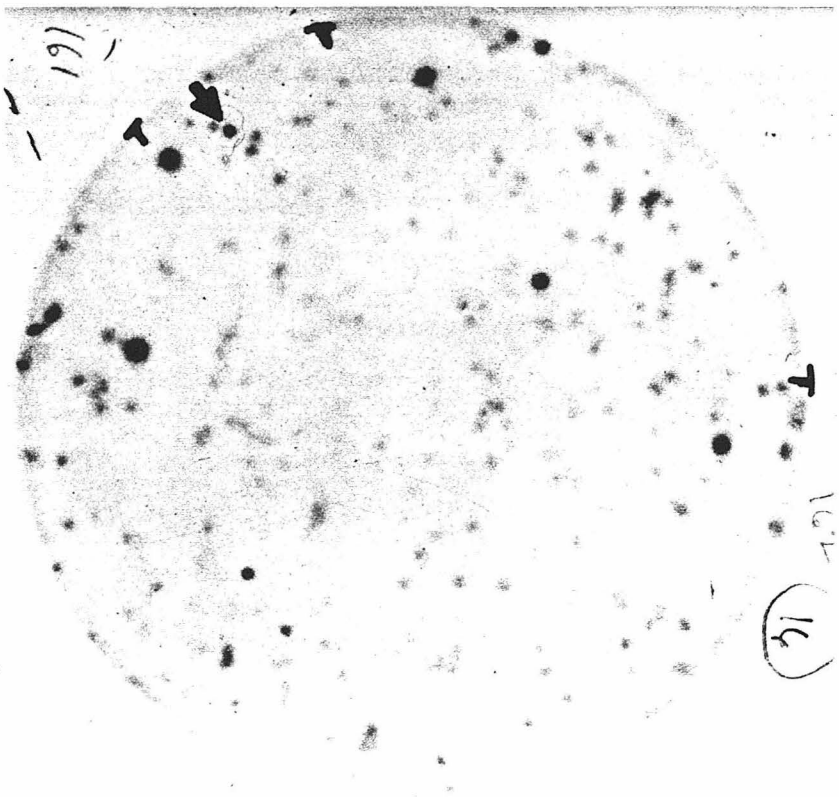
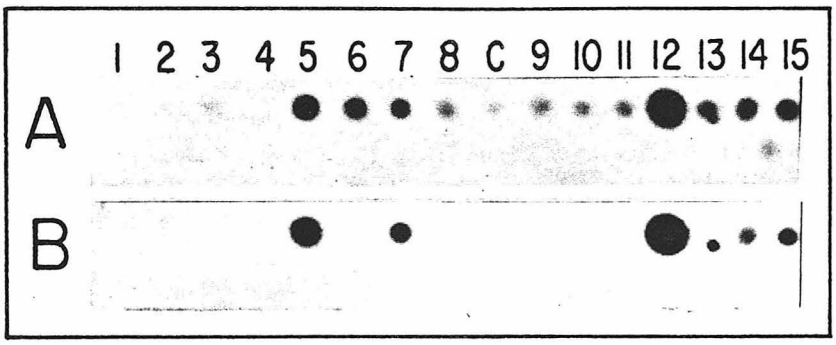


FIGURE II-7

Autoradiogram of a dot blot.
Several phage infected colonies that appeared to be mutants in a previous screening (Figure II-6) were cultured in broth and after centrifugation the phage containing supernatants were spotted on a nitrocellulose membrane as described in Methods. After a two hour hybridization to the mutagenic oligonucleotide, 5'CCACTTAGCAGAACT, at room temperature in the conditions described in the Methods section of this chapter, the membrane was washed at room temperature (A), and at 38°C (B) in 6X SSC.


1 2 3 4 5 6 7 8 C 9 10 11 12 13 14 15
A
B

was subcloned into pBR322 and the cells harboring this plasmid were shown to be ampicillin resistant at 250 ug/ml at both 37°C and 42°C. Plasmid DNA was prepared from these ampicillin resistant colonies and hybridization to the mutagenic oligonucleotide confirmed that the plasmid indeed carries the mutant sequence.

J. J. Neitzel has prepared beta-lactamase from this mutant strain and has made some preliminary observations regarding its thermal stability. These results will be discussed in Chapter IV along with the properties of other mutant proteins.

Table II shows the oligonucleotides used to produce the mutations in this chapter and in Chapter I, their estimated melting temperatures, and the temperature that was found, experimentally, to discriminate between wild type and mutant sequences. There is reasonable agreement with the predicted T_d, and there is also (as expected) a correlation between the difference in melting temperatures of DNA duplexes as a function of the number of mismatched bases and the length of the oligomer.

TABLE II

Oligonucleotide Sequence	T_d	Lowest observed discriminating temperature
AATGATGA**CC**T**C**TTTT (G)A)	42°C	24°C
AATGATGA**CC**ACTTTT (G)	42°C	37°C
ATGATG**C**GCACTTTT (A)	42°C	38°C
CCAC**TT**AGCAGAACT (A)	44°C	38°C
AGCACTTTTA**G**AGT* (A)	38°C	34°C
ACGTTC**C**TCGGGGCG** (T)	52°C	45°C
GGTGAGTA**TT**CAACC** (C)	44°C	40°C
ACATAGCAG**TA**CTTT** (A)	40°C	35°C

The sequences of several oligonucleotides used to introduce mutations in the beta-lactamase gene are shown; mismatched bases are indicated in bold face characters, and the corresponding wild type bases are shown below in parenthesis. T_d is the melting temperature estimated from the empirical relationship T_d= 4 X (G-C base pairs) + 2 X (A-T base pairs) (29). The discriminating temperature is the lowest wash temperature (in 6X SSC) at which a difference in hybridization was observed between a perfectly matched DNA duplex (oligomer/mutant) and the corresponding heteroduplex (oligomer/wild type).

*From Jim Neitzel's data.

**From Steve Schultz's data.

REFERENCES

1. Eisenstadt, E., Warren, A. J., Porter, J., Atkins, D. and Miller, J. H., Proc. Natl. Acad. Sci. USA 79, 1945-1949 (1982).

2. Foster, P. L., Eisenstadt, E. and Miller, J. H., Proc. Natl. Acad. Sci. USA 80, 2695-2698 (1984).

3. Miller, J. H., in The Operon, Miller, J. H. and Reznikoff, W. S., Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), 31-38 (1979).

4. Knott-Hunziker, V., Waley, S. G., Orlek, B. S. and Sammes, P. G., FEBS Lett. 99, 59 (1979).

5. Fisher, J., Belasco, J. G., Khosla, S. and Knowles, J. R., Biochemistry 19, 2895-2901 (1980).

6. Sigal, I.S., Harwood, B. G. and Arentzen, R., Proc. Natl. Acad. Sci. USA 79, 7157 (1982).

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

8. Pollitt, S. and Zalkin, H. J. Bacteriol. 153, 27-32 (1983).

9. Messing, J., Methods in Enzymol. 101, 20-89 (1983).

10. Wallace, R. B., Schold, M., Johnson, M. J., Dembek, P. and Irakura, K., Nucleic Acids Res. 9, 3647-3655 (1981).

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

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

13. Wallace, R. B., Johnson, P. F., Tanaka, S., Schold, M., Itakura, K. and Abelson, J. N., Science 209, 1396-1400 (1980).

14. Gillam, S. and Smith, M., Gene 8, 81 (1979).

15. Ito, H., Ike, Y. and Itakura, K., Nucleic Acids Res. 10, 1755-1769 (1982).

16. Hunkapiller, M., Kent, S., Caruthers, M., Dreyer, W., Firca, J., Giffin, C., Horvath, S., Hunkapiller, T., Tempst, P. and Hood, L., Nature 310, 105-111 (1984).

17. Greenfield, L., Simpson, L. and Kaplan, D., Biochimica et Biophysica Acta 407-365 (1975).

18. Radding, C. M., Ann. Rev. Biochem. 47, 847-880 (1978).

19. McDonell, M. W., Simon, M. N. and Studier, F. W., J. Mol. Biol. 110, 119-146 (1977).

20. Lewis, E. D., Chen, S., Kumar, A., Blanck, J., Pollack, R. E. and Manley, J. L., Proc. Natl. Acad. Sci. USA 80, 7065-7069 (1983).

21. Hanahan, D., J. Mol Biol. 166, 577-580 (1983).

22. Molecular Cloning, A Laboratory Manual, Maniatis, T., Fritsch, E. F., and Sambrook, J., Eds., (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982)

23. Gergen, J. P., Stern, R. H. and Wensink, P. C., Nucleic Acids Res. 7, 2115-2136 (1979).

24. Zoller, M. J., and Smith, M., Nucleic Acids Res. 10, 6487-6499 (1983).

25 Wallace, R. B., Personal Communication (1983).

26. Wallace, R. B., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T. and Itakura, K., Nucleic Acids Res. 6, 3543 (1979).

27. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C. Heyneker, H. L. and Boyer, H. W., Gene 2, 95-113 (1977).

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

29. Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. and Itakura, K., Proc. Natl. Acad. Sci. USA 78, 6613-6617 (1981).

CHAPTER III

STRUCTURAL VARIANTS BY ANOTHER APPROACH : IMPOSING A FUNCTION RATHER THAN IMPOSING A STRUCTURE

STRUCTURAL VARIANTS BY ANOTHER APPROACH : IMPOSING A FUNCTION RATHER THAN IMPOSING A STRUCTURE

In previous chapters, site directed mutagenesis was used as a means of deliberately imposing a known predetermined structural change on a protein in order to study its effect on function. In this chapter a different but complementary approach will be explored as a probe of structural requirements for activity: a function will be imposed in order to find out which structures will satisfy this requirement.

Imposing a function or in other words selecting a phenotype, has been an extensively used tool in classical genetics and now, in combination with the techniques of site specific mutagenesis and other recombinant DNA methods, it has become an even more powerful tool. In the study of protein structure and function, site specific mutagenesis allows one to precisely define a site or region to be studied by imposing a known structure (1): subsequently, many alternative structures can be obtained at the selected site and screened for their ability to function. For example, if a protein of interest confers a selectable phenotype, a mutation can be introduced at a specific site to produce a negative phenotype. Then, any one of a variety of random mutagenic strategies can be

applied under selective pressure (imposing a function) in order to recover revertants to a positive phenotype (functional protein). The analysis of any revertants obtained should give information on the nature of the structures allowed by the constraints of the specified function. The value of the information obtained will depend in part on how complete a search for alternative structures can be designed.

One can expect reversions at the primary site which restore function either by using the same residue found in the wild type protein or by replacing it with a different residue capable of performing the same role. In some cases one can also expect revertants that restore activity by a compensating change in structure at a secondary site (2), while retaining the mutation introduced initially at the primary site; I shall refer to these as second site revertants. These are probably more likely to occur if the primary mutation affects the folding of the protein; if the primary mutation affects a residue with a very specific role such as a catalytic residue in an enzyme, then it is far more probable to expect revertants only at the primary site. Whether substitutions at a catalytic site are possible or not, can yield significant insights into the mechanism involved.

A recent example of the use of reversion experiments as

a guide to major structural features is provided by the study of the secretion of the maltose binding protein. A deletion mutant which remained cell bound reverted to a secreted protein only when a certain number of hydrophobic residues were reinserted, restoring at least in part the length of the hydrophobic stretch lost due to the deletion (3). The classical studies on the lac repressor protein (4) are another example of this type of approach, but done with the limitation of only studying sites at which chain terminating mutations could be obtained by random mutagenesis.

Beta-lactamase also provides an interesting system in which to apply these experimental approaches. It confers a selectable phenotype, and as pointed out in previous chapters, it meets all the requirements for in vitro mutagenesis: its gene is cloned, has been entirely sequenced and the enzyme is expressed abundantly. It is secreted into the E. coli periplasmic space and when the expression is controlled by a highly efficient promoter such as the tac (5) or trp (6) promoters, up to 4 mg of pure protein per liter of culture can be obtained.

In previous experiments a double mutant of beta lactamase (ser70-->thr; thr71-->ser) was obtained by oligonucleotide directed mutagenesis (Chapter I) and was shown to have no detectable enzymatic activity; the gene for this mutant beta-lactamase with the sequence -thr70-ser71- at

the active site differs from the wild type gene by two base pairs (Table I, Chapter II). No single base pair substitution in the mutant gene can give rise to a gene encoding the wild type protein sequence (-ser70-thr71-); therefore, this double mutant can be used in reversion experiments in which base changes can ocurr only at single sites with an appreciable frequency, to ask which sequences other than the wild type (if any) can restore activity, without a high background of wild type revertants to obscure them.

With this particular arrangement in which the two conserved aminoacids at the beta-lactamase active site have been mutated, one can ask whether activity can be recovered by any single aminoacid substitution at either one of the two altered positions while maintaining the other mutant residue unchanged.

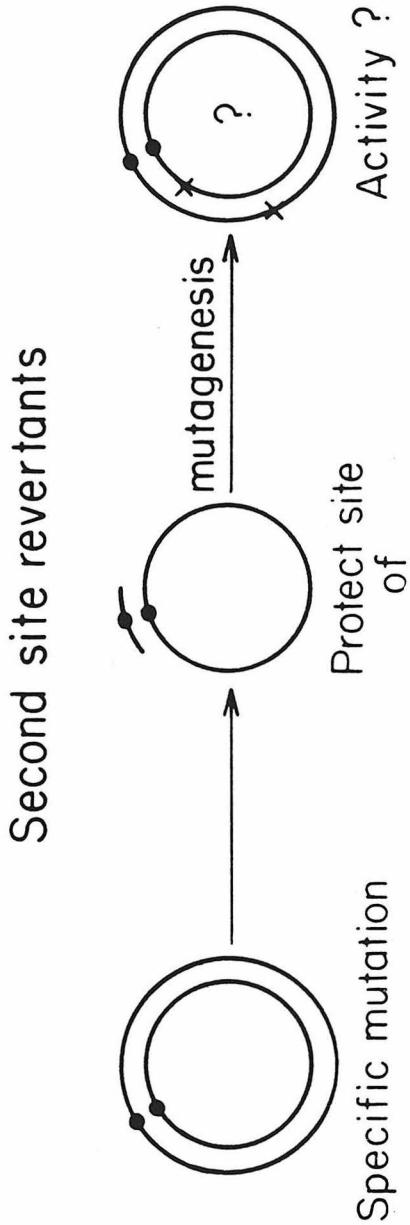
For a replicating bacterial cell one can expect a rate of spontaneous mutagenesis of the order of $1/10^6$ per locus or structural gene per generation (7) for forward mutations, and a somewhat lower rate for direct reversions, which require a mutation at a particular base pair within that gene. Consequently, the probability of n mutagenic events occurring at the same locus or gene would be no greater than $(1/106)^n$; therefore, when plating ~1×10^9 rapidly replicating cells, one should expect to see only single base pair mutations if no mutagen is used.

However, since a single base change from a given codon will only generate a particular subset of aminoacids, some aminoacid replacements will be excluded in this experimental strategy, and other approaches are necessary to explore all possibilities. This will be discussed further in the Results and Discussion section of this chapter.

Another way in which random and site specific mutagenesis can be combined advantageously is to direct mutagenesis **away** from a specified region in order to search for second site revertants, also known as intragenic suppressors. Again the inactive beta lactamase double mutant can be used (Figure III-1): annealing and enzymatically extending the oligonucleotide initially used to introduce the double mutation to an appropriate template allows one to protect the primary site while subjecting the entire template to mutagenic strategies directed at single stranded DNA. Again, protecting the primary mutation site from mutagenesis prevents a large background of active colonies due to reversion to the wild type sequence, in this case by making a specified site physically inaccessible to mutagenic agents.

FIGURE III-1

Strategy to obtain second site revertants while protecting the primary mutation site.


Second site revertants
Specific mutation
Protect site of mutation
mutagenesis
?
Activity ?

MATERIALS AND METHODS

Bacterial strains :

Mutagenesis and/or reversion experiments were done in E. coli strain LS1 (1), and E. coli JM103 was used as a host for phage M13mp8 (8) used in DNA sequencing. Culture media was L broth or L agar unless otherwise indicated.

DNA

Plasmid DNA was prepared by the alkaline lysis method (9). The plasmids used were pBR322 (10) and its derivative carrying a double mutation (Chapter I) 22G9. Sequence determination was done by the Sanger method (11) on a 700 baase pair EcoRI-PstI fragment subcloned into M13mp8; a synthetic oligonucleotide complementary to the insert at a site 50 bases away from the sequence of interest, was used as a primer.

Enzymes:

Restriction enzymes and exonuclease III were obtained from BRL, T_4 DNA ligase from New England Biolabs, and DNA polymerase (Klenow fragment) was from Boehringer-Mannheim.

RESULTS AND DISCUSSION

REVERTANTS

In order to ask questions regarding the possible aminoacid sequences at the active site of beta lactamase that might be compatible with at least some degree of catalytic activity, a simple phenotypic screen was carried out as follows.

E. coli strain LS1 harboring the doubly mutant plasmid 22G9, that encodes an inactive beta lactamase (ser70-->thr; thr71-->ser , Chapter I), was cultured in broth to a cell density of $5x10^8$ cells/ml, and then plated on L agar containing 10mg/l of ampicillin. The plates were incubated overnight at 37°C, and seven colonies were found that had reverted to ampicillin resistance. Plasmid DNA was obtained from each of these revertants (R1, R2, R3, R4, R5, R6, R7), and used to transform competent E. coli which were plated on L agar plates containing 20mg/l of tetracycline. Transformants were obtained from six of the plasmids (all except R7); R1, R2, R3, R4, and R6 also grew on 10mg/l of ampicillin as well as on tetracycline, but R5 did not.

Therefore five of the revertants had plasmid encoded ampicillin resistance. Each of the corresponding transformants was colony purified, and plasmid DNA was prepared. Plasmids were digested with EcoRI and analyzed by gel elec-

trophoresis on 1.2% agarose ; since plasmid R1 appeared to be different in size and restriction sites, it was assumed to be a contaminant and was not characterized further. Plasmids R2, R3, R4, and R6 were characterized by hybridization experiments, and R2 by DNA sequencing as well.

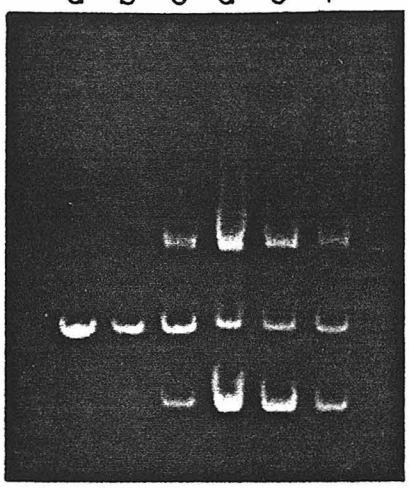
Figure III-2 shows the result of a hybridization experiment in which the plasmids from the four revertants and two control plasmids were hybridized to the oligomer 5'-^{32}p- AATGATG**AGCTCT**TTT; the two controls were pBR322 (-ser70-thr71-) and the double mutant 22G9 (-thr70-ser71-). At room temperature the hybridization probe, which codes for the sequence -ser70-ser71-, binds to all six plasmids and no significant difference in signal can be observed. However, after washing at a higher temperature it only remains bound to the plasmids from the four revertants. In Figure III-3 this result (panel B) is compared to another experiment in which the same set of plasmids were hybridized to the oligomer used to obtain the double mutant 22G9 (panel A). Clearly, the probe that matches the DNA sequence of the inactive double mutant used in the reversion experiment binds only to the DNA of the inactive parent , but not to the DNA of any of the ampicillin resistant revertants. Table I summarizes the data, showing the DNA and amino acid sequences of the controls in the region of the beta lactamase active site as well as the DNA sequence of the two hybridization probes.

FIGURE III-2

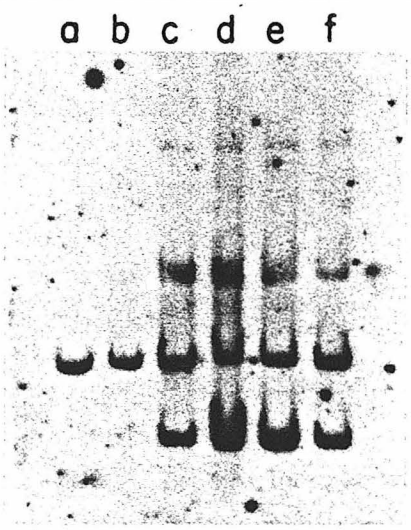
Hybridization of an oligonucleotide coding for the sequence ser70-ser71 to DNA from four revertants to ampicillin resistance: autoradiograms of a dry gel after hybridization of the oligonucleotide 5'-AATGATGAGCTCTTTT to linearized plasmid DNA from the four revertants, R2, R3, R4, R6, and two controls, pBR322 and the double mutant 22G9. The dry gel hybridization technique is described in Chapter I, Figure 5, for a similar experiment.

pBR322 22G9 REVERTANTS

A. Ethidium bromide stained gel.



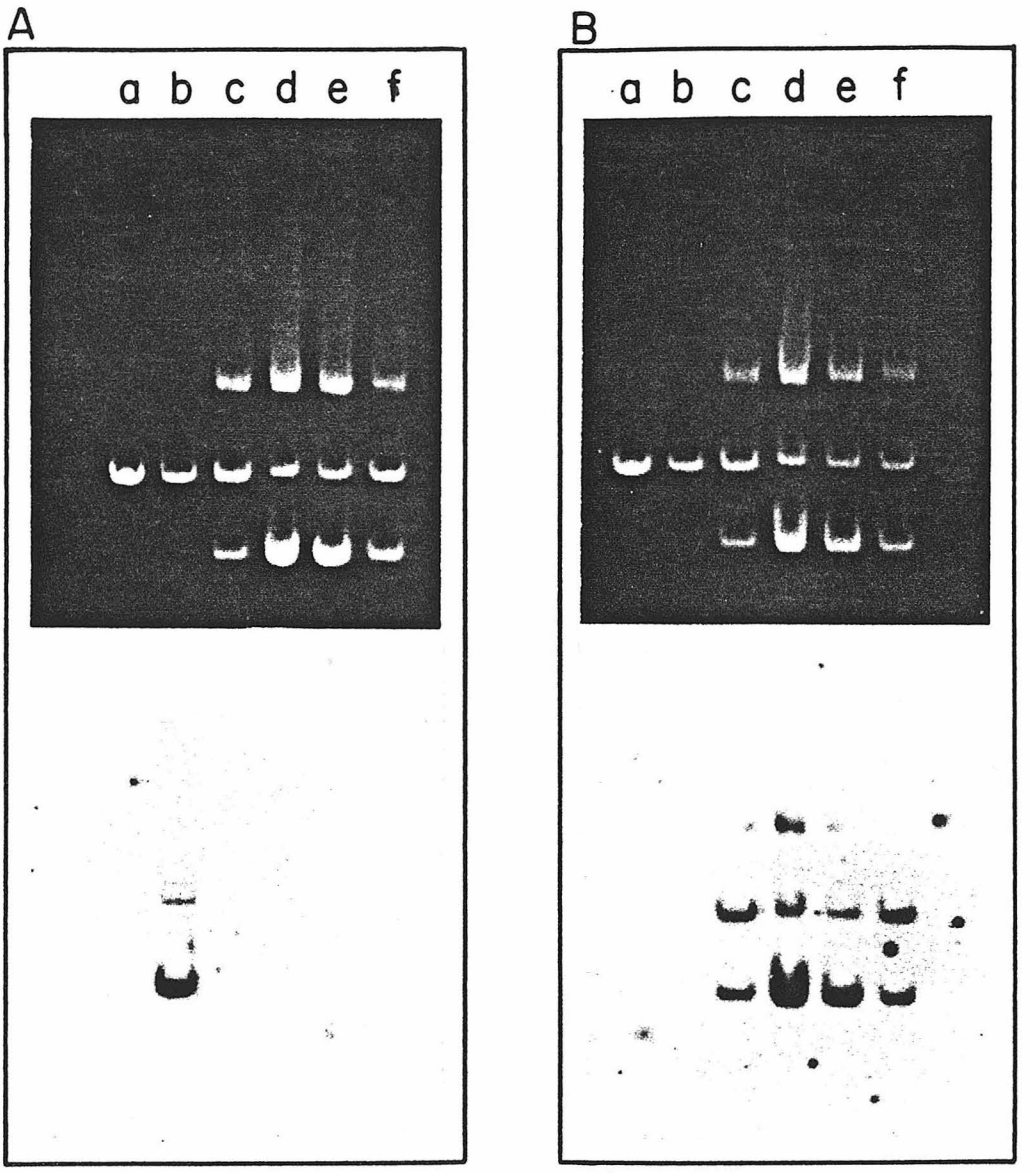
B. Room t° wash, 6X SSC



C. 50°C wash, 6X SSC

FIGURE III-3

Hybridization of two different oligonucleotides to plasmid DNA from two controls, pBR322 and 22G9, and from four revertants to ampicillin resistance. The technique and hybridization conditions are described in Chapter I, Figure 5.



PROBE: $^{5'}$AATGATGACCTCTTTT $^{5'}$AATGATGAGCTCTTTT

Lane a. pBR322 (···AGCACT···)

Lane b. 22G9 (···ACCTCT···)

Lanes c-f. Revertants to Amp^{R}

TABLE I

PLASMID	SEQUENCE	PHENOTYPE	HYBRIDIZATION PROBE SEQ*	
			ACC TCT or	AGC TCT
			-thr-ser-	-ser-ser-
pBR322	AGC ACT -ser-thr-	Amp^R	-	-
22G9	ACC TCT -thr-ser-	Amp^S	+	-
R2	?	Amp^R	-	+
R3	?	Amp^R	-	+
R4	?	Amp^R	-	+
R6	?	Amp^R	-	+

* Both probes are hexadecamers and their complete sequences are 5'-^{32}P-AATGATGACCTCTTTT-3'and
5'-^{32}P-AATGATGAGCTCTTTT-3'
The + and - singns indicate whether each of the probes hybridized to the immobilized plasmid DNA's after washing at a discriminating temperature.

These experiments demonstrate once again (12, 13) that two sequences differing by one or more bases can be distinguished from each other by hybridization to oligomeric probes.

All four revertants to ampicillin resistance exhibit similar properties while being clearly different from either of the two controls, so one can infer that their DNA sequences are identical. It is not clear whether these four clones represent four separate reversion events or whether they are siblings derived from a single revertant. DNA sequencing of revertant R2 by the Sanger method (11) confirmed the sequence AGCTCT which codes for ser70-ser71, in agreement with the hybridization results.

Note that only certain aminoacid substitutions were possible in this experiment; Table II shows the codons that can arise by a single base change at either position 70 or 71 of the beta-lactamase double mutant 22G9 (thr70-ser71).

TABLE II

ACC(thr70)-->ACA(thr)	ACG(thr)	ACT(thr)
	AAC(asn)	**AGC(ser)**	ATC(ile)
	GCC(ala)	CCC(pro)	**TCC(ser)**
TCT(ser71)-->TCA(ser)	TCG(ser)	TCC(ser)
	TAT(tyr)	TGT(cys)	TTT(phe)
	ACT(thr)	GCT(ala)	CCT(pro)

Two different serine codons could arise by a single base change in the threonine codon ACC, yet only one (AGC) was obtained. This could mean either that the four revertants are siblings or that CG to GC transversions, involving a transient CC mismatch, occur with greater frequency than AT to TA transversions with a transient AA mismatch. It is interesting to note in this regard that CC (and AG) mismatches are not corrected by the Streptococcus pneumoniae Hex mismatch repair system, whereas AA mismatches (and others) are corrected (14). It is not known whether mismatch repair in E. coli is carried out by a system with similar properties.

A number of substitutions, including thr70-->cys which would require a double base change, would not have been examined in the search for revertants described above. Other strategies are necessary to ascertain that all possible replacements have been produced at a given site, while demanding a function under selective conditions, and in order to make negative information more significant. One such approach is currently being evaluated by Steve Schultz using beta lactamase as a model system: it consists of cloning a mixture of synthetic oligomers which code for all possible replacements at one site while maintaining the wild type sequence at all other sites. Since in the case of most proteins it is possible to engineer restriction sites without altering the aminoacid sequence, it usually should

be possible to limit the length of the mixture of synthetic oligomers to a reasonable number of bases. Other approaches inlude using several different types of chemical mutagens to increase the frequency of different kinds of base changes, or the misincorporation of base analogues such as the excision resistant alpha-thionucleotides (15). However, even when all types of base changes are achieved by random mutagenesis , one must still be aware of the vastly different frequencies that separate single base changes from double and multiple base changes.

The only catalytically active structures yet found for beta lactamase all posess a primary nucleophile at position 70, either the hydroxyl group of serine in the wild type and in these revertants, or the sulfhydryl group of cysteine in thiolactamase (ser70-->cys)(6, 16). The conserved threonine at position 71, however, can be replaced by serine and possibly by other residues without destroying catalytic activity, although it is known that in S. aureus beta- lactamase replacement of threonine by the bulky, hydrophobic side chain of isoleucine at position 71 virtually abolishes activity (17). Some of the properties of the R2 beta lactamase (-ser70-ser71-) will be discussed in Chapter IV.

SECOND SITE REVERTANTS

As an example of an approach to determine whether changes outside the region of the active site could restore catalytic activity to a protein that had a threonine at position 70, the synthetic oligonucleotide 5'-AATGATGAC CACTTTT was hybridized to a plasmid DNA template prepared from the inactive beta lactamase double mutant 22G9 (-thr70-ser71-; ACC TCT) and extended in the conditions described below and as shown schematically in Figure III-1. The binding of the synthetic oligonucleotide will protect the region of the active site from mutation during the in vitro replication of the template carried out in the presence of limiting amounts of each of the four deoxynucleotide triphosphates in turn. In these conditions one can expect an increased incidence of nucleotide misincorporation at sites complementary to the limiting nucleotide.

The template was prepared from plasmid DNA by the method described in detail in Chapter II, Methods section (b). Briefly, it consists of introducing single stranded nicks using an endonuclease (HpaII in this case) in the presence of ethidium bromide and treating the nicked material with E. coli exonuclease III. In Chapter II, I argued that a limited exonuclease digestion would be preferable when preparing a template for mutagenesis. In this case, however, exonuclease digestion was deliberately extensive in order to obtain completely single stranded

circles so that the only primer available for DNA polymerase would be the synthetic oligonucleotide. Nicked DNA (20 ug) was digested for 60 minutes at 37°C with 40 units of E. coli exonuclease III in the appropriate buffer conditions(see Chapter II); the enzyme was heat inactivated in the presence of chloroform and the solution was lyophilyzed.

160 pmoles of the synthetic oligonucleotide representing a 40 fold molar excess was annealed to the template in a 12.5 ul volume of 0.5 M NaCl (and the salts carried over from the previous reaction) by heating in a boiling water bath for 5 minutes and allowing it to cool slowly to room temperature. The annealed material was divided into five aliquots and five reaction mixtures were prepared containing 0.5mM ATP and the concentrations of deoxynucleotide triphosphates indicated below.

Reaction	dATP	dGTP	dCTP	dTTP
A	0.5uM	0.5mM	0.5mM	0.5mM
G	0.5mM	0.5uM	0.5mM	0.5mM
C	0.5mM	0.5mM	0.5uM	0.5mM
T	0.5mM	0.5mM	0.5mM	0.5uM
Control	0.5mM	0.5mM	0.5mM	0.5mM

The buffer and salt concentrations were adjusted to 100 mM NaCl, 10 mM $MgCl_2$, 20 mM Tris-HCl pH 7.5, 10 mM DTT (dithiothreitol) and the final reaction volume was 50ul.

The reactions were started by adding 4 units of the Klenow fragment of DNA polymerase I from E. coli, and 100 units of T_4 DNA ligase. They were incubated 15 minutes at room temperature and overnight at 15°C. Then 1 ul of a 10mM solution of the limiting deoxynucleotide triphosphate was added to each reaction, as well as another aliquot (as above) of each of the two enzymes. Incubation was continued for one hour at room temperature, and the reaction products were purified by extraction with phenol and chloroform and precipitated with ethanol. The pellets were dried , redissolved in 40ul of TE buffer and an aliquot was used to transform competent E. coli LS1. The cells were plated on L agar plates containing either 10 ug/ml of Ampicillin or 20 ug/ml of tetracycline and incubated overnight at 37°C.

The total possible number of transformants examined (from all four reactions) was 15,000, as determined from the control transformation plated on tetracycline. This refers to the transformation done with the control DNA which was repaired in non-mutagenizing conditions. No active second site revertants were isolated by this procedure.

In order to draw a significant conclusion from a negative result such as this, it would be necessary to screen a much larger number of molecules subjected to mutagenizing conditions. Even then, some caution would be appropriate because one can never be absolutely certain that every pro-

tein structure with threonine at position 70 was examined in the search for second-site revertants. This negative result, however, is consistent with other data that suggests that a primary nucleophile at position 70 is an essential requirement for lactamase activity.

REFERENCES

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

2. Helinski, D. R. and Yanofsky, C., J. Biol. Chem. 238, 1043 (1963).

3. Bankaitis, V. A., Rasmussen, B. A. and Bassford, Jr., P. J., Cell 37, 243-252 (1984).

4. Miller, J. H., Coulondre, C., Hofer, M., Schmeissner, V., Sommer, H., Schmitz, A. and Lu, P., J. Mol. Biol. 131, 191 (1979).

5. DeBoer, H. A., Comstock, L. J. and Vasser, M., Proc. Natl. Acad. Sci. USA 80, 21 (1983).

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

7. Lewin, B., Gene Expression 3 (Wiley Interscience), 57 (1980).

8. Messing, J., Methods in Enzymol. 101, 20-89 (1983).

9. Molecular Cloning, A Laboratory Manual, Maniatis, T., Fritsch, E. F., and Sambrook, J., Eds., (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982)

10. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L. and Boyer, H. W., Gene 2, 95-113 (1977).

11. Sanger, F., Nicklen, S. and Coulson, A. R., Proc. Natl. Acad. Sci. USA 74, 5463-5467 (1977).

12. Wallace, R. B., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T. and Itakura, K., Nucleic Acids Res. 6, 3543 (1979).

13. Wallace, R. B., Schold, M., Johnson, M. J., Dembek, P. and Itakura, K., Nucleic Acids Res. 9, 3647-3656 (1981).

14. Claverys, J.-P., Mejean,V., Gasc, A.-M. and Sicard, M., Proc. Natl. Acad. Sci. USA 80, 5956-5960 (1983).

15. Shortle, D., Grisafi, P., Benkovic, S. J. and Botstein, D., Proc. Natl. Acad. Sci. USA 79, 1588 (1982).

16 Sigal, I. S., Harwood, G. B. and Arentzen, R., Proc. Natl. Acad. Sci. USA 79, 7157-7160 (1982).

17. Ambler, R. P., Phil. Trans. R. Soc., London Ser. B 289, 321 (1980).

CHAPTER IV

SOME MUTANTS OF BETA-LACTAMASE AND THEIR PROPERTIES

SOME MUTANTS OF BETA-LACTAMASE AND THEIR PROPERTIES

The beta-lactamase from E. coli is a small, monomeric, globular protein of molecular weight 28,500 (1), with approximately 40% alpha helical content (2). Its crystal structure has not yet been solved at high resolution, but a four angstrom resolution electron density map has been reported (2).

There are several different beta lactamases from various bacterial sources, which Ambler (1) has classified as classes A and B; a third class C, including the chromosomally encoded beta lactamases of E. coli K12 (ampC) and Pseudomonas aeruginosa shares sequence homology distinct from that characterizing class A enzymes (3). Class A beta lactamaseses appear to have a common evolutionary origin and to be related to transpeptidases (4); Class A enzymes include the beta lactamases from B. cereus, B. licheniformis, S. aureus and E. coli (the subject of this study), all of whose aminoacid sequences are known and which exhibit considerable sequence homology and share a number of mechanistic features. Class B enzymes of molecular mass 23,000 are unrelated, showing little or no aminoacid sequence homology and a different catalytic mechanism which requires a metal cofactor, usually ZnII (1); Class C en

zymes are unrelated as well, having a different molecular weight and aminoacid sequence. In addition, there are many beta lactamases that have not been classified due to insufficient aminoacid sequence data and mechanistic information.

The class A beta lactamases are serine hydrolases which catalyze the ring opening of bicyclic beta lactams such as penicillins and cephalosporins, but are not capable of hydrolyzing linear peptides (5); it has been recently shown, however that these enzymes are capable of hydrolyzing certain monocyclic beta lactams (6), and even acyclic depsipeptides ($RCONH-CR_1R_2COO-CHR_3COOH$) (7) albeit at a very slow rate. The apparent order of substrate preference (bicyclic beta lactam > monocyclic beta lactam > depsipeptides >> linear peptides) appears to reflect the susceptibility of each type of compound to nucleophilic cleaveage. This is consistent with the proposed mechanism of catalysis which entails a nucleophilic attack of the active site serine hydroxyl group on the carbonyl carbon of the substrate. Serine 70, (in Ambler's numbering system), (1) has been covalently modified by the irreversible inhibitor 6-beta bromopenicillanic acid in the B. cereus beta lactamase I (8, 9), and similar inhibitors (10) modify the corresponding serine in the homologous plasmid encoded E. coli enzyme. A covalent acyl-enzyme intermediate was demonstrated spectroscopically using the very poor

substrate cefoxitin and was shown to have the characteristics of an aliphatic ester consistent, therefore, with an acylserine (11).

The aminoacid sequence information on class A beta lactamases (1) indicates that the enzymes are made of a single polypeptide chain varying in length from 264 to 273 aminoacids for the mature soluble proteins. There is a conserved serine-threonine dyad at the active site, but the role of the threonine residue in catalysis is not understood, although there is a report of virtually complete loss of activity in a mutant S. aureus beta lactamase with isoleucine (instead of threonine) adjacent to serine. No other aminoacid residues have been directly implicated in the catalytic mechanism, and although the existence of a second internal nucleophile or general base is plausible, by analogy with serine proteases, none has been demonstrated; reference was made by Ambler to another almost inactive mutant of S. aureus beta lactamase in which conserved aspartate 179 was substituted by asparagine, but since the characterization has not gone beyond identifying the aminoacid lesion in a randomly obtained mutant, no conclusions can be drawn about the role of the residue.

Evidence indicates these enzymes are biosynthesized as precursors, and in the case of the E. coli enzyme encoded by the plasmid pBR322 the processing of the precursor to the mature enzyme and its secretion into the periplasmic

space have been studied (12, 13, 14).

Koshland and Botstein (12, 13) demonstrated that beta lac tamase is translocated posttranslationally across the inner membrane and that 23 N terminal aminoacids are removed during or immediately after translocation to yield the mature periplasmic enzyme. Mutations in the signal sequence and in other regions of the protein were used to characterize different steps in the process (15). Whereas the carboxyl terminus is not required for for translocation, it seems to be necessary for the enzyme to be completely released as a soluble molecule in the periplasm. Apparently the energy required for export of beta lactamase is fulfilled by the total protonmotive force, with no specificity for either the membrane potential or the pH gradient alone (16).

Other mutants of beta lactamase have been obtained by Shortle and Botstein (unpublised results) by segment directed mutagenesis, and screening for those that had reduced resistance to ampicillin. Interestingly, in the N terminal segment of the mature enzyme including the first 48 aminoacids the mutations found with reduced catalytic activity had alterations in one or another of the three conserved residues found in this region of the protein. These mutants have not been characterized further, and so one can only speculate on the reasons why each one affects the function of the enzyme; the possibilities include al-

teration of the final folded form of the protein, changes in the kinetics of the folding pathway, and even direct involvement of these residues in secretion, in substrate binding or in catalysis.

In order to study the catalytic properties of the enzyme, several specific aminoacid substitutions were designed and introduced in the active site region. As mentioned above the role of the active site serine had been fairly well established by chemical modification and kinetic studies by 1980, but the possible role of the adjacent conserved threonine was not; in addition, it seemed possible that perhaps threonine could function in place of serine in catalysis. In order to gain further understanding of the role of these residues and the possible use of a secondary nucleophile in catalysis, we set out to obtain mutants of beta lactamase containing all of the permutations of serine and threonine at positions 70 and 71.

The first one obtained by oligonucleotide-directed mutagenesis was the double mutant described in Chapter I . The mutant sequence -ser70-ser71- was obtained as a revertant of this inactive mutant that had regained enzymatic activity (Chapter III). The mutant sequence, -thr70-thr71- was obtained by oligonucleotide-directed mutagenesis as described in Chapter II.

These aminoacid substitutions were not expected to affect the overall conformation of the protein to any large

extent, so it was somewhat surprising to discover early that the active mutant -ser70-ser71- appeared to be less resistant to proteolytic digestion than the wild type protein, and that its isolation required growing the bacteria at or below 30^{0}C. This observation suggested that the stability of the native conformation of the protein was altered with respect to that of the wild type, or that the folding kinetics had been altered enough to allow proteolysis to occur before the mature protein could adopt the native conformation.

Since more of the enzyme's properties than just catalysis might have been altered, study of the effects of these mutations not only on the activity of the mature protein, but also on the processing and secretion of the precursor seemed essential. For example, the occurrence of internal sequences in the processed protein that nevertheless influence secretion has ample precedent (17, 12). To this end the mutants were transferred by *in vivo* recombination methods to a variant of the temperate Salmonella phage P22 that carries a copy of the beta lactamase gene, and which had been used previously (12) to study the secretion of beta lactamase. The genetic transfer of the three active site mutants was done in David Botstein's laboratory as described below in the Results and Discussion section.

One additional mutant of beta lactamase will be discussed in this chapter. The E. coli beta lactamase has been

shown to have a disulfide bond (18) between cysteines 77 and 123, the only two cysteines present in the polypeptide chain. Evidence indicates that, whereas in the precursor *in vivo* the half cystines are reduced, as one might expect (12, 18), the enzyme precursor does form a disulfide bond if synthesized by in vitro translation. In vivo the formation of the disulfide bond occurs simultaneously with the processing of the precursor, and the role, if any, of the disulfide in the translocation and processing of the enzyme can only be tested with certainty by substituting one of the two cysteines with a different residue.

In addition a beta lactamase with a single cysteine residue would be useful in preparing heavy atom derivatives for crystallographic studies, and in studies of the role of the disulfide bond in stabilizing the three dimensional structure of the protein. The study of disulfide bonds which are buried, as in this and many other known cases (19), can be hindered by the fact that one needs to denature the protein in order to expose the bond to reducing agents. In many cases disulfide bonds occur because the position of the half cystines in the polypeptide chain is such that they come in close proximity to each other in at least one conformation of minimal energy. This conformation may or may not represent the overall energy miniumum of the processed, mature protein, and so studies of the thermal stability of a beta lactamase lacking a disulfide bridge,

can indicate whether its native conformation does indeed represent a global energy minimum even when the disulfide bond has been removed. When studies are undertaken by denaturing and reducing a protein, it is not always possible to regain an active protein by renaturation. Indeed, one can sometimes show that the irreversibility of a denaturation step is due to the fact that the protein becomes trapped in an ordered but inactive conformation, distinct from the native conformation, as in the case of human alpha interferon in certain conditions (20). However, the inability to refold into the native conformation might also be due to destabilizing van der Waals contacts between two bulky -SH groups, whether derivatized or not. Therefore, studies on the role of disulfide cross links can be carried out more reliably in proteins in which the disulfide bond is removed genetically rather than chemically, as this avoids the uncertainty of interpreting the influence of bulky substituents.

To these ends, I introduced in E. coli beta lactamase the substitution cys77-->ser by oligonucleotide directed mutagenesis as described in detail in Chapter II. I chose serine because it occurs at the corresponding site in another class A beta lactamase (S. aureus) and the hydroxyl group is probably small enough to be accomodated opposite a reduced sulfhydryl group without disrupting the structure, while still maintaining a potential hydrogen bond acceptor

at position 77; alanine might have been another good choice because of its smaller size and the fact that it also occurs at position 77 in B. cereus beta lactamase, another class A enzyme.

In this Chapter, the results of pulse chase experiments designed to determine the rate of processing of the beta lactamase wild type and mutant precursors will be presented, as well as some of the properties of the three active site mutant proteins, and of the cys77-->ser mutant. Secretion studies of the latter have not yet been undertaken .

Materials and Methods

Phage and bacterial strains.

The P22 phage and Salmonella typhimurium strains used in the work described in this Chapter were from David Botstein's strain collection. The relevant characteristics of the bacterial strains are listed below.

Strain DB4673 was used because it is more readily transformed than other Salmonella strains. DB6142 is a PolA mutant which does not allow plasmid replication. DB4381, hisG46 del (bio-uvrB), is particularly sensitive to damage by ultraviolet light, and was used as the host in labeling and pulse chase experiments. DB7000, leuA-am414 is wild type S. typhimurium. DB7609 is a gal^-, amp^+ (low) derivative of DB7000 used in combination with the parent strain to identify phages which are either sensitive or

resistant to ampicillin.

The P22 phage strains used were P22Ap3lpfr1, also called P22bla which carries a copy of the ampicillin resistance gene identical to the one encoded by pBR322 in lieu of sequences in the secondary immunity region ImmI. The phage P22bla 1752 carries the wild type ampicillin resistance gene, and P22bla 4393 is a derivative that carries an ampicillin resistance gene with a large deletion (13).

Red plate test for ampicillin resistance on P22bla strains:

This method has been described in detail (21), so this is a brief description of the principle. This method was developed to rapidly assay a large number of phage for their ampicillin phenotype. Phage are plated on an agar plate which contains galactose, tetrazolium dye and 1 ug/ml of ampicillin, using a mixture of two tester strains: DB7000, a gal^+, amp^- strain, and DB7609, a gal^-, amp^+ strain. Its ampicillin resistance is very low, and is not due to secreted beta lactamase, therefore this strain does not save neighboring amp^- cells. The galactose and tetrazolium act together as a color indicator which stains galactose fermenting cells red, leaving gal^- cells colorless. Both tester strains are sensitive to P22 infection and so phage plate whether or not they carry beta lactamase. The gal^- strain, which is amp^+ will give rise to colorless plaques when infected; the amp^- cells

however will only survive when infected by ampicillin resistant phage, and since they ferment galactose, they will give rise to red plaques on the indicator plates. Thus ampicillin resistant phage plate as red plaques and sensitive phage as colorless plaques.

Labeling of phage-encoded proteins:

Labeling of phage-encoded proteins after infection, and pulse chase experiments were done as described by Koshland and Botstein (13). A 30 second pulse of ^{35}S-methionine was used, and samples were taken every 20 seconds after addition of the chase mixture unless otherwise indicated. Each sample was transferred immediately into an Eppendorf tube containing gel loading sample buffer preheated to 90°C as described. The samples were analyzed by SDS-polyacrylamide gel electrophoresis using 12% gels according to Laemmli (22); the gels were fixed in 10% acetic acid, 50% ethanol before drying and exposing to Kodak XR5 film at room temperature. Autoradiograms were traced using a Cary 219 spectrophotometer with the appropriate attachments.

Results and Discussion

Three mutants in the region of the active site of beta lactamase with the sequences ser70-ser71-, thr70-thr71-, and thr70-ser71-, have been obtained as described in previous chapters; all three have been shown to cross react with

antibody raised against native wild type beta lactamase (23), but only one of them (-ser70-ser71-) is catalytically active. All three mutants appear to be more susceptible to proteolytic damage than the wild type enzyme during the early stages of purification (23), and it is only possible to obtain good yields when the host cells are cultured at or below 30°C and the cell extracts are kept cold throughout the early stages of purification. This is in contrast with the wild type enzyme which is quite stable even if grown at 37°, and handled at higher temperatures. These observations led to the conclusion that a more thorough study of the functions and physiology of these mutants would be in order, since very minor, conservative changes in protein sequence produced, in addition to the predicted effects on catalysis, some unexpected changes in other properties of the enzyme. Recent results (24) indicate that this is not an isolated phenomenon, and that seemingly inocuous point mutations may be disastrous for a protein's expression by simply making it more sensitive to proteases. Indeed, in some cases the rate of proteolysis *in vivo* could be greater than the overall rate at which the protein is synthesized thereby giving the false appearance that the mutant protein is functionally inactive.

In order to compare the rate of processing of the wild type and mutant beta lactamase precursors, the genes for the mutant beta lactamases were transferred to a previously

characterized phage system(12), and pulse labeling and pulse chase experiments were done. Other systems such as maxi cells (18) could be used for the same purpose, but the one used here is the same one used in prior beta lactamase studies.

Phage-specified proteins are easily identified by SDS-polyacrylamide gel electrophoresis because they can be labeled specifically after infection of a host that has been irradiated with ultraviolet light. Koshland and Botstein used this method to investigate the synthesis and maturation of beta-lactamase in S. typhimurium; the structural gene for beta lactamase had been previously incorporated (12) into the bacteriophage P22 as a simple substitution of the beta lactamase structural gene for the dispensable secondary immunity region(immI) of P22. The active site mutants of beta lactamase were transferred to this phage derivative, designated P22-bla, as described below and the resulting phage strains were used in labeling experiments.

Transfer of active mutants:

S. typhimurium DB4673 was transformed with pBR322-R2, the plasmid encoding the mutant beta-lactmase gene of sequence -ser70-ser71-. The resulting ampicillin resistant strain was subsequently infected at a multiplicity of infection of 0.1 pfu/cell with phage P22bla-4393, a strain carrying a large deletion in the beta lactamase gene, which

encompasses the region of the active site. Phage progeny from this infection were used to infect strain DB6142, a polA mutant in which the plasmid cannot replicate, at a multiplicity of infection of 0.1. After plaque purification, the phage progeny was used to infect cells at a multiplicity of infection of 10, and ampicillin resistant lysogens were selected on L agar plates containing 50 ug/ml ampicillin. After colony purification of the lysogen, it was cultured in broth and induced by ultra violet light irradiation. The resulting phage progeny was plated and a single plaque was used to infect a 100 ml broth culture to prepare phage stock by standard procedures.

Transfer of inactive mutants:

A strain DB4673 lysogen of the phage P22-bla 1752 (Amp^r, wt), was transformed with plasmid DNA from each of the two inactive mutants of beta lactamase (-thr70-ser71- and -thr70-thr71-). After transformation, the lysogen was induced and the phage progeny was titrated and used to in fect strain DB6142 at a multiplicity of infection of 10. These cells were plated on L agar containing a low amount (2 ug/ml) of tetracycline, and incubated at 37°C until some colonies began to appear (this took 48 hours). Presumably these slightly tetracycline resistant colonies had in the bacterial chromosome a cointegrate of the plasmid and the phage DNA of the form $-bla^+-tc^+-bla^-$; the low level of tetracycline resistance would be due to

the fact that these cells carry a single copy of the tetracycline resistance gene as opposed to 20 to 30 copies when the gene is encoded by a high copy number plasmid. Such a cointegrate when induced should give rise to two kinds of P22bla phage: one carrying the wild type (Amp^r) alele, and one carrying the mutant (Amp^-) alele of the beta lactamase gene. The presumed cointegrate was cultured in broth, induced by ultraviolet light irradiation, and plated on a combination of two tester strains designed to detect both Amp^r and Amp^s strains using a color reaction (see Methods). As expected, both red and colorless plaques were obtained; single colorless plaques (Amp^s) were replated and then used to prepare phage stock. The Amp^s phenotype of the phage accompanied by the occurrence of precursor and mature forms of beta-lactamase demonstrated in labeling experiments (see below) was considered evidence that the transfer of the mutant beta lactamase genes had taken place.

Pulse chase experiments:

Labeling of phage encoded proteins and pulse chase ex periments were done in the conditions described in the methods section.

At 37°C the incorporation of label into both precursor and mature forms of beta lactamase is rapid, so that even sampling every 15 seconds, after a 30 second pulse of ^{35}S-methionine, it is difficult to observe the

rate of processing accurately. Experiments done at this temperature did not reveal any differences between the processing of the wild type enzyme and any of the three active site mutants. In order to determine whether there were subtle differences, the experiments were carried out at 30°C for one of the mutants, since the process is slower and can be observed with greater resolution.

Figure IV-1 is an autoradiogram of ^{35}S-methionine labeled products of a typical pulse chase experiment analyzed by SDS polyacrylamide gel electrophoresis. A wild type control and the active mutant R2 (ser70-ser71-) are shown side by side ; each lane represents a sample taken at the specified time after the chase was initiated. The intensity of the bands was measured by doing a densitometer tracing, and integrating each peak. Figure IV-2 shows the quantitative result; each point was normalized to the sum of the label present in each lane of the gel. There is no significant difference between the rate of processing of the mutant and the control, and this was reproducible in three independent experiments.

Similar results were obtained with the two inactive mutants. Figures IV-3 and IV-4 show the results of pulse chase experiments done with the phage encoded mutant -ser70-->thr, and the double mutant ser70-->thr, thr71-->ser respectively. The results shown in figure IV-4 are unfortunately obscured by background labeling of cellu-

FIGURE IV-1

Autoradiogram of ^{35}S-methionine labeled proteins from a pulse-chase experiment comparing the rate of processing of mutant beta-lactamase thr71-->ser. Phage infected cells were pulse-labeled for 30 seconds and chased with chloramphenicol; samples were taken at different times after the chase and analyzed by gel electrophoresis through a 12% polyacrylamide SDS gel in reducing conditions. Lane (a) is a control infected with a phage strain carrying a large deletion in the beta-lactamase gene. Lanes (b) through (i) are extracts from cells infected with phage carrying the mutant beta-lactamase gene R2 (thr71-->ser); the length of the chase, in seconds, for lanes (b) through (i) respectively was 10, 30, 50, 70, 90, 120, 150, and 300. The same time intervals were used for the control infected with phage carrying the wild type beta-lactamase gene in lanes (j) through (q). The labeling and chase were done at 30°C.

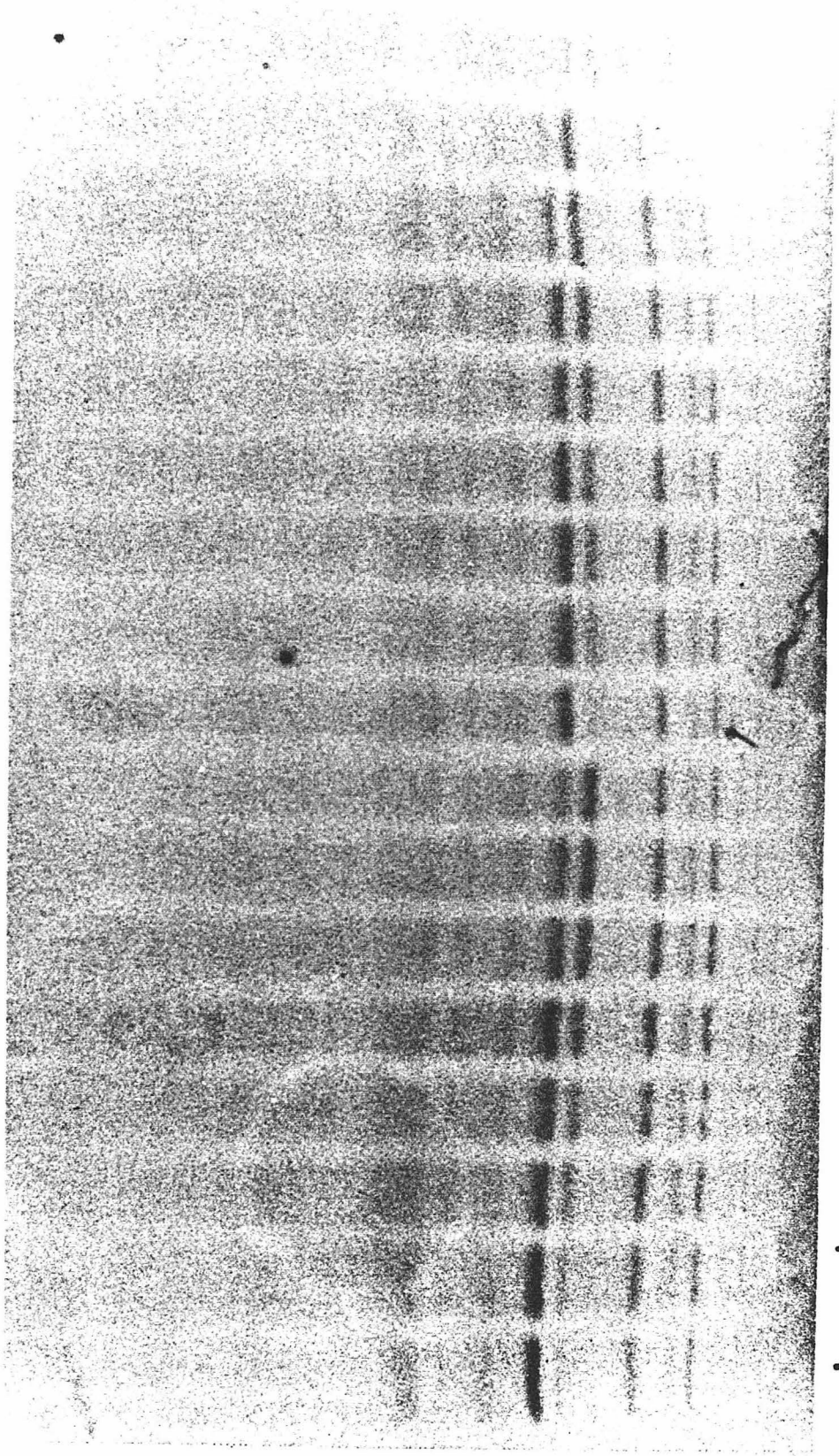

a b c d e f g h i j k l m n o p q
b p o n m l k j i h g f e d c b a

FIGURE IV-2

Quantitation of pulse-chase analysis of wild type and mutant beta-lactamase gene products.

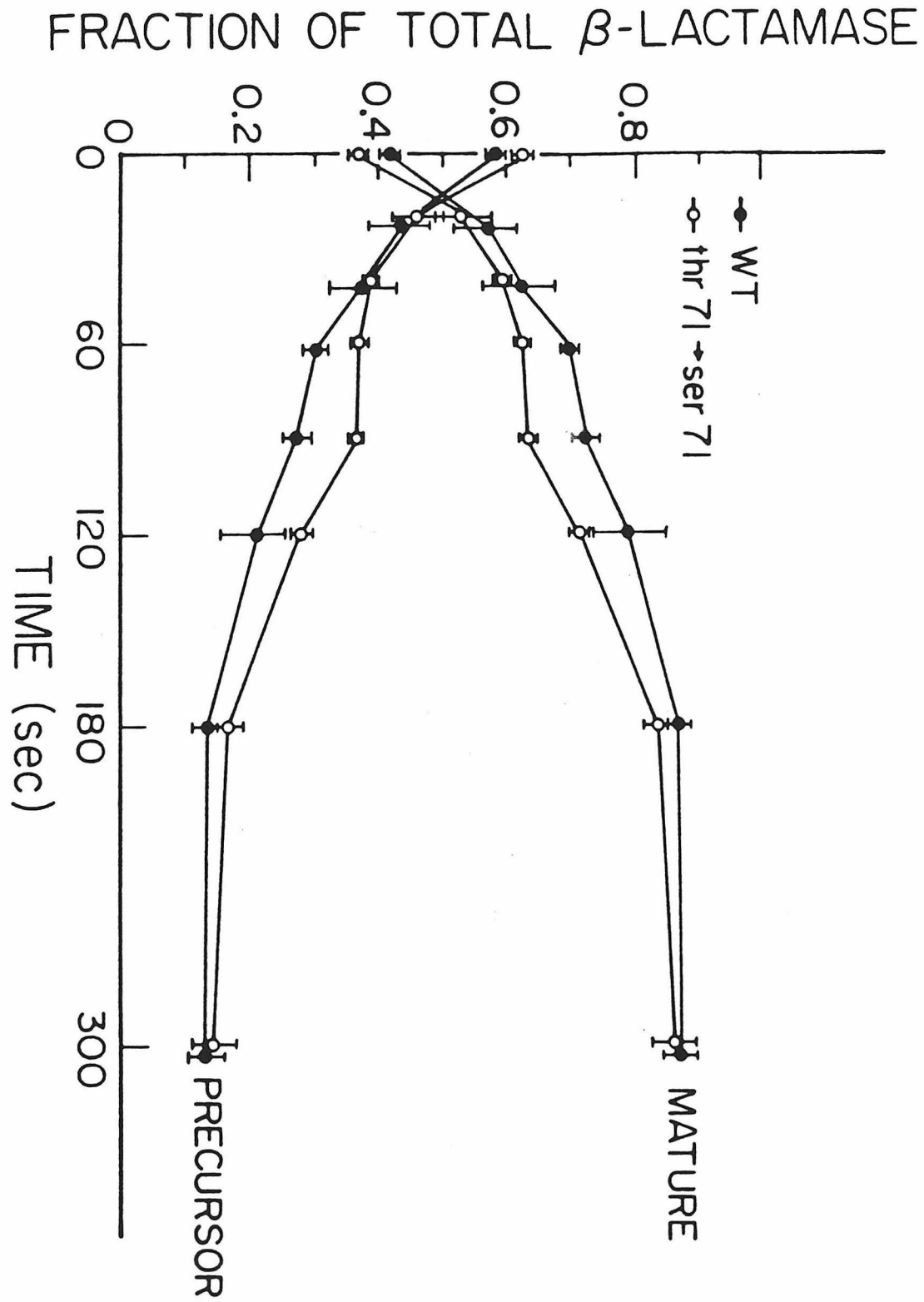

FRACTION OF TOTAL β-LACTAMASE
0
0.2
0.4
0.6
0.8
WT
thr 71→ser 71
0
60
120
180
300
TIME (sec)
PRECURSOR
MATURE

FIGURE IV-3

Autoradiogram of ^{35}S-methionine labeled proteins from a pulse-chase experiment comparing the rate of processing of mutant beta-lactamase ser70-->thr. The experiment was done as described in the text and in Figure IV-1. Lane (a) is a control infected with a phage strain carrying a large deletion in the beta-lactamase gene. Lanes (b) through (i) are extracts from cells infected with phage carrying the mutant beta-lactamase gene (ser70-->thr); the length of the chase, in seconds, for lanes (b) through (i) respectively was 10, 30, 50, 70, 90, 120, 150, and 300. The same time intervals were used for the control infected with phage carrying the wild type beta-lactamase gene in lanes (j) through (q). The labeling and chase were done at 37°C.

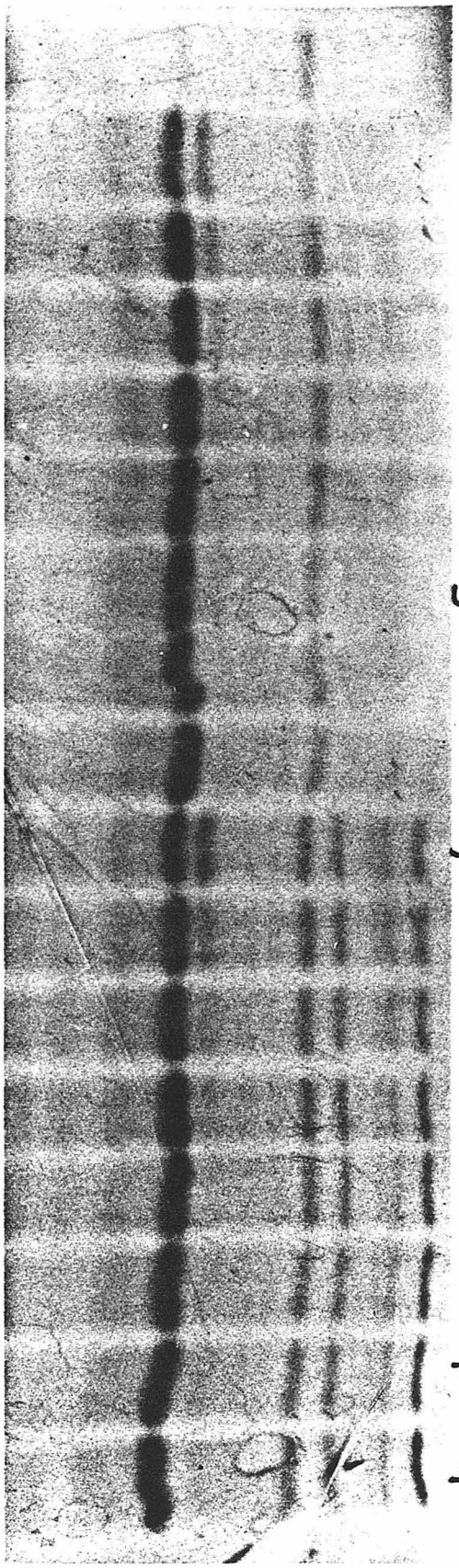

a b c d e f g h i j k l m n o p q

FIGURE IV-4

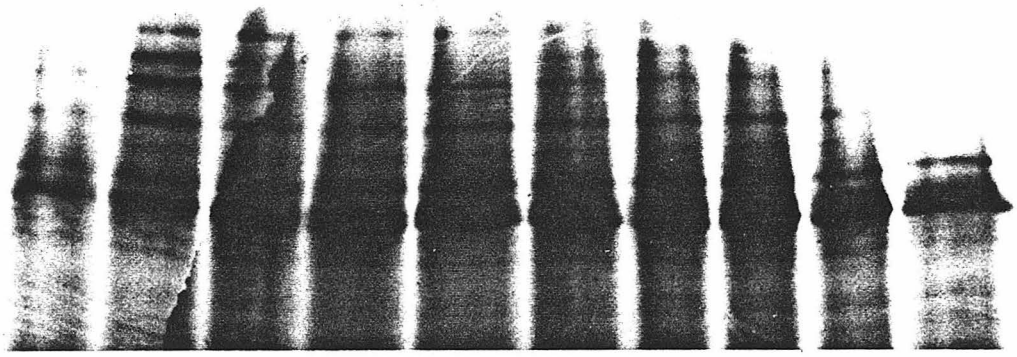
Autoradiogram of ^{35}S-methionine labeled proteins from a pulse-chase experiment comparing the rate of processing of mutant beta-lactamase ser70-->thr/thr71-->ser. The experiment was done as described in the text and in Figure IV-1. The labeling and chase were done at 37°C.

Panel A: Lane (a) is a control consisting of uninfected cells; lane (b) is a control infected with a phage strain carrying a large deletion in the beta-lactamase gene. Lanes (c) through (j) are extracts from cells infected with phage carrying the wild type beta-lactamase gene ; the length of the chase, in seconds, for lanes (c) through (j) respectively was 0, 30, 60, 90, 120, 150, 180 and 300.

Panel B: Lane (a) is a control infected with phage carrying a partially deleted beta-lactamase gene. Lanes (b) through (m) are extracts from cells infected with phage carrying the mutant beta-lactamase gene (ser70-->thr); the length of the chase, in seconds, for lanes (b) through (m) respectively was 0, 30, 60, 90, 120, 150, 180 240, 300, 360, 420 and 480.

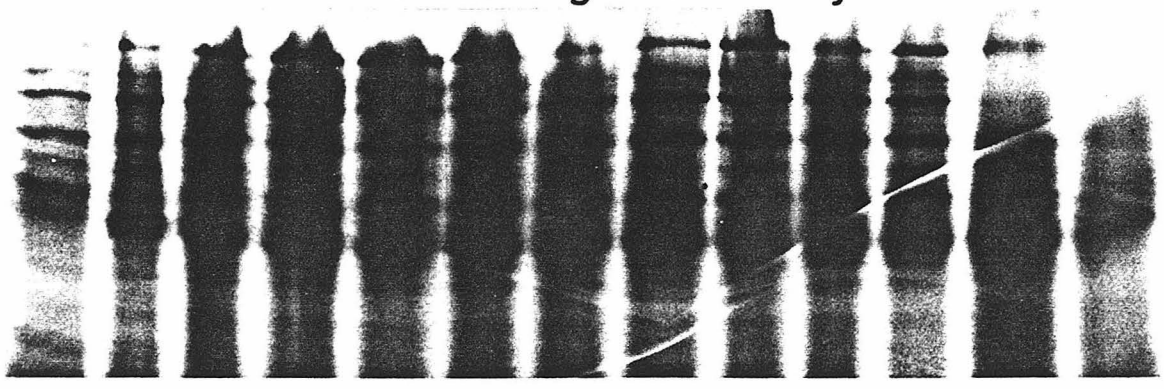
A.

a b c d e f g h i j



B.

a b c d e f g h i j k l m



lar proteins, as well as by poor electrophoretic resolution of precursor and mature beta lactamase, but there is no apparent indication of any dramatic difference between the rate of processing of the wild type and the mutant protein. (Attempts to repeat the experiment some months later were unsuccessful because after preparing fresh phage stock the beta lactamase gene apparently had been lost. Since it seemed unlikely that any striking differences would be found, the transfer of this mutant to phage P22bla was not repeated.)

To summarize, of the three beta lactamase active site mutants that I have obtained only one retains catalytic activity whereas the other two are inactive. Another active site mutant of beta lactamase, thiolactamase, has been reported in the literature and may represent the only substitution at position 70 which gives an active enzyme (23). The three active site mutants appear to be less stable than the wild type protein, and the active revertant, in addition, has a reduced k_{cat}/K_m value for benzylpenicillin (23). On the other hand, the beta lactamase (cys77-->ser) mutant lacking the disulfide bond is active and stable in normal assay conditions (kinetic parameters have not been determined accurately yet, but preliminary results indicate that any differnces with the wild type enzyme will be small).

It appears that the disulfide bond is only incidental,

in that it is not required to maintain the enzyme in its native conformation, since an active conformation can be regained by the cys77-->ser mutant after thermal denaturation (23). When a crude extract from either the wild type or the cys77-->ser mutant strain, containing beta lactamase and other periplasmic proteins, including some protease(s), is assayed at different temperatures ranging from 30°C to 60°C, the following takes place: both wild type and mutant enzymes suffer a small but comparable loss of activity as the temperature is increased from 30°C to 50°C. At approximately 55°C the activity of the mutant enzyme drops rapidly, while that of the wild type enzyme continues to lose activity at a lower rate as temperature is increased. If the heat inactivated mutant enzyme is cooled again to 30°C at least 50% of the activity is regained , indicating that the loss of activity was due to unfolding of the protein rather than proteolysis. In contrast, if the active revertant (-ser70-ser71-) is subjected to thermal denaturation in similar conditions, it loses activity rapidly above 40°C. In this case, the loss of activity is irreversible, due to proteolysis (23). This difference would seem to indicate a change in the unfolding kinetics of the active site mutant with respect both to the wild type and to the cys77-->ser mutant. On the other hand, the initial stages of the unfolding pathway are probably very similar in both the wild type and the mutant

without the disulfide bridge. The other two active site mutants which are inactive also appear to be more sensitive to proteolytic damage during purification than the wild type enzyme, but they have not been charactrized further in this respect.

How does this speak of the structure of beta lactamase and its mechanism ? The thermal stability of a protein is determined by numerous parameters, and one must distinguish between those that influence the overall stability or, in other words, the work required to transfer the protein from the native state to the completely denatured state, and those that affect the stability of the native protein and are responsible for producing local, transient unfolding in the native protein . For small globular proteins the thermal denaturation can be considered a two stage cooperative transition which occurs at a characteristic temperature (24). Below this melting temperature, noncooperative predenaturational changes occur, and the frequency of these local changes in compact structure increases with temperature (25). Hydrogen exchange studies, used to judge how local unfolding brings internal groups into contact with the solvent, indicate that in many cases, different peptide hydrogens are exchanged at different rates, probably as a result of the difference in unfolding energies for different parts of the protein (24). The aminoacid substitutions at the active site of beta-lac-

tamase may have the effect on the enzyme's stability of increasing the rate of local transient unfolding below the melting temperature, thus making some regions of the enzyme more accessible to solvent and to proteolytic digestion. In the case of the active mutant thr71-->ser, for example, the removal of a methyl group may have the effect of reducing the number of favorable hydrophobic interactions that stabilize that region of the protein, allowing greater local mobility; on the other hand, the addition of a methyl group in one case and its repositioning in another may have similar destabilizing effects on the local packing in the two inactive mutants, ser70-->thr and ser70-->thr;thr71-->ser.

In contrast, the mutant lacking the disulfide bond behaves very much like the wild type enzyme at lower temperatures and this can be easily understood if one considers that the contribution of a disulfide bond to the stability of the protein is of an entirely different nature. As J. M. Thornton pointed out (19) most available evidence suggests that disulfides do not dictate the folding of the polypeptide chain, but they do contribute to the stability of the folded protein by reducing the entropy of the denatured state still containing the intact disulfide(s), relative to that of the folded protein, as contrasted to the entropy of a denatured state in which the disulfides have been removed. The contribution of a disulfide bond to the

overall stability of a protein at temperatures and in solvent conditions that favor the folded native conformation is negligible. Therefore, the reduced stability of the cys77-->ser mutant should only be manifest at or near the melting temperature if no local instability has been introduced into the protein by the side chain substitution; this appears to be the case. Often, disulfide bridges occur in proteins that have to function in harsh, sometimes unpredictable extracellular environments (Eg., trypsin, chymotrypsin, trypsin inhibitor, etc.), and such stabilizing cross-links contribute to the preservation of the proper folding even in the presence of potentially denaturing conditions. Interestingly, in the case of the four class A beta lactamases of known sequence, the E. coli enzyme is the only one with a disulfide bond. There does not seem to be any obvious selective advantage for E. coli, and perhaps the existence of these class A beta lactamases is a case of divergent evolution from an ancestor which itself evolved to operate at a higher temperature or in an otherwise harsher environment.

The alterations in the three dimensional structure of beta-lactamase produced by the substitutions at the active site, are probably subtle since the active revertant is still a relatively efficient enzyme, and there is no effect on the maturation and secretion of the enzyme. The effects of these substitutions in the inactive mutants can almost

certainly be attributed entirely to the functional properties of the residues present at position 70, and only marginally to changes in the tertiary structure.

The reduced value of k_{cat} and slightly increased K_m for benzylpenicillin with the active revertant R2, indicate that the effect of the mutation thr71-->ser is primarily to decrease the ability of the mutant enzyme to stabilize the transition state relative to the enzyme bound substrate although substrate binding is also very slightly poorer. In the inactive mutants, however the determining factor is clearly that a secondary hydroxyl of threonine is not capable of catalyzing the reaction, probably due to steric hindrance by the methyl group. The sulfhydryl group of thiolactamase (26), a primary nucleophile, does catalyze the reaction as does the primary hydroxyl of serine. No other residues have been found to be active at position 70 and no compensating mutations have been found (Chapter III) that restore activity when threonine is present at position 70. The complete lack of activity of the double mutant ser70-->thr; thr71-->ser, indicates that the residue at position 71 cannot substitute for the serine at 70 in providing the nucleophile that opens the beta lactam ring, confirming the assignment of ser70 as the catalytic residue in the wild type enzyme.

A more detailed interpretation of these results in terms of the structure of the protein will require the long

awaited high resolution crystal structure which still has not been solved. A more detailed understanding of the changes in the catalytic mechanism of these mutant proteins will require further studies with different substrates and inhibitors, in order to establish the appropriate comparisons with the wild type beta lactamase. Of particular interest will be the effect of these structural changes in the enzyme on the relative rates of acylation and deacylation.

References

1. Ambler, R. P., Phil. Trans. R. Soc. B 289, 321-331 (1980).

2. DeLucia, M. L., Kelly, J. A., Mangion, M. M., Moews, P. C. and Knox, J. R., Phil. Trans. R. Soc. Lond. B, 289, 374-376 (1980).

3. Jaurin, B. and Grundstrom, T., Proc. Natl. Acad. Sci. USA 78, 4897-4901 (1981).

4. Tipper, D. J. & Strominger, J. L., Proc. Natl. Acad. Sci. USA 54, 1133-1141 (1965).

5. Saz, A. K., Lowery, D. L. and Jackson, L. L., J. Bacteriol. 82, 298-304 (1961).

6. Pratt, R. F., Anderson, E. G. and Odeh, I., Biochem. Biophys. Res. Commun. 93, 1266-1273 (1980).

7. Pratt, R. F. and Govardhan, C. P., Proc. Natl. Acad. Sci. USA 81, 1302-1306 (1984).

8. Cohen, S. A. and Pratt, R. F., Biochemistry 19, 1996-4003 (1980).
9. Knott-Hunziker, V., Waley, S. G., Orlek, B. S., and Sammes, P. G., FEBS Lett. 99, 59-61 (1979).

10. Fisher, J., Charnas, R. L., Bradley, S. M. and Knowles, J. R., Biochemistry 20, 2726-2731 (1981).

11. Fisher, J., Belasco, J. G., Khosla, S. and Knowles, J. R., Biochemistry 19, 2895-2901 (1980).

12. Koshland, D. and Botstein, D., Cell 30, 893-902 (1982).

13. Koshland, D. and Botstein, D., Cell 20, 749-760 (1980).

14. Winston, F. and Botstein, D., J. Mol. Biol. 152, 209-232 (1981).

15. Koshland, D. and Botstein, D., Cell 30, 903-914 (1982).

16. Baker, E. P. and Randall, The EMBO Journal 3, 895-900 (1984)

17. Bassford, P. J., Silhavy, T. J., and Beckwith, J. R., J. Bacteriol. 139, 19-31.
18. Pollitt, S. and Zalkin, H., J. Bacteriol. 153, 27-32 (1983).

19. Thornton, J. M., J. Mol. Biol. 151, 261-287 (1981).

20. Morehead, H., Johnston, P. D., and Wetzel, R., Biochemistry 23, 2500-2507 (1984).

21. Koshland, D., Ph. D thesis, Massachusets Institute of Technology, (1982).

22. Laemmli, U. K., Nature 227, 680-685 (1970).

23. Neitzel, J. J., unpublished results.

24. Villafranca, E., personal communication; Sauer, R., personal communication.

25. Sancar, A., Hack, A. M. and Rupp, W. D., J. Bacteriol. 137, 692 (1979).

26. Privalov, P. L., Advances in Protein Chemistry 33, 167-241 (1979).

27. Nakanishi, M., Tsuboi, M., and Ikegami, A., J. Mol. Biol. 75, 673-682 (1973).

28. Sigal, I. S., DeGrado, W. F., Thomas, B. J., and Petteway, S. R., J. Biol. Chem. 258, 5327-5332 (1984).

RESEARCH PROPOSITIONS

INTRODUCTION TO

PROPOSITIONS I AND II

PHOSPHOGLYCERATE KINASE: AN INTERESTING SYSTEM TO PROBE STRUCTURAL REQUIREMENTS FOR PROTEIN FUNCTION

PHOSPHOGLYCERATE KINASE: AN INTERESTING SYSTEM TO PROBE STRUCTURAL REQUIREMENTS FOR PROTEIN FUNCTION.

INTRODUCTION

Phosphoglycerate kinase (PGK) catalyses the phosphoryl transfer reaction 1,3-diphospho-D-glycerate + ADP---> 3-phospho-D-glycerate + ATP. It is an enzyme that has been purified to homogeneity from a wide variety of sources, and has been studied intensively for more than thirty years (1,2). PGKs obtained from different sources are invariably monomers with molecular weights of about 45,000, having similar catalytic properties and a highly conserved amino acid composition (1,2). The entire amino acid sequences of the yeast (3), horse muscle (4), and human red blood cell (5) proteins are known, and there is remarkable homology between the equine and human PGKs, with only 14 amino acid differences in a 416 residue chain; both have about 65% homology with the yeast enzyme. The structural genes for the yeast (6,7) and the human (8,9) enzymes have been cloned and their DNA sequences have also been determined.

High-resolution X-ray studies on the yeast (3) and horse muscle (4) enzymes have yielded detailed structural

information, and it is clear that the enzymes from both sources are highly homologous. The single polypeptide chain folds into two widely separated domains of approximately equal size corresponding to the C-terminal and N-terminal halves of the polypeptide chain, which are connected by a narrow waist through which the polypeptide chain passes only twice (the last few C-terminal residues complete the N-terminal domain).

It has only recently become possible to generate deliberately and accurately any predetermined variant of a protein by introducing changes in the primary sequence that can range from specific amino acid substitutions to major restructuring. This ability, made possible by the now widely applied techniques of in vitro mutagenesis and recombinant DNA, has widened the range of structure-function relationships that can be investigated, since we are no longer limited to the study of naturally occurring variants or to those variants that can be obtained by random mutagenesis and classical genetic methods (10).

The prerequisites for deliberately altering the sequence of a naturally-occurring protein to study the relationship between its structure and function are to have the structural gene cloned and expressed in a

suitable vector-host system, to know the entire DNA sequence of the gene, and (for some forms of specific mutagenesis) to have the ability to synthesize oligodeoxynucleotides. In addition, it is highly desirable to have available detailed structural and mechanistic information, preferably from more than one source, to design the most informative changes in primary structure and to subsequently interpret the results accurately. The PGK system meets these requirements.

The gene for the yeast enzyme is cloned and expressed (7) and its entire sequence has been determined (8). The structural gene for the human X-chromosome-linked enzyme has been cloned by J. Singer, et al., in A. D. Riggs's laboratory (8). (A full-length cDNA clone has been obtained recently (J. Singer-Sam, personal communication.) Its entire sequence is known (9) although it has not been expressed. As mentioned above, detailed structural information is available for the yeast enzyme and for the human enzyme by virtue of its homology with the horse-muscle enzyme. Furthermore, enzymatic phosphoryl transfer reactions have been the subject of many studies at the molecular, physical-organic chemical level, and a variety of methods have been devised to scrutinize their mechanism (11).

These facts, namely, the availability of structural genes from two very different eukaryotic sources for an enzyme whose structure and catalytic mechanism have been intensively studied (1-5,13-22), make this a system that is ideally suited for directed mutagenesis experiments. In addition to the study of specific problems in catalysis, this system is useful for experiments designed to shed light on general sequence requirements for different forms of secondary and tertiary structure. It also affords an excellent opportunity to study the phenomenon of domain movement, which is possibly a general requirement of kinase reactions (12).

The types of mutagenesis that can be done with this system fall into three categories:

1. Specific amino acid substitutions.
2. Random amino acid substitutions within a particular segment of the gene.
3. Major restructuring.

Oligonucleotide-directed mutagenesis is ideally suited to study a glutamate-histidine interaction that is postulated to play an important role. Substantial evidence supports domain movement on binding the two substrates, and it has been suggested that weakening of this interaction (glu 190-his 388) is the initial step

in domain movement. The substitution of glutamate 190 in the yeast enzyme by (a) glutamine, (b) aspartate, or (c) asparagine would be informative, because it would subtly alter the nature of the interaction while minimizing the changes in the size and shape of the side chain at position 190.

A rotation of one domain with respect to the other about residue 187 in the yeast enzyme or the equivalent region in the horse enzyme is consistent with experimental observations (14,4); therefore this position (and neighboring residues) is another site of considerable interest. For example, substitution of leucine 187 by proline would, in all likelihood, interfere with domain movement if this is indeed the "hinge" region it appears to be.

Segment-directed mutagenesis in this region probably would also prove fruitful, since many variants could be produced, and their ability or inability to undergo the postulated conformational change could be correlated with amino acid substitutions and provide information on the nature of the hinge region. Mutations that result in a loss of function can be used later to identify any alternative residues capable of restoring biological activity. This sort of data would be a useful addition to the growing body of information used to predict protein

structure from primary sequence.

Major restructuring of a protein can also provide information, especially on the role of entire regions. Since the genes for human and yeast enzymes are available, the construction of human-yeast hybrid proteins is possible and could be informative. A protein consisting of an N-terminal domain from one source and a C-terminal domain from the other source would be interesting, and its activity or inactivity would give information on the degree of sequence variation that can occur without disruption of the overall structural and functional requirements of the protein. If the resulting hybrids are inactive, the requirements for restoration of activity could be studied.

When the genes for a given protein are available from more than one source but no detailed structural information exists, the construction of hybrids could be used to determine whether the extent of sequence conservation is sufficient to produce homologous three-dimensional structures and to determine whether nonconserved residues are critical or not. Performing this experiment in a system for which high-resolution X-ray data is available would test the validity of these assumptions.

These are only a few examples of mutagenic experi-

ments that could be done on this system. In the following section, I will describe in detail the strategy for cloning the human-yeast hybrid genes, and for introducing some of the directed-point mutations discussed.

PROPOSITION I

CONSTRUCTION OF HUMAN-YEAST PGK HYBRIDS

The separation of PGK into two domains is clear, but defining a point in the primary sequence as the boundary between those two domains is somewhat arbitrary. The so-called waist region of the protein extends approximately from the carboxyl end of beta strand F in both the horse and the yeast enzymes to some point in alpha helix 7 in the horse enzyme, or to the equivalent helix V in yeast. At this point, the polypeptide chain crosses from the N-terminal domain to form the C-terminal domain. The last few C-terminal residues stack on the N-terminal domain, and contribute to the structure of the waist region as the polypeptide chain passes through it a second time (3,4).

As a first approximation, one can chose the region that first passes through the waist region as the arbitrary boundary. Later one could refine this definition by taking into consideration in the construction of hybrids, that the last 10 to 12 C-terminal residues constitute part of the N-terminal domain, and that therefore they should have the same origin as the rest of that domain. Residues 184 through 186 in yeast (186 through 188 in horse and human enzymes) are conserved in the three species, and connect beta strand F with

helix V (or 7). Because a splice in a conserved sequence would be the least disruptive, this appears to be a good choice. Although the amino acid sequence in this segment is conserved, the DNA sequence is not. Since there are no restriction sites conveniently located at this site, it is necessary to create a junction at this point by other means.

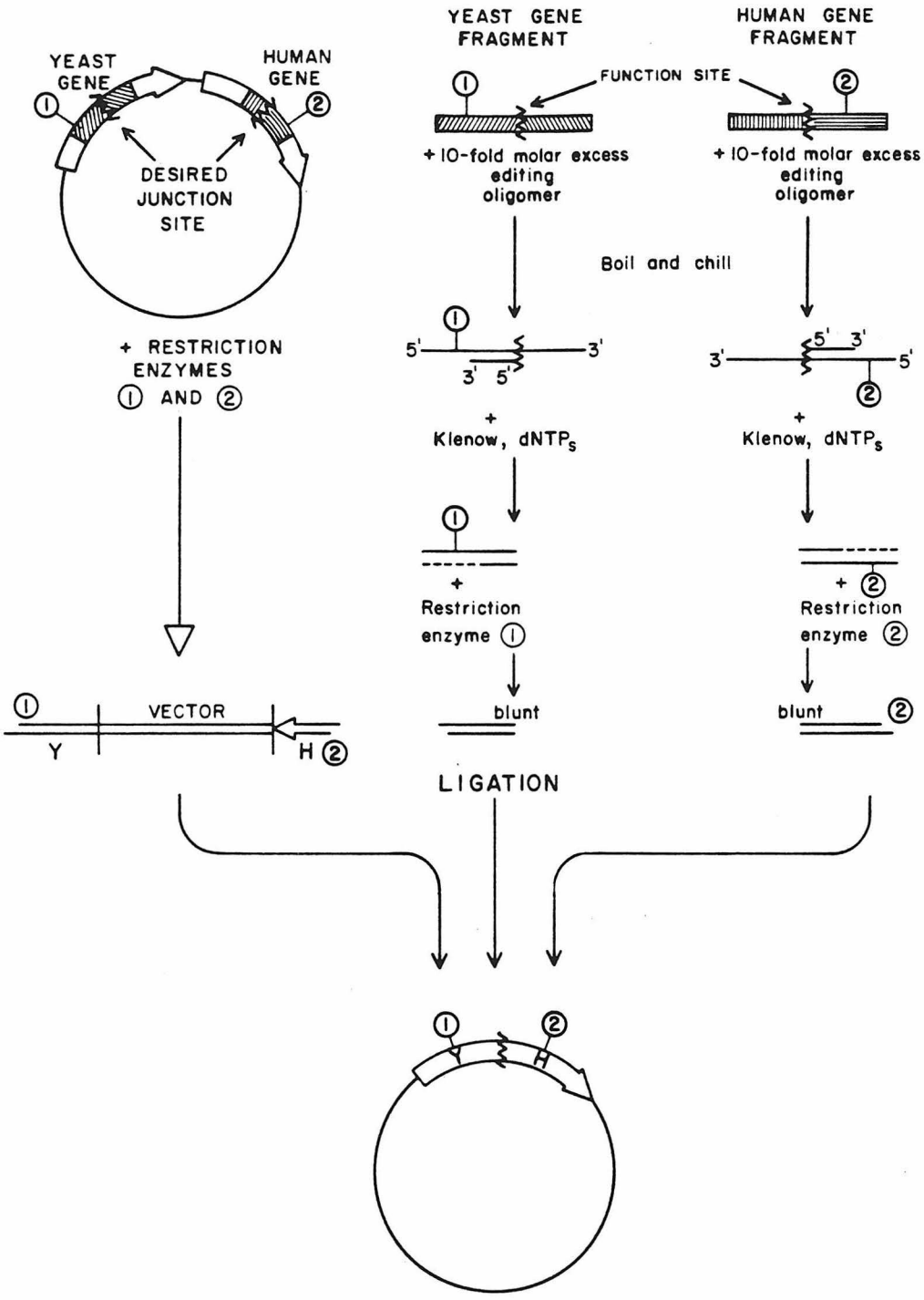
Several techniques are available to accomplish this end, but one stands out because of its simplicity and effectiveness. In addition to its template-dependent polymerization activity, the Klenow fragment of DNA polymerase I from E. coli is capable of digesting single-stranded DNA in the 3' to 5' direction (23), but digestion stops when a perfectly complementary double-stranded region is reached. This property has been used to create blunt-ended junctions at a specific site in a DNA sequence with the aid of a synthetic oligodeoxy-nucleotide that acts as an "editing primer" (24,25).

The oligomer complementary to the DNA sequence at the site to be "edited" is designed so that its 5' end constitutes the junction that needs to be created. A suitable restriction fragment of a few hundred base pairs in length containing the region to be modified is heat-denatured and then chilled in the presence of excess

oligomer. The synthetic DNA is annealed specifically, and its 3' end is extended when DNA polymerase (Klenow fragment) and the four deoxynucleotide triphosphates are added, while the bases on the initial restriction fragment, which lie 3' to the annealing site, are digested. This reaction results in a double-stranded, blunt-ended DNA fragment, which extends from the 5' end of the oligomer (the junction) to the intact 5' end of the DNA strand that served as a template. In addition to this, if necessary, the sequence at the junction can be modified by designing the oligonucleotide with an altered sequence at its 5' end. The 3' editing activity of DNA polymerase will digest any bases that are noncomplementary to the 5' end of the oligonucleotide, and will faithfully copy the sequence of the oligomer. These methods can be adapted to produce hybrids of the yeast and human PGK genes to be used later to direct the synthesis of hybrid PGK enzymes.

The overall strategy is depicted in Figure 1. If both genes (or appropriate 5' and 3' terminal segments) are cloned sequentially in the same vector and in the same relative orientation, a double restriction digest can generate a fragment containing the vector and the terminal segments of each gene. Internal fragments of each gene are edited to create a blunt junction, and

FIGURE 1


YEAST GENE
HUMAN GENE
DESIRED JUNCTION SITE
+ RESTRICTION ENZYMES ① AND ②
YEAST GENE FRAGMENT
FUNCTION SITE
+ 10-fold molar excess editing oligomer
Boil and chill
Klenow, dNTPs
Restriction enzyme ①
blunt
HUMAN GENE FRAGMENT
+ 10-fold molar excess editing oligomer
Klenow, dNTPs
Restriction enzyme ②
blunt
VECTOR
Y
H
LIGATION

restricted to produce specific cohesive ends that are compatible with those of the segments attached to the vector. A three-fragment ligation reaction will yield a circular DNA molecule consisting of the vector and a reconstructed hybrid PGK gene with the half corresponding to the N-terminal domain of the protein from one source and the C-terminal half from another, precisely joined at a specific predetermined site in the sequence.

Figure 2A shows an overall restriction map of the yeast and human PGK genes, including only those sites that pertain to the experimental strategy. Figure 2B shows in detail the selected junction site, as well as the "editing" oligonucleotides. Because of the differences in the DNA sequence of the two genes, two junctions are needed to create the two hybrid proteins $N_{human}C_{yeast}$ and $N_{yeast}C_{human}$ (where N and C represent the N-terminal and carboxy-terminal domains respectively). Therefore, four oligonucleotides are required. The four editing reactions are shown in Figure 3.

A PstI-XbaI fragment of the human gene, including all of the structural gene and parts of the 5' and 3' untranslated regions, can be cloned into the corresponding sites in the vector M-13 mp11, after removal of the vector's single HindIII site. The yeast gene, contained

Figure 2

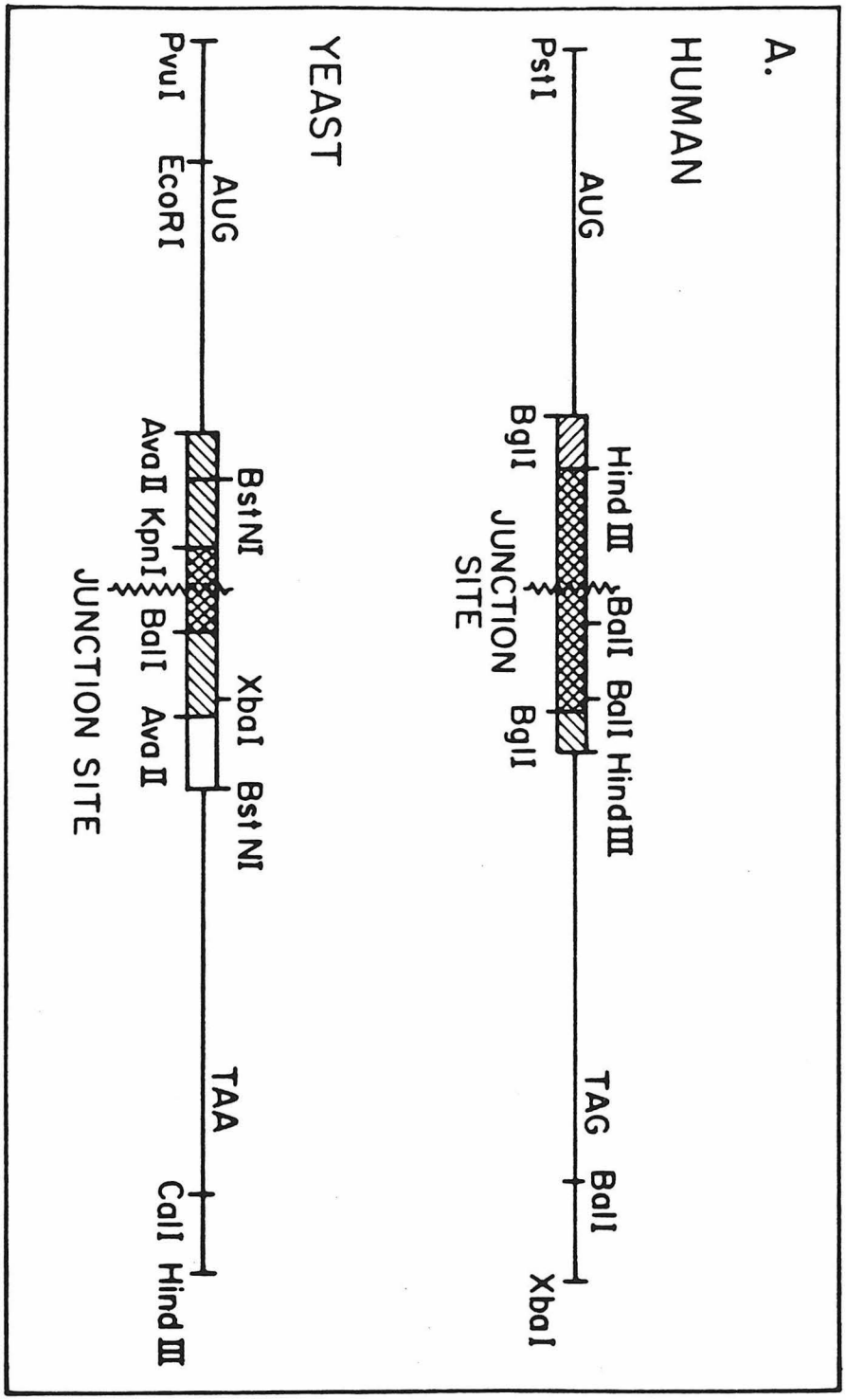
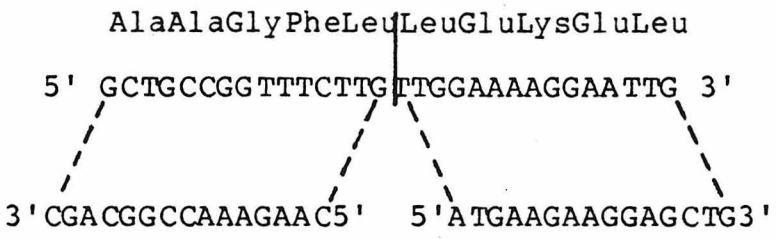

A.
HUMAN
PstI
AUG
BglI
HindIII
BalI
BalI
HindIII
JUNCTION SITE
BglI
TAG
BalI
XbaI
YEAST
PvuI
EcoRI
AUG
AvaII
KpnI
BalI
AvaII
JUNCTION SITE
BstNI
XbaI
BstNI
TAA
CalI
HindIII

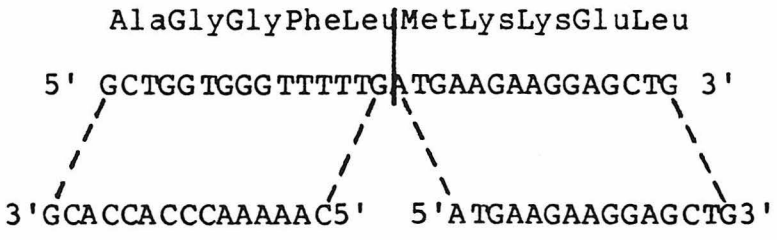
FIGURE 2 B

Sequence of the selected junction site and of the editing primers.

YEAST GENE JUNCTION SITE



HUMAN GENE JUNCTION SITE



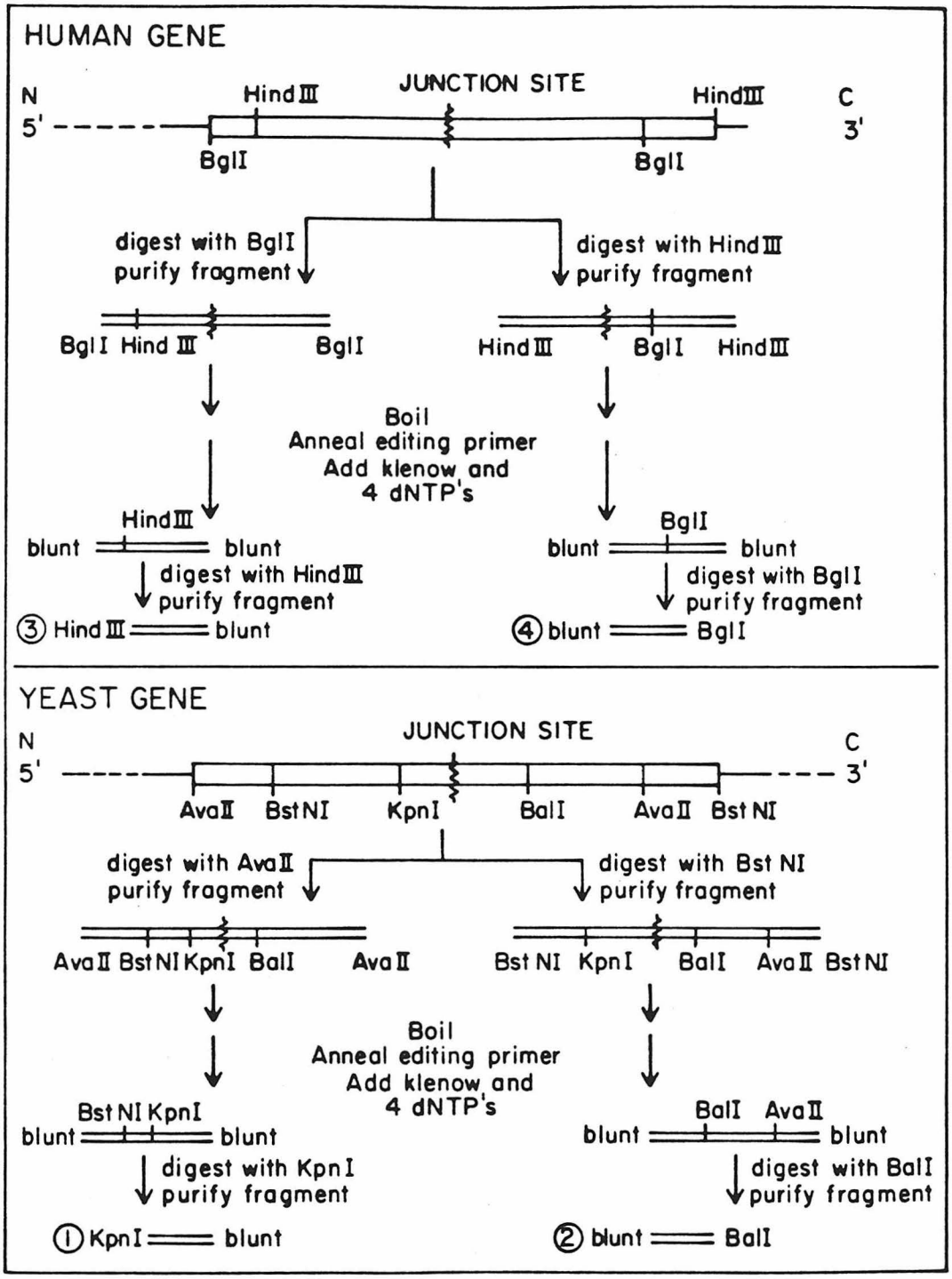
in a Pvu I-Hind III fragment can be cloned by blunt-end ligation, either into the Pst I site 5' to the human gene, or into the Eco RI site 3' to the human gene, producing vectors $N_\gamma C_H$ and $N_H C_\gamma$ respectively. Clones carrying both genes can easily be identified by using the "editing" primers as hybridization probes. Since the yeast gene can enter each site in either of two orientations, the clones carrying it in the desired orientation must be identified, either by their restriction patterns or by hybridization to the "editing" oligonucleotides.

Vector $N_\gamma C_H$ is digested with KpnI and BglI, which have no target sites in the M13 sequence, but do have sites on the yeast and human genes respectively. The largest resulting fragment is used in a three-fragment ligation reaction with yeast KpnI-blunt fragment (No. 1 in Fig. 3) and human blunt-BglI fragment (No. 4 in Fig. 3) that resulted from "primer-editing" reactions.

Vector $N_H C_\gamma$ can be used in a similar set of reactions, digesting with HindIII and BalI in this instance. The vector fragment, the human Hind III-blunt fragment (No. 3 in Fig. 3) and the yeast blunt-Hha II fragment (No. 2 in Fig. 3) are ligated to yield the desired hybrid gene. Figure 4 summarizes these reactions.

The three-fragment ligation reactions, as well as

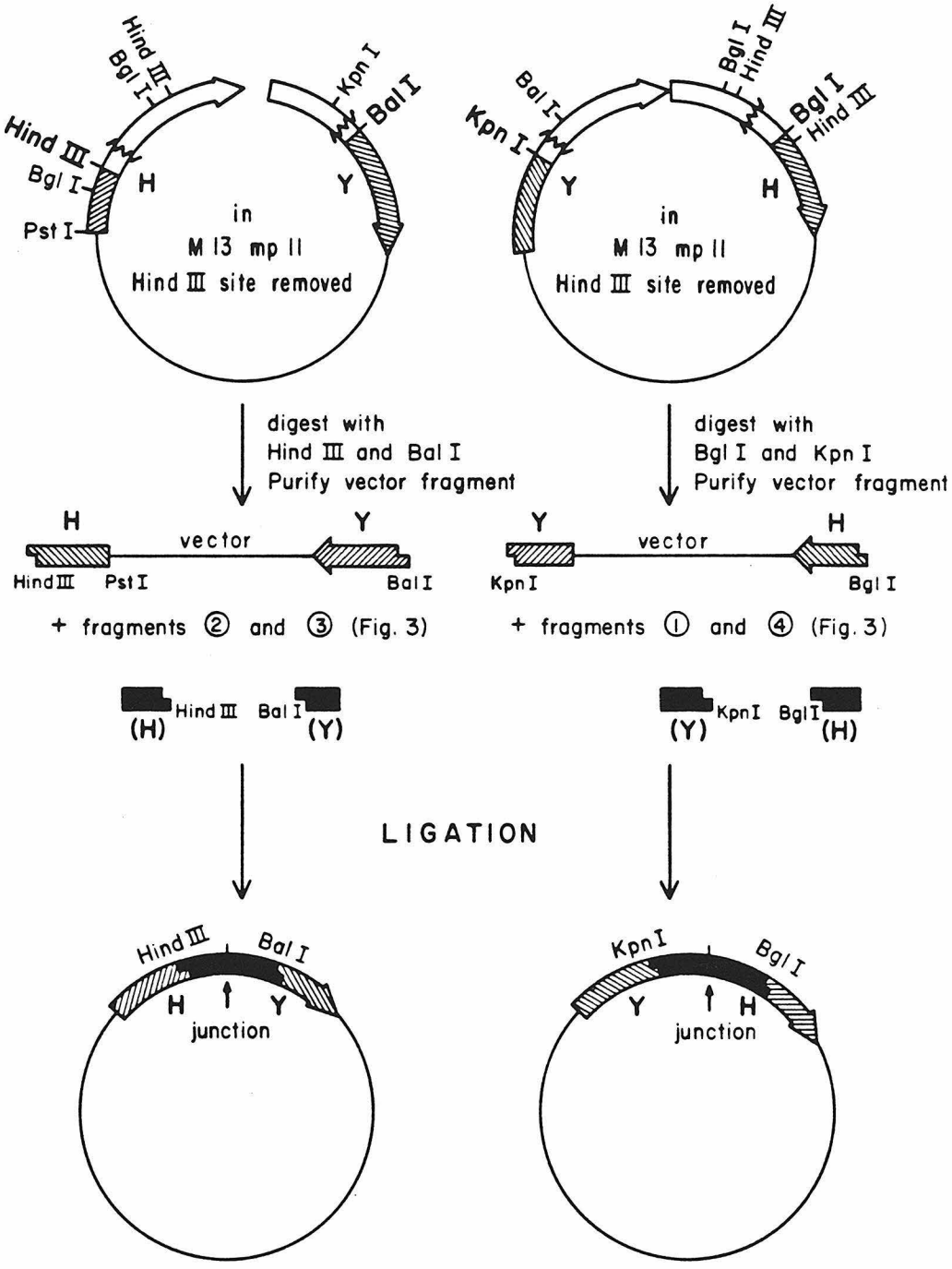
Figure 3 Primer editing reactions



the primer-editing reactions, have been reported to be efficient (24,25). The choices of restriction sites and fragments from each gene to be used in generating these two hybrids have been dictated primarily by practical considerations, such as the occurrence of target sequences in the human and yeast genes, and the availability of restriction enzymes from commercial sources. Evident alternatives do exist, but the proposed strategy is designed to utilize existing resources to their best advantage, using as few steps as possible to accomplish the desired objective. If having both the yeast and human PGK genes in the same vector resulted in stability problems because of homology in the DNA sequences (in spite of being in the same relative orientation), it would be simple to modify the strategy to overcome this by starting with vectors having single inserts.

Once the hybrid genes are obtained and their sequences verified, it will be necessary to transfer them to an expression vector. The yeast gene already has been adapted for expression in both yeast and *E. coli* (2,26) and work is underway to obtain expression of the human gene in *E. coli* (27) and in yeast (26). Therefore, once the N_YC_H and N_HC_Y hybrids are obtained it should be relatively straightforward to arrange for their expression

FIGURE 4


Hind III
Bgl I
Hind III
Bgl I
H
Pst I
Kpn I
Bal I
Y
in
M 13 mp 11
Hind III site removed
Bal I
Kpn I
Y
Bgl I
Hind III
Bgl I
Hind III
H
in
M 13 mp 11
Hind III site removed
digest with
Hind III and Bal I
Purify vector fragment
digest with
Bgl I and Kpn I
Purify vector fragment
H
vector
Y
Hind III
Pst I
Bal I
Y
vector
H
Kpn I
Bgl I
+ fragments ② and ③ (Fig. 3)
+ fragments ① and ④ (Fig. 3)
(H)
Hind III
Bal I
(Y)
(Y)
Kpn I
Bgl I
(H)
LIGATION
Hind III
Bal I
H
Y
junction
Kpn I
Bgl I
Y
H
junction

in pgk$^-$ strains (6) and for purification of the resulting proteins by adapting existing procedures. There are four methionines in the yeast enzyme and fourteen in the human enzyme; therefore, each hybrid protein would give rise to different sets of peptides, distinct from those generated by either of the parent proteins when treated with cyanogen bromide. This will be useful to confirm the sequences at the protein level.

If the resulting enzymes are active, the initial characterization should focus on determining kinetic parameters, stability to heat and denaturing agents, etc., in an effort to establish what differences and similarities exist between the hybrids and the parent enzymes. If the hybrid proteins are inactive, it would be more interesting initially to try to determine what changes in the structure (if any) would restore catalytic activity, although ligand binding and stability properties should also be determined to assess the extent of the structural disruption responsible for loss of activity.

In either case, the determination of the structure by X-ray crystallography should be undertaken. The detailed nature of the experiments that should be done to characterize the resulting proteins will be determined, to some extent, by their properties, and a further

description of the possibilities and experimental approaches designed to explore them are beyond the scope of this proposal.

PROPOSITION II

STUDY OF THE EFFECT OF VARIOUS SPECIFIC AMINOACID SUBSTITUTIONS ON THE MECHANISM OF DOMAIN MOVEMENT IN PHOSPHOGLYCERATE KINASE

The carboxyl group of glutamate 190 in helix V of the amino terminal domain of yeast PGK extends toward beta sheet N of the carboxyl domain, forming an interesting interaction with histidine 388. H. C. Watson et. al. proposed a "trigger mechanism" in which the cumulative effects of small conformational changes, which probably occur on binding of both substrates, weaken the interaction between glu190 and his388 thus initiating domain movement (3).

The importance of the interaction between glutamate 190 and histidine 388 in the mechanism of domain movement of phosphoglycerate kinase could be better understood if presumably non disruptive substitutions are introduced at one of the two sites. The proposed substitutions are glu190-->gln, glu190-->asp and glu190-->asn which would would probe the importance of the charged group and the length of the side chain in forming this interaction. The resolution of the crystallographic data (2.5 Angstroms) leaves the precise relative orientation of the two side chains of interest somewhat uncertain, but they have been represented as being hydrogen bonded (3). Visual inspection of a graphic representation of the PGK molecule in an Evans and Sutherland PS 300 graphics terminal revealed no other

residues in the waist region sufficiently close for any obvious interactions with either residue; if the glutamate residue is substituted in the computer by the shorter, less flexible aspartate, the carboxyl group can still come into close proximity with the histidine residue in at least one possible side chain conformation. This suports the idea that these changes would not disrupt the protein's structure.

Since these substitutions are likely to weaken the his-glu interaction, in the resulting mutant proteins domain movement would be facilitated, and this would have the effect of producing a more active enzyme if the proposed mechanism is correct.

The yeast PGK structural gene is contained in an Eco RI-HindIII restriction fragment which can be subcloned into the single stranded phage vector M13mp9 for the purposes of introducing the specific substitutiuons. Synthetic oligonucleotides can be used to mutate the GAA glutamate codon to CAA (gln), AAC (asn) or GAC (asp) by oligonucleotide-directed mutagenesis as described in Chapter II of this thesis. A computer search for homologies between the oligonucleotides and both the vector plus strand and the cloned fragment should be done; if this reveals homologous competing sites, a shorter fragment of the PGK gene and/or a different vector can be used.

After obtaining and confirming the mutations by DNA sequencing, the gene fragments should be introduced into a yeast expression vector and used to transform an appropriate pgk^- yeast host (6) in order to produce the mutant proteins.

Unless the aminoacid substitutions are destabilizing to the proteins, or alter their folding pathway making them more sensitive to proteolysis, these mutant enzymes should be active and produced in large quantities. If they are active, purification and determination of kinetic parameters should be undertaken in order to assess the nature and magnitude of the differences between the mutants and the wild type enzyme. Inactive mutants in this case would not be informative about domain movement, since they would probably be folding mutants, but they could be used to gain understanding of the requirements for proper folding, if it were possible to produce them in adequate amounts. If active mutants are obtained, they should be crystalized for structure determination; it is likely that crystals of active mutants would be isomorphous with those of the parent enzyme, simplifying the problem of structure determination.

REFERENCES

1. Scopes, R.K. (1973) in, THE ENZYMES, Vol. 8, 3rd ed., P.D. Boyer, ed., pp. 335-351, Academic Press, New York.

2. Blake, C.C.F. and Rice, D.W. (1981) Phosphoclycerate kinase. PHIL. TRANS. R. SOC. LOND. A. 293:93-104.

3. Watson, H.C., Walker, N.P.C., Shaw, P.J., Bryant, T.N., Wendell, P.L., Fothergill, L.A., Perkins, R.E., Conroy, S.C., Dobson, M.J., Tuite, M.F., Kinsgsman, A.J., and Kingsman, S.M. (1982) Sequence and structure of yeast phosphoglycerate kinase. EMBO JOUR. 1:1635-1640.

4. Banks, R.D., Blake, C.C.F., Evans, P.R., Haser, R., Rice, D.W., Hardy, G.W., Merrett, M., and Phillips, A.W. (1979) Sequence, structure, and activity of phosphoglycerate kinase: a possible hinge-bending enzyme. NATURE (London) 279:773-777.

5. Huang, I-Y., Welch, C., and Yoshida, A. (1980) Complete amino acid sequence of human phosphoglycerate kinase. J. BIOL. CHEM. 255:6412-6420.

6. Hitzeman, R.A., Clarke, L., and Carbon, J. (1980) Isolation and characterization of the yeast 3-phosphoglycerokinase gene (PGK) by an immunological screening technique. J. BIOL. CHEM. 255:12073-12080.

7. Hitzeman, R.A., Hagie, F.E., Hayflick, J.S., Chen, C.Y., Seeburg, P.H., and Derynck, R. (1982) The primary structure of the Saccharomyces cerevisiae 3-phosphoglycerate kinase. NUCL. ACIDS RES. 10:7791-7808.

8. Singer-Sam, J., Simmer, R.L., Keith, D.H., Shively, L., Teplitz, M., Itakura, K., Gartler, S.M., and Riggs, A.D. (1983) Isolation of a cDNA clone for human X-linked 3-phosphoglycerate kinase by use of synthetic oligodeoxyribonucleotides as a detection probe. PROC. NATL. ACAD. SCI. USA 80:802-806.

9. Michelson, A.M., Markham, A.F., and Orkin, S.H. (1983) Isolation and DNA sequence of a full-length cDNA clone for human X chromosome-encoded phosphoglycerate kinase. PROC. NATL. ACAD. SCI. USA 80:472-476.

10. Dalbadie-McFarland, G. and Richards, J.H. (1983) In vitro mutagenesis: powerful new techniques for studying structure-function relationships in proteins. In, ANNU. REPORTS MEDICINAL CHEMISTRY, Vol. 26.

11. Knowles, J.R. (1980) Enzyme-catalyzed phosphoryl transfer reactions. ANNU. REV. BIOCHEM. 49:877-919.

12. Anderson, C.M., Zucker, F.H., and Steitz, T.A. (1979) Space-filling models of kinase clefts and conformation changes. SCIENCE 204:375-380.

13. Roustan, C., Fattoum, A., Jeanneau, R., and Pradel, L-A. (1980) Yeast 3-phosphoglycerate kinase: sulfate and substrate binding, their effect on the conformational state of the enzyme. BIOCHEMISTRY 19:5168-5173.

14. Bacharach, A.D.E., Markland, F.S., Pellino, A., and Weber, B.H. (1977) Modification of yeast 3-phosphoglycerate kinase: isolation and sequence determination of a nitrated active-site peptide and isolation of a carboxyl-modified active-site peptide. BIOCHEM. BIOPHYS. RES. COMMUN. 74:165-171.

15. Meyer, M.C. and Westhead, E.W. (1976) Interaction of sulfate ion with a critical tyrosine residue in yeast phosphoglycerate kinase detected through the tetranitromethane reaction. FEBS LETTERS 71:25-28.

16. Pickover, C.A., McKay, D.B., Engelman, D.M., and Steitz, T.A. (1979) Substrate binding closes the cleft between the domains of yeast phosphoglycerate kinase. J. BIOL. CHEM. 254:11323-11329.

17. Scopes, R.K. (1978) The steady-state kinetics of yeast phosphoglycerate kinase. EUR. J. BIOCHEM. 85:503-516.

18. Tanswell, P., Westhead, E.W., and Williams, R.J.P. (1976) Nuclear-magnetic-resonance study of the active-site structure of yeast phosphoglycerate kinase. EUR. J. BIOCHEM. 63:249-262.

19. Huang, I-Y., Fujii, H., and Yoshida, A. (1980) Structure and function of normal and variant human phosphoglycerate kinase. HEMOGLOBIN 4:601-609.

20. Fujii, H. and Yoshida, A. (1980) Molecular abnormality of phosphoglycerate kinase-Uppsala associated with chronic nonspherocytic hemolytic anemia. PROC. NATL. ACAD. SCI. USA 77:5461-5465.

21. Blake, C.C.F. and Evans, P.R. (1974) Structure of horse muscle phosphoglycerate kinase. J. MOL. BIOL. 84:585-601.

22. Pegoraro, B. and Lee, C-Y. (1978) Purification and characterization of two isozymes of 3-phosphoglcerate kinase from the mouse. BIOCHIM. BIOPHYS. ACTA 522:423-433.

23. Brutlag, D. and Kornberg, A. (1972) Enzymatic synthesis of deoxyribonucleic acid. XXXVI. A proofreading function for the 3'-->5' exonuclease activity in deoxyribonucleic acid polymerases. J. BIOL. LCHEM. 247:241-248.

24. Goeddel, D.V., Shepard, H.M., Yelverton, E., Leung, D., Crea, R., Sloma, A., and Pestka, S. (1980) NUCL. ACIDS RES. 8:4057-4074.

25. Cabilly, et al., manuscript in preparation.

26. Hitzeman, R.A., personal communication (1983)

27. Singer-Sam, J. and Riggs, A.D., personal communication (1983)

PROPOSITION III

SOME APPROACHES TO INCREASE UNDERSTANDING OF THE MECHANISM OF TELOMERE REPLICATION

The replication of of the ends of linear eukaryotic chromosomes is a problem that has not been solved although several hypotheses have been put forth (1). Since all known DNA polymerases require a primer, when a linear DNA molecule is replicated, a 5' terminal gap is generated in the daughter strand when the primer is removed after initiation of synthesis. Several proposed models to explain the mode of replication of chromosomal ends are variations of the idea that palindromic sequences, or sequences capable of forming hairpin loops at the end of the replicating chromosome are involved, and that specific nicks are introduced at some point to resolve the daughter strands.

Progress in understanding the mode of telomere replication has been facilitated recently by the construction of self replicating linear plasmids with Tetrahymena macronucleus ribosomal DNA ends capable of functioning as telomeres in yeast (2), and the cloning of yeast telomeres using one such linear plasmid (2).

The structure of linear DNA termini appears to be very similar in many lower eukaryotes, and within a same organism, it is likely that all telomeric ends are identical

(1). The structure of Tetrahymena rDNA telomeric ends has been studied and found to contain tandem repeats of the hexanucleotide 5'-CCCCAA-3'; several single stranded, one nucleotide gaps were found at specific sites in the C_4A_2 strand and one in the complementary G_4T_2 strand. Interestingly, when these ends function as telomeres for linear plasmids in yeast, the position of these gaps is altered to a pattern characteristic of yeast telomeres, which is similar but not identical (2). No covalently, or tightly bound proteins have been isolated (3), but some mechanism for maintaining the specific gaps and replicating these ends must exist which relies on some specialized set of enzymatic activities. These enzymes will of necessity have affinity for the substrate, $(C_4A_2)_n$, and therefore it should be possible to isolate them by affinity chromatography using a similar method to that used by Nordheim et al. (4) to isolate Z DNA binding proteins.

In order to prepare an affinity column for this purpose it would be necessary to prepare a relatively large amount of DNA, so synthesizing both strands of an oligomer would probably be the more straightforward approach. The number of repeats of the hexanucleotide present in the synthetic oligomer would be limited by practical considerations and could not exceed six or seven without impacting significantly on the overall yield of the synthesis. The

number of repeats present in the Tetrahymena rDNA ends is of the order of 50, so by using a much shorter segment of DNA one runs the risk of not taking full advantage of all possible contacts between DNA and protein(s); however, it is hard to envision a situation in which a protein would not bind preferentially to a fragment of its normal substrate, unless a very complex tertiary structure of the DNA were required for binding. Since the proteins that would be interesting are those involved in mediating the replication of the termini, it seems reasonable to assume that this is not the case.

Tetrahymena extracts enriched for nuclear proteins could be fractionated on an affinity matrix derivatized with the synthetic DNA duplex and any proteins obtained in this way could be further purified in a second affinity fractionation step using carrier E. coli DNA in order to reduce contamination by proteins non specifically bound to DNA. Any fractions with C_4A_2 binding activity should be assayed for nuclease activity, the ability to excise nucleotides from double stranded DNA, and their ability to alter the electrophoretic mobility of DNA, using both random sequence DNA and $(C_4A_2)_n$ as substrates, as an initial form of characterization. Further biochemical characterization would depend on preliminary findings.

In this way, by identifying some of the enzymatic activities responsible for telomere replication it would be

possible to refine the existing models.

Using synthetic DNA for this experiment would have the added advantage of providing material for physical and chemical characterization of this CCCCAA sequence to determine whether it dictates some unusual sequence dependent secondary or tertiary structural characteristic.

REFERENCES

1. Blackburn, E. H. and Szostak, J. W., Ann. Rev. Biochem. (1984) 53: 163-94.

2. Szostak, J. W. and Blackburn, E. H., Cell 29: 245-255 (1982).

3. Blackburn, E. H., Budarf, M. L., Challoner, P. B., Cherry, J. M., Howard E. A., et al. Cold Spring Harbor Symp. Quant . Biol., 47: 1195-1207.

4. Nordheim, A., Tesser, P., Azorin, F., Young, H. K., Moller, A., and Rich, A., Proc. Natl. Acad. Sci. U.S.A. 79: 7729-7733 (1982).

PROPOSITION IV

STUDY OF THE POSSIBLE ROLE OF Z-DNA IN DEVELOPMENT

The salt induced, cooperative helix-helix transition that the alternating deoxypolynucleotide poly d(GC)·poly d(GC) undergoes in solution was first described by Pohl and Jovin (1972) (1). The transition can be induced by raising the concentration of NaCl (transition midpoint 2.7 M), $MgCl_2$ (trasition midpoint 0.8 M) or by adding ethanol to a final concentration of 60% v/v. It produces virtually complete inversion of the CD spectrum (1,2) and substantial changes in the ORD (1) and laser Raman spectra (3) of the alternating copolymer.

The characteristics of the high salt form of oligo d(GC) were studied further by Patel et al. (1979) (4) using proton and ^{31}P NMR solution methods in an attempt to elucidate structural aspects of the conformation. The NMR data yielded information on the nature of the conformational transition and some major features of the high salt helix. It was shown to exhibit an A-DNA sugar pucker and glycosidic torsion angle at the purine residues, alternating with a B-DNA type sugar pucker and glycosidic torsion angle at the pyrimidine residues. The phosphate backbone formed by two types of alternating non-equivalent phosphates gave rise to two ^{31}P resonances, in contrast

to the single peak observed in the low salt, B conformation.

When the crystal structures of $d(CpG)_3$ and $d(CpG)_2$ were solved (5,6,7) a novel left-handed conformation for DNA emerged: the atomic resolution data gave detailed information on the structure of the self-complementary oligomers as well as the organization of water molecules and ions found in the crystal structure. It became apparent that there is a family of closely related left-handed DNA conformations which can differ in specific ways, but which share certain major characteristics (8). In general, the structure is an antiparallel double helix with Watson-Crick base pairing and a left-handed helical sense. The guanine residues adopt a syn conformation relative to the corresponding deoxyribose, which is found in a C3' endo conformation. The cytsoine residues on the other hand are in the familiar anti configuration, associated with a deoxyribose that has C2' endo pucker as in B-DNA. The zigzag course pursued by the ribose-phosphate backbone led to the name Z-DNA.

Right-handed B-DNA has a major and minor groove, but there is only a single groove in Z-DNA, analogous to the B-DNA minor groove. The outer, convex surface of Z-DNA is analogous to the concave major groove of B-DNA. The G-C base pair forms part of the outer surface of left-handed DNA, with the bases far more accessible to solvent and

ligands than in B-DNA: it has been shown that bromination of the C8 of guanosine as well as methylation of cytosine C5 stabilize the Z conformations (9,10).

The assignment of Z conformation to the high salt form of alternating d(GC) has been made on the basis of the solution NMR data mentioned above, and a recent study of the laser-Raman spectra of crystals and solutions of $d(CpG)_3$ and solutions of poly d(GC) (11). This conformation does not appear to be restricted to alternating d(GC) since there is evidence that other alternating purine-pyrimidine sequences can form left-handed helices. A Z-type conformation was observed for poly d(AC)·poly d(GT) by fiber diffraction methods (12) and changes in the CD spectrum of the same polymer can be induced in solution at high concentrations of CsCl or CsF (13,14).

The potential role in biological systems of a reversible, sequence dependent conformational change in DNA could be extensive since it represents an additional mode of encoding biological information in the nucleotide sequence of DNA. Varied evidence supporting this idea is rapidly accumulating. Cytological data indicate the presence of DNA segments with Z forming potential in several eukaryotic nuclei in species of Drosophila (9), Chironomus, Stylonichia, and several rat tissues (15, 16); proteins that bind Z-DNA specifically, have been isolated from Drosophila (17) and from SV40 minichromosomes (15); two Z-compatible se-

quences in the transcriptional enhancer region of SV40 appear to be required for enhancer function (18). The subject has been reviewed by Rich et al. (15) and these are only some of the more striking examples that indicate that Z-DNA does indeed have a biological role. How extensive a role it plays in the control of gene expression is yet undetermined and it would be interesting to design experiments to address that question.

In a cytological study of several different rat tissues (16) using specific anti Z-DNA antibodies, it was found that some nuclei are intensely stained whereas others are not. In particular, in the seminiferous tubules of the testes, where several developmental stages in the maturation of spermatozoa are observed, only certain cell types at a particular stage of differentiation are stained with labeled anti Z-DNA antibodies. Although the significance of results obtained by staining fixed tissues should be interpreted with caution, this suggests a possible role for Z-DNA in development, and it would be interesting to find out whether this is the case by an independent approach.

One can argue that in order to stabilize DNA in a Z conformation in some cells and not in others Z-DNA binding proteins are differentially expressed in different cell types. Since a number of proteins that specifically bind Z-DNA have been isolated from Drosophila (17), it would be feasible to clone their genes in order to use them as hy-

bridization probes and investigate their expression at different stages in development. Either protein sequence information or antibodies against the Z-binding proteins should be available soon.

If antibodies against Z-binding protein(s) become available first, then the method of choice would be to screen an expression library, constructed using a vector such as lambda gtll (21).

With protein sequence data, synthetic oligonucleotide probes (19) can be designed to screen a cDNA library. If only very highly degenerate mixed oligomeric probes are possible due to the protein sequence, a longer (>50 bp), therefore more specific, unique sequence probe can be designed by chosing a single base at degenerate sites based on the preferred codon usage of the organism from which the library was made (20). Some "guesses" will be incorrect, but the added length of the probe should compensate for this, with the added benefit of greatly reducing the probability of obtaining "false positives", a problem commonly encountered when using short, mixed sequence probes.

Once the gene is cloned and its identity is confirmed by comparison of its sequence to that of the protein, it can be used to study the expression of its product in different tissues and at different stages of differentiation, as well as to study the structure of the gene(s) in the genomic DNA.

REFERENCES

1. Pohl F. M., and Jovin, T. M., J. Mol. Biol. 67, 375-396 (1972).

2. Pohl, F. M., Nature 260, 365-366 (1976).

3. Pohl, F. M., Ranade, A. and Stockburger, M., Biochim. et Biophys. Acta 335, 85-92 (1973).

4. Patel, D. J., Canuel, L. L. and Pohl, F. M., Proc. Natl. Acad. Sci. USA 76, 2508-2511 (1979).

5. Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Mare, G. and Rich, A., Nature 262, 680-686 (1979).

6. Drew, H. R., Takano, T., Tanaka, S., Itakura, K. and Dickerson, R. E., Nature 286, 567 (1980).

7. Crawford, J. l., Kolpak, F. J., Wang, A. H.-J., Quigley, G. J., van Boom, J. H., van der Marel, G. and Rich, A., Proc. Natl. Acad. Sci, USA 77, 4016-4020 (1980).

8. Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., van der Marel, G., van Boom, J. H. and Rich, A., Science 211, 171-176 (1981).

9. Nordheim, A., Pardue, M. L., Lafer, E. M., Moller, A., Stollar, B. D. and Rich, A., Nature 294, 417-422 (1981).

10. Behe, M. and Felsenfeld, G., Proc. Natl. Acad. Sci. USA 78, 1619-1628 (1981).

11. Thamann, T. J., Lord, R. C., Wang, A. H.-J and Rich, A., Nucl. Ac. Res. 9, 5443 (1981).

12. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W., and Ratliff, R. L., Nature 283, 743-745 (1980).

13. Vorlickova, M., Kypr, J., Stokrova, S. and Sponar, J., Nucl. Ac. Res. 10, 1071 (1982).

14. Zimmer, C., Tymen, S., Marck, C. and Juschlbauer, W., Nucl. Ac. Res. 10, 1081 (1982).

15. Rich, A., Nordheim, A. and Wang, A. H.-J., Ann. Rev. Biochem. 53, 791-846 (1984).

16. Morgenegg, G., Celio, M. R., Malfoy, B., Leng, M. and Kuenzle, C. C., Nature 303, 540-43 (1983).

17. Nordheim, A., Tesser, P., Azorin, F., Ha Kwon, Y., Moller, A. and Rich, A., Proc. Natl. Acad. Sci. USA 79, 7729-7733, 1982.

18. Khouri, G. , Gruss, P., Cell 33, 313-314 (1983)

19. Itakura, K., Rossi, J. J. and Wallace, R. B., Ann. Rev. Biochem. 53, 323-56 (1984).

20. Anderson, S. and Kingston, I. B., Proc. Natl. Acad. Sci. USA 80, 6838-6842, 1983.

21. Young, R. A. and Davis, R. W., Proc. Natl. Acad. Sci. USA 80, 1194-1198, 1983.

PROPOSITION V

A STUDY OF UNFOLDING AND REFOLDING OF E. COLI BETA LACTAMASE

The mechanism of folding and unfolding of S. aureus beta-lactamase has been studied by a variety of methods, including circular dichroism, optical rotatory dispersion (1, 2), NMR (3), and UV difference spectroscopy (4, 5), using several denaturants such as urea, guanidinium chloride and some biguanide salts.

These studies led to the identification of a stable unfolding intermediate predominant at intermediate concentrations of urea or guanidinium chloride, which is significantly more expanded than the native state, but which has a substantial content of secondary structure (1, 2). NMR data indicate that the residues in the N terminal third of the molecule are affected by the transition from the native state N, to the intermediate state H (4), and a multidomain model has been proposed for the structure of this protein which would allow a large change in shape in the intermediate state H if the domains or folding regions separate without themselves becoming unfolded (2). Additional studies by urea-gradient polyacrylamide gel electrophoresis (6) confirm the existence of state H and indicate that unfolding proceeds by a two-state transition to fully unfolded protein, but that refolding is complex and involves the formation of the partially folded intermediate H (7).

Urea-gradient electrophoresis is a useful technique that can yield a substantial amount of information regarding the mode of folding and unfolding of a protein (6). A band of protein is subjected to electrophoresis through a gel slab in which a transverse gradient of denaturant has been established; thus a continuous band of protein migrates in a direction that is perpendicular to the gradient. As the continuous band of protein moves electrophoretically under gradually varying conditions, a pattern is generated that can be interpreted in terms of the variation of protein conformation with the changing conditions. The technique is particularly useful in identifying conformational states which are interconverted slowly relative to the duration of electrophoresis; some of the limitations are that this method may not be sensitive to small conformational changes, and that not all conformational transitions will produce a measurable difference in mobility. It has proven useful in the case of S. aureus beta-lactamase (7) which unfolds and refolds slowly at room temperature (1).

S. aureus beta lactamase is a class A enzyme, according to Ambler's classification (8), whose sequence is 30% homologous to that of the E. coli beta-lactamase. Their three dimensional structures are likely to be similar, particularly since 19% of the residues are completely conserved in these and two other class A beta lactamases (8). It would

be interesting therefore to find out whether the unfolding and refolding of E. coli beta-lactamase is similar to that of the S. aureus enzyme. The latter has no disulfide bridges, therefore the initial comparison should be established with the mutant of E. coli beta lactamase that lacks the single disulfide bond present in the wild type enzyme (Chapters II and IV of this work).

This mutant E. coli beta-lactamase could be analyzed by urea-gradient electrophoresis, starting with either native enzyme or completely denatured enzyme, in similar conditions to those used for the S. aureus enzyme. If evidence is found of similar behavior, the interpretation of results would be greatly aided by the large amount of information already available for the S. aureus enzyme.

Subsequently, similar experiments could be carried out to compare the behavior of this mutant beta-lactamase to that of the wild type enzyme with its disulfide bond intact, to determine how it influences the unfolding and in particular the refolding of the enzyme. Aquiring this kind of information about E. coli beta lactamase could help design experiments to understand the cause of the instability observed in some of the other mutant beta-lactamases described in this thesis. Spectroscopic studies would probably be needed in addition, but this method appears to be a good initial step toward understanding the effect of

various mutations on the folding pathway of E. coli beta-lactamase.

REFERENCES

1. Robson, B. and Pain, R. H., Biochem. J. 155, 325-330; 331-344 (1976).

2. Carrey, E. A. and Pain, R. H., Biochimica et Biophysica Acta. 533, 12-22 (1978).

3. Thomas, R. M., Feeney, J., Nicholson, R. N., Pain, R. H. and Roberts, G. C. K., Biochem. J. (1983) 215, 525-529.

4. Robson, B. and Pain, R. H., Biochem. J. (1976) 155, 331-344.

5. Mitchinson, C., Pain, R. H., Vinson, J. R. and Walker, T., Biochimica et Biophysica Acta 743, 31-36 (1983).

6. Goldenberg, D. P. and Creighton, T. E., Analytical Biochemistry 138, 1-18 (1984).

7. Creighton, T. E. and Pain, R. H., J. Mol. Biol. 137, 431-436 (1980).

8. Ambler, R. P., Phil. Trans. R. Soc. Lond. B 289, 321-331 (1980).