PROTEINS: Structure, Function, and Genetics 14:29–44 (1992)

Probing β-Lactamase Structure and Function Using Random Replacement Mutagenesis

Timothy Palzkill and David Botstein
Department of Genetics, School of Medicine, Stanford University, Stanford, California 94305

ABSTRACT **A new analytical mutagenesis technique is described that involves randomizing the DNA sequence of a short stretch of a gene (3–6 codons) and determining the percentage of all possible random sequences that produce a functional protein. A low percentage of functional random sequences in a complete library of random substitutions indicates that the region mutagenized is important for the structure and/or function of the protein. Repeating the mutagenesis over many regions throughout a protein gives a global perspective of which amino acid sequences in a protein are critical. We applied this method to 66 codons of the gene encoding TEM-1 β-lactamase in 19 separate experiments. We found that TEM-1 β-lactamase is extremely tolerant of amino acid substitutions: on average, 44% of all mutants with random substitutions function and 20% of the substitutions are expressed, secreted, and fold well enough to function at levels similar to those for the wild-type enzyme. We also found a few exceptional regions where only a few random sequences function. Examination of the X-ray structures of homologous β-lactamases indicates that the regions most sensitive to substitution are in the vicinity of the active site pocket or buried in the hydrophobic core of the protein. DNA sequence analysis of functional random sequences has been used to obtain more detailed information about the amino acid sequence requirements for several regions and this information has been compared to sequence conservation among several related β-lactamases.**

Key words: in vitro mutagenesis, antibiotic resistance, protein structure-function

INTRODUCTION

A major goal of protein structure and function studies is to identify amino acid residue positions critical to the structure and/or function of the protein. The central difficulty is that not all residue positions contribute equally to the structure and function of a protein, i.e., they do not have the same "information content."[1–3] At some residue positions the precise chemical characteristics of a side chain are essential for function, and therefore no other amino acid will substitute for the wild-type residue.[4–6] In contrast, at other positions a specific side chain characteristic is not necessary for either structure or function and a wide range of different amino acids at that position can produce a functional protein. It is those residue positions with stringent side chain requirements that contribute most critically to the structure and function of the protein[4–7] and are thus said to have high information content.[1]

A functional selection approach to the problem of determining which are the most critical amino acid residues has been described by several groups.[1,8–12] The strategy consists of randomizing one to three positions by site-directed mutagenesis methods, selecting for functional protein, and then sequencing to determine the identity of the allowable substitutions at each position. This allows the importance of individual positions to be readily assessed.

In this study we describe a conceptually similar but methodologically unique technique, random replacement mutagenesis that makes practical assessments of the information content in many different regions of a protein's amino acid sequence, allowing inference of which regions are important determinants of the protein's structure and/or function. The strategy involves randomizing three to six contiguous codons in the gene of interest to form a "library" containing all or nearly all possible substitutions for the region randomized. The percentage of random sequences that results in a functional protein is then assessed by an in vivo assay. If only a few random sequences provide a functional protein, then the region randomized has a high information content and we infer it must be important to the structure and/or function of the protein. In contrast, if a large percentage of sequences can function, then the region is relatively unimportant from both the structural and functional point of view. By repeating the process several times in different regions of the gene, it is

Received August 15, 1991; revision accepted October 25, 1991.

Address reprint requests to Dr. David Botstein, Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305.

possible to obtain an estimate of what fraction of all residue positions are actually important for the structure and function of the protein as well as where those important regions lie. Detailed information about the sequence requirements of individual residues is obtained by sequencing functional substitutions. The amino acid side chain characteristic(s) critical for function should be conserved among the set of functional substitutions.

We have applied the method to TEM-1 β-lactamase.[13] β-lactamases provide for bacterial resistance to β-lactam antibiotics such as the penicillins and cephalosporins by catalysing the hydrolysis of the β-lactam ring and generating inactive products. A large number of different β-lactamases have been described and their presence in bacterial pathogens represents a serious threat to antibiotic therapy.[14] Based on primary sequence homology, β-lactamases have been grouped into four classes.[15] Classes A, C, and D involve a serine residue in the active site.[16,17] Class B enzymes are less abundant and require a catalytic zinc for activity.[18] The active site serine β-lactamases belong to a larger family of penicillin-recognizing enzymes that includes the penicillin binding proteins (PBPs), which are the lethal targets of β-lactam antibiotics.[19] All of these enzymes contain the active site serine and a conserved lysine three residues downstream from the serine as well as a conserved triad of K(S/T)G between the active site serine and the C-terminus.[19] The homology suggests a similar enzyme mechanism and is consistent with the suggestion that β-lactamases have evolved from the cell wall peptidases.[20] The crystal structures of several class A enzymes,[21–24] a class C enzyme[25] and a cell wall transpeptidase/carboxypeptidase[26] are available, and they have similar three-dimensional structures that also supports a common evolutionary origin for the enzymes.

The class A β-lactamases are plasmid-encoded and are widely distributed in both Gram-negative and Gram-positive bacteria.[14,27] High resolution crystal structures of the class A *Staphylococcus aureus* and *Bacillus licheniformis* β-lactamases have located the active site.[22,24] The results suggest that the catalytic residues include Glu-166 and Lys-234 in addition to Ser-70 and Lys-73; numbering according to Ambler.[15] The amino acid sequences of 14 class A β-lactamases have been aligned and the proposed catalytic residues are conserved.[28] Site-directed mutagenesis on each of the residues also supports their proposed role in catalysis.[29–32]

TEM-1 β-lactamase is a class A enzyme encoded by the *bla* gene, which is present on the Tn3 transposon.[33] TEM-1 is an attractive enzyme for genetic studies because a selection for function exists that allows the easy detection of functional substitutions among a large population of nonfunctional mutants. Although a high-resolution structure of the TEM-1 enzyme has not yet been reported, its homology to the class A *Staphylococcus aureus* and *Bacillus licheniformis* β-lactamases allows a correlation of the sensitivity of a region with its position in the three-dimensional structure.[22,24] Finally, TEM-1 is the most common β-lactamase in Gram-negative bacteria and thus is a clinically important enzyme.[27,34,35] The information gained on structure-activity relationships for TEM-1 β-lactamase is therefore relevant to the design of β-lactamase resistant antibiotics.

MATERIALS AND METHODS

Materials

Oligonucleotides used for construction of the "excision" linkers, random linkers, primers used for DNA sequencing, and primers for PCR were custom synthesized by the oligonucleotide synthesis facility at Genentech. Plasmid pJ1519 was a gift from J. Wells.[36] Plasmid pBG66 was used in all experiments described here. It is a derivative of pJ1519 created by deleting a 210 bp BglI DNA fragment.

Bacterial Strains

E. coli strain HB101 (hsdS20 (rB^-, mB^-), supE44, ara14, λ^-, galK2, lacY1, proA2, rspL20, xyl-5, mtl-1, recA13, mcrA(+), mcrB(−)) was used for random library constructions, and TG1 (Δ[lac-pro], supE, thi, hsdDS/F'traD36, $proA^+B^+$, lacI, lacZM15) was used for assaying ampicillin resistance and for single-strand DNA preparations.

Media and Reagents

E. coli strain HB101 was grown in LB medium.[37] *E. coli* strain TG1 was maintained on glucose minimal medium plates and growth in liquid culture for single-strand DNA isolation was in M9 minimal medium and also 2X YT medium.[37] All enzymes were purchased from New England Biolabs, except calf intestinal phosphatase (Boeringer Manniheim) and DNAseI (Worthington). XGal was obtained from Clonetech.

Plasmid Isolation

Plasmid DNA was prepared from *E. coli* by the alkaline lysis procedure.[38] The DNA was purified in cesium chloride step gradients.[39] Plasmid minipreps were made from 1.5 ml cultures using the alkaline lysis protocol of Birnboim and Doly.[40] Single-strand plasmid DNA was prepared for DNA sequencing as described in Maniatis et al.[41]

DNA Sequencing Reactions

The dideoxy chain termination method of DNA sequencing[42] was applied to single-strand plasmid DNA templates. Custom synthesized oligonucleotides were used to prime synthesis from defined sites within the *bla* gene.

A

```
                          EcoRI       BspMI
5'-GCAGGTCTGCAGGAATTCCTCGAGACCTGC-3'
3'-CGTCCAGACGTCCTTAAGGAGCTCTGGACG-5'
   BspMI
```

B

```
                          EcoRI       BspMI
N      NNNNNNNNGCAGGTCTGCAGGAATTCCTCGAGACCTGCNNNN      NNNNN
NNNNN      NNNNCGTCCAGACGTCCTTAAGGAGCTCTGGACGNNNNNNNNN      N
               BspMI
```

Fig. 1. **A.** Structure of the "excision linker" used in the first step of mutagenesis procedure to create defined deletions in the *bla* gene. **B.** Illustration of the digestion pattern of the BspM1 enzyme, which results in a defined DNA deletion after release of the "excision linker" from the *bla* gene. N represents any nucleotide.

Generation of Linker Insertions in *bla*

The first step of the mutagenesis procedure was to create a collection of individual linker inserts. The L37-42, L157-160, L182-186, and L196-200 (see Fig. 4) "excision linker" inserts were inserted randomly after DNAseI cleavage as described by Heffron et al.[43] The L72-74, L134-136, L146-149, L224-227, L250-252, L259-262 and L286-289 (see Fig. 4) inserts were made by ligating the linker to plasmid that had been linearized by partial restriction endonuclease digestion.[44] The L69-72, L102-104, L130-132, L168-170, L233-235, L235-237, and L238-241 linker inserts were made by site-specific insertion using the Polymerase Chain Reaction.[45] The method involved synthesizing PCR primers that anneal on opposite strands with their 5′ ends at the same nucleotide position. This position defined the insertion point. At the 5′ end of the primers was an additional nonhomologous "tail" that consists of one strand of the "excision linker" (Fig. 1). A second set of PCR primers was paired with the first set, one for each of the "tailed" primers. These primers lay on either side of the *bla* gene and were used in conjunction with the "tailed" primers in the PCR reactions to create two DNA fragments. At one end of each fragment was one-half of the "excision linker" to be inserted and at the other end was a restriction site that is 5′ or 3′ of the *bla* gene depending on the DNA fragment (the two PCR fragments overlap the 5′ and 3′ ends of the *bla* gene, respectively). Digestion of the PCR fragments with the appropriate 5′ or 3′ restriction enzyme was then performed and the fragments were gel purified.[38] The pBG66 plasmid was digested with the same 5′ and 3′ enzymes and the vector fragment was gel purified.[38] A 3-way ligation between the cleaved vector fragment and the PCR DNA fragments was then performed to create the complete pBG66 plasmid with the linker inserted at the predetermined position. These inserts have the same structure as the inserts generated by other methods. All of the PCR-generated linker insertions were targeted to the active site region.

Generation of Random Replacement Libraries

Plasmids containing the excision linker were purified on cesium chloride gradients. Each individual linker insertion was taken through the following procedure; 1.5 μg of linker insert containing plasmid was digested with 2 units of the restriction endonuclease BspM1 in 50 mM KCl, 10 mM Tris-Cl (pH 7.5), 10 mM $MgCl_2$, and 100 μg/ml bovine serum albumin (BSA) in a total volume of 40 μl at 37°C overnight (~16 hr). After the digestion, 1 unit of calf intestine phosphatase (CIP) was added and the reaction incubated at 37°C for 30 min. CIP was inactivated by incubation at 70°C for 15 min. To the reaction mix dNTPs were then added to a concentration of .05 mM along with 10 units of Klenow DNA polymerase and the reaction was incubated at 15°C for 30 min. The reaction products were then separated on a 4% acrylamide gel. The linearized plasmid was visualized after ethidium bromide staining and the band was cut from the gel and electroeluted in dialysis tubing.[38] The eluted DNA was precipitated in ethanol and resuspended in 10 μl TE. A 50-fold molar excess of a linker with random DNA at its ends (R-lac linker) was then ligated to the total eluted plasmid DNA (~1.2 μg) overnight at 23°C in 10 mM Tris-Cl (pH 7.5), 5 mM $MgCl_2$, 1 nM ATP, 20 mM DTT, 100 μg/ml BSA, and 100 units of T4 DNA ligase in a 50-μl total volume. The reaction was inactivated by incubation at 70°C for 10 min. The reaction mix was adjusted to 150 mM NaCl and 80 units of XhoI was added. The reaction mix was incubated at 37°C for 2hr. There is a XhoI site embedded in the original linker so this digestion ensured

TABLE I. Sequence Complexity of Random Libraries*

Random library	Total possible replacements	Number of colonies pooled	Probability least common %	Probability most common %
L37-42	9.6×10^6	2.3×10^4	<0.1	15.3
L69-71	2.8×10^3	1.6×10^4	61.5	99.9
L72-74	1.6×10^3	1.6×10^4	62.5	99.9
L102-104	6.4×10^3	4.8×10^3	8.0	99.4
L130-132	6.4×10^3	2.1×10^4	27.5	99.9
L134-136	6.0×10^3	1.1×10^4	80.8	99.9
L146-149	3.2×10^3	1.7×10^5	99.5	99.9
L157-160	3.2×10^4	6.0×10^4	5.7	99.9
L166-168	5.2×10^5	1.4×10^5	89.4	99.9
L168-170	2.0×10^3	2.1×10^4	47.9	99.9
L182-186	2.2×10^5	3.4×10^4	0.2	92.9
L196-200	8.0×10^5	1.5×10^4	<0.1	55.1
L224-227	6.4×10^3	4.0×10^4	91.1	99.9
L233-235	1.6×10^3	7.4×10^4	98.9	99.9
L235-237	2.4×10^3	2.5×10^5	97.7	99.9
L238-241	6.4×10^3	1.8×10^5	93.2	99.9
L250-252	1.6×10^3	7.8×10^3	38.1	99.9
L259-262	1.6×10^3	1.3×10^4	53.6	99.9
L286-289	1.6×10^3	5.6×10^2	3.3	70.7

*Table shows the data used to calculate the probability that the random libraries contain all possible amino acid substitutions. Details of the calculations are in Materials and Methods.

that any plasmid that did not digest with BspM1 was not included among the R-lac linker inserts. The XhoI digestion was phenol:chloroform extracted, ethanol precipitated and resuspended in 4 μl TE. This DNA was used to electroporate *E. coli* HB101 using a Gene Pulser apparatus (Bio-Rad). The efficiency of transformation using this method was routinely 5×10^8 transformants per μg plasmid DNA. The transformants were selected on LB plates containing 12.5 μg/ml chloramphenicol and 100 μg/ml XGal. The R-lac linker contains a *lacO* site so transformants, which contained an R-lac linker were blue.[46] White colonies represented recircularized plasmids without the R-lac linker. This was confirmed by isolating plasmid DNA from blue and white colonies and checking for the R-lac linker by restriction enzyme digests. In all experiments blue colonies constituted 95% or more of the transformants. All transformants were pooled together and plasmid DNA was isolated and purified on a cesium chloride gradient.

Two μg of pooled, R-lac linker-containing plasmid was digested with 2 units of BspM1, under the same buffer conditions as those described above, overnight at 37°C in 50 μl total volume. dNTPs were added to .05 mM and 10 units of Klenow DNA polymerase was added and incubated at 15°C for 30 min. The reaction products were separated on a 4% polyacrylamide gel and the plasmid band was visualized by ethidium bromide staining. The plasmid band was cut from the gel and the DNA was isolated by electroelution.[38] The eluate was extracted with phenol:chloroform, precipitated with ethanol, and resuspended in 10 μl TE. The digested, blunt-ended, gel-purified plasmid was recircularized in ligation conditions as described above. The reaction was heat inactivated at 70°C for 10 min and adjusted to 50 mM KCl. 50 units of NruI were then added and the digest was incubated at 37°C for 2 hr. An NruI site is embedded in the R-lac linker; therefore this digestion eliminated any R-lac linker containing plasmids that were not digested by BspM1. The reaction mix was then extracted with phenol:chloroform, precipitated with ethanol, and resuspended in 4 μl TE. The recircularized DNA was electroporated into *E. coli* HB101 and plated on LB plates containing 12.5 μg/ml chloramphenicol and 100 μg/ml XGal. More than 95% of the colonies in all experiments were white, indicating that the *lacO* portion of the R-lac linker had been released, leaving only the random DNA replacement. Many transformants (from 10^3 to 10^5) (Table I) were pooled and plasmid DNA was extracted and purified on a cesium chloride gradient. This plasmid collection was the final random substitution library, which contains a large collection of random DNA substituting for *bla* DNA.

Selection of Functional Replacement Mutants

To select for functional replacements, the library containing random DNA in a defined region was used to transform *E. coli* strain TG1 by a $CaCl_2$ procedure.[38] The transformation mixture was plated on LB plates containing, respectively, 1 mg/ml ampicillin, 10 μg/ml ampicillin, or 12.5 μg/ml chloramphenicol.

The L72-74 experiment (Table I) is used as an example for the calculations. The percent functional

replacements at 10 μg/ml and 1 mg/ml ampicillin for each experiment was calculated by dividing the number of transformants/ml resistant to 10 μg/ml (or 1 mg/ml) ampicillin by the number of transformants/ml resistant to 12.5 μg/ml chloramphenicol after adjusting the chloramphenicol number to compensate for the number of replacements containing a STOP codon. For L72-74, the probability of not having a STOP codon is [(61/64) × (61/64) × (1)] = 0.9. The 1 for the third position reflects the fact that only the first nucleotide of the codon was randomized, which eliminates, in this case, the possibility of a STOP codon. The adjusted number of chloramphenicol transformants/ml for L72-74 is given by 0.9 × number of chloramphenicol transformants/ml.

The conversion of the overall percent functional replacements to the percent functional per codon was calculated by taking the n^{th} root of the absolute fraction of functional replacements where n equals the number of codons randomized. For L72-74, n = 2.2. The fraction reflects only one nucleotide of the third position being randomized, which results in only four different amino acids being sampled, giving 4/20 = 0.2. The total possible number of different amino sequences for a replacement experiment was calculated based on the number of different amino acids sampled at each codon randomized. For L72-74 the calculation is 20 × 20 × 4 = 1600 different replacement mutants possible. The 4 again reflects partial mutagenesis of the third position.

The pool size in Table I is the number of different random substitution mutants present in the library. It is simply the number of R-lac linker containing blue colonies pooled. The number of white colonies pooled after the release of the R-lac linker is also important because if it is less than the pool size above, it will limit the library size. However, in all the experiments performed, the number of white colonies pooled in the final step was at least three times the blue colony pool size. The probability that the pool size in each experiment was large enough to contain the most probable (i.e., Leu Leu Leu) and least probable (i.e., Trp Trp Trp) replacement mutants (Table I) was calculated using the Poisson distribution $P = \lambda^x e^{-\lambda}/x!$. For these calculations $\lambda = np$, where n = pool size, p = probability of the least or the most common replacement, and x = number of times the sequence occurs in pool size n. For these calculations x = 0. The probability that the given sequence occurs is then $1-e^{-np}$; i.e., the probability that the sequence occurs one or more times in the pool.

The average percentage of functional replacements for the 66 codons mutagenized was calculated by adding the percent functional replacements per codon numbers for each of the 66 codons and dividing by 66.

RESULTS AND DISCUSSION

Random Replacement Mutagenesis

To randomize three to six codons of a gene as a unit a mutagenesis method was developed based on the properties of the restriction endonuclease BspM1, which is a class-IIS restriction enzyme that recognizes a non-palindromic, double-strand nucleotide sequence and cleaves both strands at a precise distance from the recognition sequence.[47] In the first step of the procedure, an "excision linker" was constructed containing BspM1 recognition sites at each end (Fig. 1A).[48] The linker was so named because when inserted into a gene and then released by BspM1 digestion, it creates a defined deletion in the gene (Fig. 1B).[48] The mutagenesis experiments were begun by generating a collection of over 100 individual "excision linker" insertions into the *bla* gene (Fig. 2A). The linker insertions were generated by three methods: linker insertion mutagenesis with DNAseI,[43] linker insertion following partial restriction enzyme digestions,[44] and with a site-specific method using the polymerase chain reaction (see Materials and Methods). The position of individual linker inserts was initially restriction mapped using the EcoRI site present in the "excision linker" (Fig. 1A). The precise position of the DNAseI generated inserts was determined by DNA sequencing. It was found that those inserts generated after DNAseI treatment (L37-42, L157-160, L182-186, L196-200) were not simple insertions but rather contained small additional adjacent deletions of *bla* DNA coincident with the linker insertion (see Fig. 4).

As illustrated in Figure 2B, the next step in the procedure was to take one of the collection of linker inserts and digest the plasmid containing the linker insert with BspM1. This releases the excision linker and, after treatment of the digested plasmid with Klenow polymerase to create blunt-ends, resulted in an eight nucleotide deletion of the *bla* gene at the original position of the linker insert. To replace the deleted nucleotides with nucleotides of random sequence, a second linker was ligated to the blunt-ended plasmid DNA (Fig. 2B). This linker (R-lac linker) contains four base pairs of random sequence at each end along with embedded BspM1 sites and the *lacO* sequence. After the ligation, the mix was transformed into *E. coli* using electroporation to generate the maximum number of transformants. Colonies produced by transformants with a plasmid that contains a linker were blue due to the *lacO* sequence binding and titrating the *lacI* gene product away from the chromosomal *lacO* site.[46] Each blue colony represents an independent random substitution mutant, and thus the number of blue colonies pooled indicates the number of different random substitution mutants included in the library. All of the colonies resulting from the transformation were

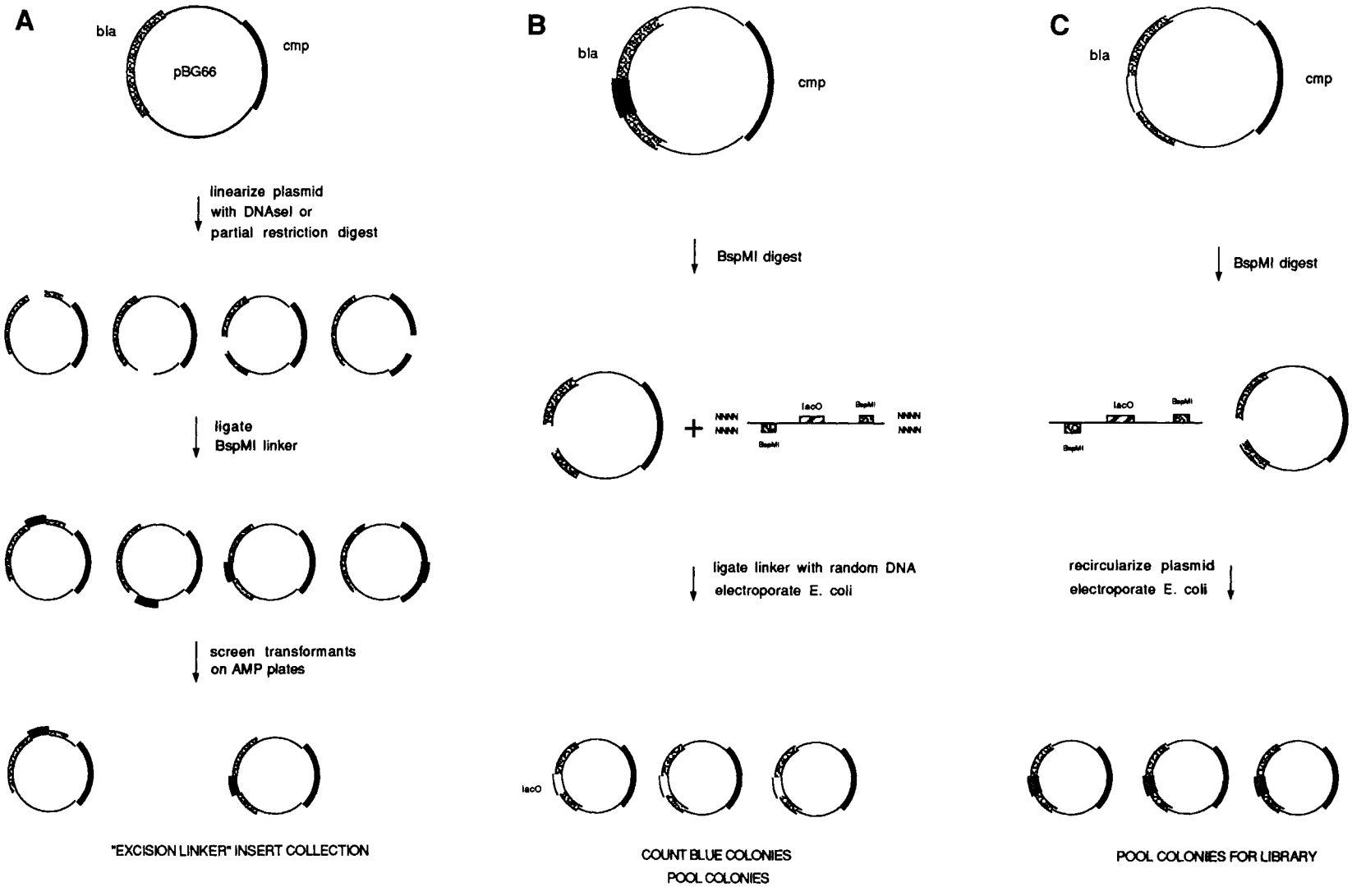

Fig. 2. Legend appears on page 35.

then pooled and the plasmid DNA extracted and purified.

Plasmid DNA from the pooled library was then digested with BspM1 to release the *lacO* section of the linker and leave the random DNA associated with the plasmid (Fig. 2C). The plasmid was then treated with Klenow DNA polymerase to generate blunt ends and DNA ligase to recircularize the plasmid. The result was the replacement of 8 bp of the *bla* gene with 8 bp of random sequence DNA. For those inserts of the excision linker that were associated with small deletions, the deletion after BspM1 digestion was larger than 8 base pairs (see Fig. 4). This was compensated for by using an R-lac linker that has an increased amount of random DNA at the ends so that after BspM1 digestion the amount of random DNA remaining with the plasmid was equal to the increased deletion size. Therefore, more codons were randomized in these cases. After recircularization of the plasmid, the ligation mix was used to electroporate *E. coli*. The transformants are pooled and the plasmid DNA was extracted and purified to create a plasmid library containing a set of random substitutions for a particular region of the gene. The random sequences were thus present in the form of a plasmid library containing the set of all or nearly all possible random substitutions for a particular region of the gene (Table I). Random libraries are designated by a label consisting of the letter L (library) followed by the position of amino acid sequences randomized; numbering according to Ambler.[15] For example, a library consisting random substitutions at positions 69 through 71 is designated L69-71.

Percentage of Functional Replacements

The percentage of random sequences in any library that result in a functional protein was taken to be an indication of the importance of the region randomized to the structure and/or function of the protein. To determine the percent functional replacements, each library was transformed into *E. coli* and plated on three sets of selective agar plates: (1) 1 mg/ml ampicillin, which is the maximal concentration of ampicillin on which *E. coli* containing a wild-type copy of the *bla* gene on the pBG66 plasmid will grow; therefore, this concentration selected for wild type levels of β-lactamase activity; (2) 10 μg/ml ampicillin, which is the minimal concentration of ampicillin that still selects unambiguously for β-lactamase function; and (3) 12.5 μg/ml chloramphenicol, which selects for the presence of the pBG66 plasmid without selecting β-lactamase activity.

By counting the number of transformants/ml able to grow on each type of plate, it was possible to determine the percentage of random sequences in each library that were consistent with wild type and minimal levels of β-lactamase function. The numerator for the calculations was the number of transformants/ml capable of growing on 1 mg/ml and 10 μg/ml ampicillin plates, respectively, and the denominator was the number of transformants/ml able to grow on the chloramphenicol plates.

The use of selections at both high and low ampicillin concentrations allowed functional random substitution mutants to be subdivided into two phenotypic classes, wild type and partial function. Catalysis of cleavage of the β-lactam bond of ampicillin by β-lactamase requires that the enzyme fold into a precise three-dimensional structure to make the proper contacts with the antibiotic. Thus altered enzymes that possess even minimal activity against ampicillin must retain the basic fold of β-lactamase even though the structure might be destabilized to some degree. The ability of substitution mutants selected at the high ampicillin concentration to function at a level similar to wild-type suggests these mutants contain relatively neutral amino acid substitutions.

The results of the 19 mutagenesis experiments are shown in Table II. Because the number of codons randomized was not the same in each experiment and some codons were only partially randomized, the absolute percentage of functional replacements cannot be compared from experiment to experiment. Therefore, the absolute percentages were normalized to represent the percent functional substitutions on a per codon basis for each mutagenesis experiment. As expected, the frequency of random replacements passing the selection was less for 1 mg/ml ampicillin than for 10 μg/ml ampicillin. The percent active substitutions for 10 μg/ml and 1 mg/ml ampicillin versus the position of the randomized amino acids in the β-lactamase sequence is plotted in Figure 5. It is apparent that a surprisingly large fraction of replacements are functional. For the minimal selection, the percent of functional replacements per codon varied from 5.7% for L260-262 to 85.1% for L182-186. The average percentage of replacements active at 10 μg/ml ampicillin for the 66 codons mutagenized in the 19 experiments was 44%. This translates to approximately (20 × .44) = 8.8 different amino acids, on average, that are able to

Fig. 2. **A.** Diagram illustrating the creation of a collection of "excision linker" inserts into the *bla* gene. The *bla* and *cat* genes present in the pBG66 starting plasmid are labeled. **B.** Diagram illustrating the generation of a library of pooled inserts of the R-lac linker. The starting plasmid is from the "excision linker" insert collection of A. The N's in the R-lac linker represent random nucleotides. The position of the *lac* operator sequence and the embedded BspM1 sites in the R-lac linker are shown schematically. **C.** Diagram illustrating the release of the interior portion of the R-lac linker from members of the R-lac linker library following BspM1 digestion. Note that the random nucleotides remain associated with the plasmid. Recircularization of the plasmid and electroporation of *E. coli* with the ligation mix results in a library of random substitutions.

TABLE II. Results of 19 Random Replacement Experiments*

Random library	Percent functional 10 μg/ml AMP	Percent functional 1 mg/ml AMP	Percent functional per codon 10 μg/ml AMP	Percent functional per codon 1 mg/ml AMP
L37-42	15.80	1.57	69.1	43.6
L69-71	10.25	0.50	37.9	10.5
L72-74	6.50	1.62	28.9	15.3
L102-104	41.38	0.76	73.0	17.5
L130-132	2.30	0.01	26.0	2.7
L134-136	12.61	4.52	47.1	32.4
L146-149	62.08	3.20	81.3	22.4
L157-160	0.25	0.01	15.4	5.0
L166-168	0.08	0.02	4.0	2.0
L168-170	0.50	0.08	9.5	4.2
L182-186	55.93	5.03	85.1	43.6
L196-200	36.96	1.65	79.1	38.3
L224-227	4.70	0.15	33.6	9.8
L233-235	0.22	0.02	6.2	2.1
L235-237	9.21	0.23	24.7	6.7
L238-241	44.61	2.46	78.9	26.8
L250-252	11.11	3.00	38.1	20.3
L259-262	0.18	0.05	5.7	3.2
L286-289	0.20	0.02	5.9	2.1

*Table shows the data used to calculate the percent functional replacements per codon. Details of the calculations are in Materials and Methods.

replace the wild-type amino acid at each residue position and still retain the β-lactamase fold. This number should be regarded as the number of different amino acids able to function at a residue position if the rest of the protein retains the wild-type sequence since combining substitutions, each of which individually reduces the activity of the protein somewhat, will likely result in a protein that is inactive. Nevertheless, the number suggests that the "fold" of the β-lactamase protein is very tolerant to substitutions.

For the 1 mg/ml ampicillin selection, the percent of functional replacements per codon varied from 2.1% for L233-235 and L286-288 to 43.6% for L37-42. The average percentage of active replacements for the 66 codons mutagenized was 20%, which translates to approximately (20 × .2) = 4 different amino acids per residue position, on average, that are able to function at a residue position. Since these mutants phenotypically resemble wild type, these results indicate that the number of possible neutral mutations in the *bla* gene must be extremely large.

Although TEM-1 β-lactamase is, in general, very tolerant of amino acid substitutions, there are regions of the protein that essentially cannot tolerate substitutions (see Fig. 5). For the 1 mg/ml ampicillin selection, 8 of the 19 random libraries including L130-132, L157-160, L166-168, L168-170, L233-235, L235-237, L260-262, and L286-288 contained a very low percentage of functional random sequences (< 5%) and thus represent important regions. Interestingly, 5 of these 8 regions (L130-132, L166-168, L168-170, L233-235, and L235-237) are located within the active site pocket on the three-dimensional structure of the homologous *S. aureus* and *B. licheniformis* β-lactamases.[22,24] The position of the random libraries on the three-dimensional structure of the homologous *S. aureus* β-lactamase are shown in Figure 4. Thus regions likely to be directly involved with the function of the protein were especially sensitive to amino acid substitutions.

The L157-160 and L260-262 regions represent a different circumstance. They are far from the active site but are largely buried in the interior of the protein in the *S. aureus* and *B. licheniformis* structures.[22,24] One possibility is that they are constrained by structural requirements for buried positions. It is known from studies on the hemoglobins and the λ-repressor protein that positions buried in the interior of proteins are very sensitive to amino acid substitutions and the L157-160 and L260-262 results are consistent with this.[1–3]

Finally, the L286-288 region is an exposed surface of the protein in the extreme C-terminal α-helix. Interestingly, it has been shown previously that this C-terminal α-helix is required for release of β-lactamase, after transit of the cytoplasmic membrane, into the periplasmic space.[49] The L286-288 results suggest that there may be very specific sequence requirements at the C-terminus for proper secretion.

The shapes of the graphs in Figure 5 showing the percent functional replacements at 1 mg/ml and 10 μg/ml ampicillin are similar. However, there is not a constant proportion that relates the 1 mg/ml ampi-

cillin results with the 10 μg/ml ampicillin results. The relative importance of some regions seems to vary with the stringency of the functional selection. For example, the L72-74 and L102-104 libraries had similar percentages of functional replacements at 1 mg/ml ampicillin but at 10 μg/ml ampicillin the L72-74 library had a much lower percentage of functional replacements (see Fig. 5). Thus although the regions randomized in these libraries were indistinguishable in terms of importance at the wild-type level of function, the minimal selection revealed that the L72-74 region makes a more important contribution to the basic requirements for a folded, functional protein. This example illustrates that another layer of information about the importance of any given region of a protein to that protein's structure and function can be extracted by varying the stringency of the functional selection.

The Sequence of Functional Replacements

To obtain more detailed information about the sequence requirements for β-lactamase function, we determined the sequence of 128 replacement mutants from 8 different random libraries. The individual sequences are listed in Figure 6. The primary sequences of 14 class A β-lactamases have been reported and aligned.[28] This allows a comparison of the tolerance of residue positions to substitution in the random replacement experiments with their conservation among the 14 class A enzymes. The sequencing results are summarized and compared to the sequence conservation of positions in the class A enzymes in Figure 7.

The 8 libraries for which sequencing data is available are from 5 contiguous regions of the enzyme (see Fig. 7). Each of these regions is discussed with reference to its position in the three-dimensional structure of the homologous β-lactamases and its conservation among the class A enzymes.

Residues 37-42

This region is covered by the single L37-42 library (see Fig. 6). Positions 37–40 are at the C-terminal end of α-helix 1 and positions 41 and 42 are in a turn in the *S. aureus* and *B. licheniformis* structures (see Fig. 4).[22,24] Glu37 is conserved in all class A β-lactamases, whereas Ala-42 is replaced by Gly in only 2 of the 14 class A sequences.[28] Among the replacement mutants, both Glu37 and Ala42 can be substituted and retain activity on 1 mg/ml ampicillin. Because only 1 base pair of the Glu37 codon was randomized only 4 amino acids were sampled at this position (Fig. 3). Only two base pairs of the Ala-42 codon were randomized so only 12 amino acids were sampled at this position. Therefore, a wider range of amino acids than that seen may function at these positions. Positions 38–41 vary freely among both the class A enzymes and the TEM-1 replacement mutants. Since there was no clear pattern of sequence conservation among the L37-42 replacements at 10 μg/ml ampicillin, a number of nonfunctional replacements were sequenced. As seen in Figure 6, lysine or arginine was always found at position 42 among the nonfunctional replacements. This suggests that the only sequence requirement for minimal function might be not to have a lysine or arginine at position 42. This is an example of a region that is especially tolerant of substitutions.

Residues 69-74

This region is covered by the L69-71 and L72-74 libraries (Fig. 6). Positions 69-74 are at the N-terminus of a long α-helix in the *S. aureus* and *B. licheniformis* structures (Fig. 4).[22,24] This region contains the conserved catalytic residues Ser70 and Lys73. Consistent with an important role for these residues, they are conserved among the functional replacement mutants with only serine substituting at position 70 and only lysine or arginine substituting at position 73. These results are in agreement with previous site-directed mutagenesis experiments that indicated substitutions of Ser70 or Lys73 abolished β-lactamase function.[29,30] However, a Cys70 substitution has been shown to possess partial β-lactamase activity.[50] Cys70 was not recovered among the sequenced 10 μg/ml ampicillin selected replacements; however, too few replacements were sequenced to allow any conclusions.

Thr71 is well conserved among the class A β-lactamases and is also conserved among the 1 mg/ml ampicillin selected replacement mutants. However, the mutants functional at 10 μg/ml ampicillin vary freely. This result is consistent with the Thr71 site-saturation mutagenesis results of Schultz and Richards.[51] They found that only cells containing mutants with Tyr, Trp, Asp, Lys, or Arg at position 71 had no resistance to ampicillin. None of these amino acids were found among the functional replacement mutants. Our results show Met69, Phe72 and Val74 are also tolerant of substitutions (Fig. 6). These results suggest Met69, Thr71, Phe72, and Val74 are not critical for binding or catalysis.

Residues 103-105

This region is covered by the single L103-105 library (Fig. 6). Positions 103–105 form a loop that is in the vicinity of the substrate binding pocket of the *S. aureus* and *B. licheniformis* structures (Fig. 4).[22,24] Despite being well conserved among the class A enzymes (serine is found in 1 of the 14 enzymes),[28] Val103 is tolerant of substitutions. The pattern of substitutions among the 1 mg/ml ampicillin selected mutants suggests a hydrophobic side chain is the predominant requirement at this position.

Glu104 lacks sequence conservation among both the class A enzymes and the functional replacement mutants (Fig. 7). It is noteworthy that a Lys104

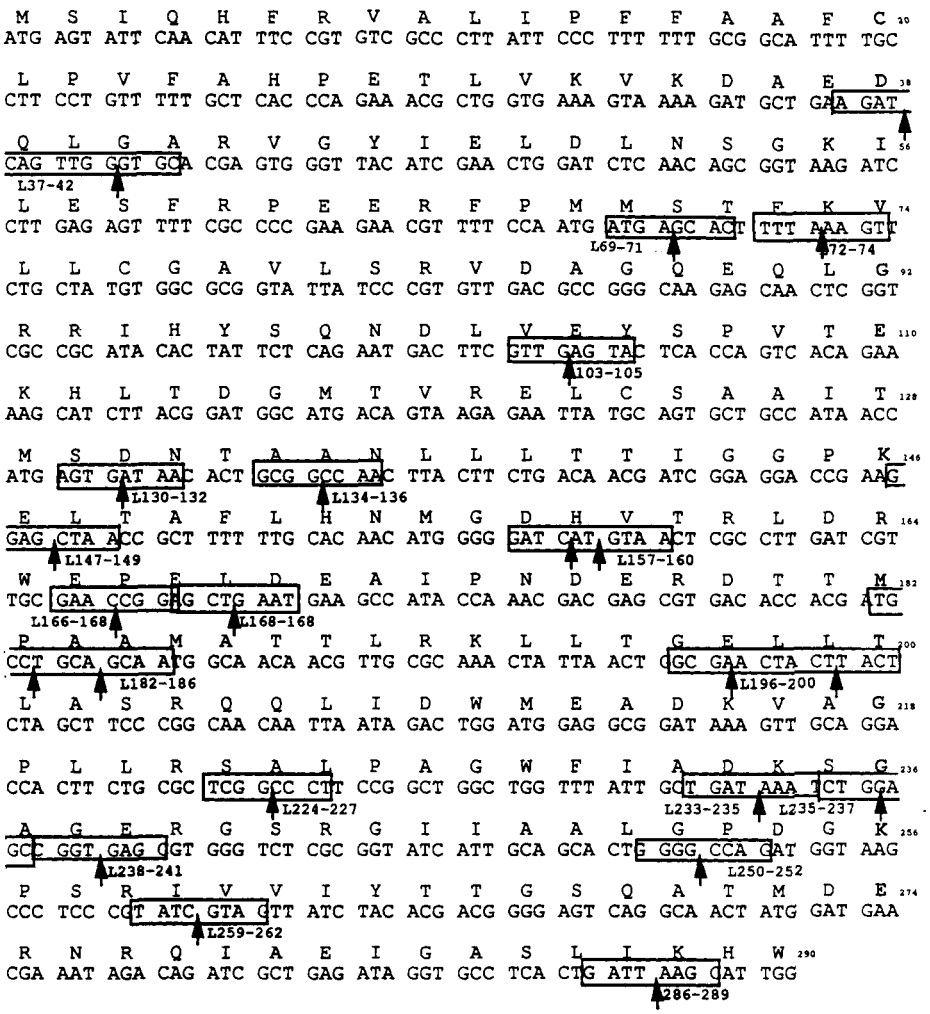

Fig. 3. Positions of the BspMI linker insertions in the *bla* gene. The arrows indicate the insertion points of the linkers. The boxed nucleotides indicate the nucleotides which are randomized in the experiment. Those inserts with two arrows represent BspMI linker inserts that were accompanied by small deletions. The nucleotides between the arrows are deleted. Amino acid numbering according to Ambler.[15]

substitution has been found in TEM enzymes from clinical isolates with increased hydrolytic activity against the third generation cephalosporin ceftazidime.[52,53] This illustrates that a residue position need not have high information content to mediate a change in enzyme structure and activity when mutated.

Tyr105 is moderately conserved among both the class A enzymes and the 1 mg/ml ampicillin selected replacements (Fig. 7). This residue is also conserved among the penicillin-recognizing enzyme superfamily.[19] Despite being conserved it is not apparent from the class A crystal structures what the role of position 105 is.[54] The variety of sequences that function among the 10 μg/ml ampicillin selected mutants suggest the residue is not essential for binding or catalysis (Fig. 6).

Residues 196-200

This region is covered by the single L196-200 library (Fig. 6). Positions 196–200 are in a turn between two α-helices in the *S. aureus* and *B. licheniformis* structures (Fig. 4).[22,24] Gly196, Glu197, Leu197, and Thr200 all vary freely among both the class A β-lactamases and the functional replacement mutants indicating these residues are not essential for the structure or activity of the enzyme. The results for Leu199 are striking in that the position is completely conserved among the 14 class A enzymes and yet varies among the functional replacement mutants (Fig. 7). It is possible that simultaneous changes at the other positions randomized in this experiment can compensate for changes at position 199. However, there is no identifiable pattern among the sequences (Fig. 6). This observation

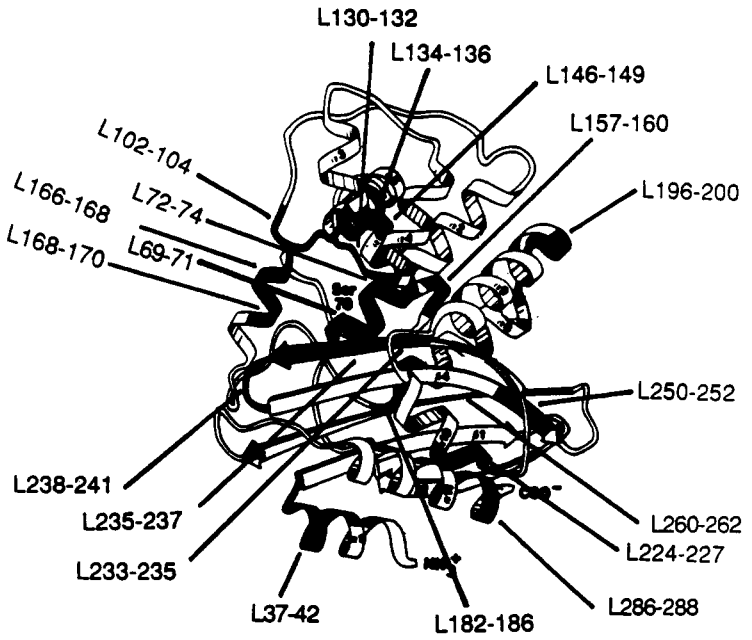

Fig. 4. Position of random libraries on a ribbon diagram of the homologous *S. aureus* β-lactamase.[22] Darkened regions correspond to the position of random libraries. Lines point to the position of individual libraries.

along with the very high percentage of functional replacements for the library (Fig. 5) suggests that if compensating mutations occur they must do so at a high frequency.

Residues 233-241

This region is covered by the L233-235, L235-237, and L238-241 libraries (Fig. 6). Positions 233–238 are part of a β-sheet that lines one side of the active site pocket in the *S. aureus* and *B. licheniformis* structures (Fig. 4). Positions 240 and 241 make up a turn between β-sheets. Note that position 239 does not exist in TEM-1 with the numbering system of Ambler.[15] Asp233, Lys234, Ser235, and Gly236 are strongly conserved among the class A enzymes (Fig. 7). Lys234, Ser235, and Gly236 make up the K(S/T)G box that is conserved in the penicillin-recognizing superfamily of enzymes.[19] Consistent with their evolutionary conservation these residues are also conserved among the 1 mg/ml ampicillin selected mutants. The conclusions from the sequences follow.

The aspartate residue at position 233 must be essential for wild-type levels of activity since only aspartate is found among the replacements functional at 1 mg/ml ampicillin. However, the aspartate seems not to be absolutely required for a folded functional protein since a variety of different amino acids are able to substitute for aspartate at 10 μg/ml ampicillin.

The lysine at position 234 is also required for wild-type activity as indicated by sequence conservation among the 1 mg/ml ampicillin selected replacements. A hydrogen bond donor group seems minimally required in the side chain of the residue at position of 234 to fulfill its role in the structure or function of the enzyme. This is evident from the fact that only amino acids with hydrogen bond donor groups among the 10 μg/ml functional replacements.

Position 235 is strongly constrained in that only serine was found among the 1 mg/ml ampicillin selected mutants. Replacement 1-2 from the 10 μg/ml class in Figure 6 clearly shows that even the conservative substitution of threonine for serine perturbs function somewhat. These results indicate position 235 has a very specific side chain requirement to fulfill its role in the function of the enzyme and also suggest that its role involves a hydrogen bonding interaction.

The glycine at position 236 is required for wild-type activity since only glycine is found among the replacements functional at 1 mg/ml ampicillin. However, the glycine is not essential for a functional enzyme since several different amino acids can substitute for glycine at 10 μg/ml ampicillin.

The conservation of hydrogen bonding groups among the functional replacements at positions 234 and 235 shows that the random replacement strategy is useful in both identifying important residues and in suggesting the possible function of the residues. Based on the X-ray structure of the *S. aureus*

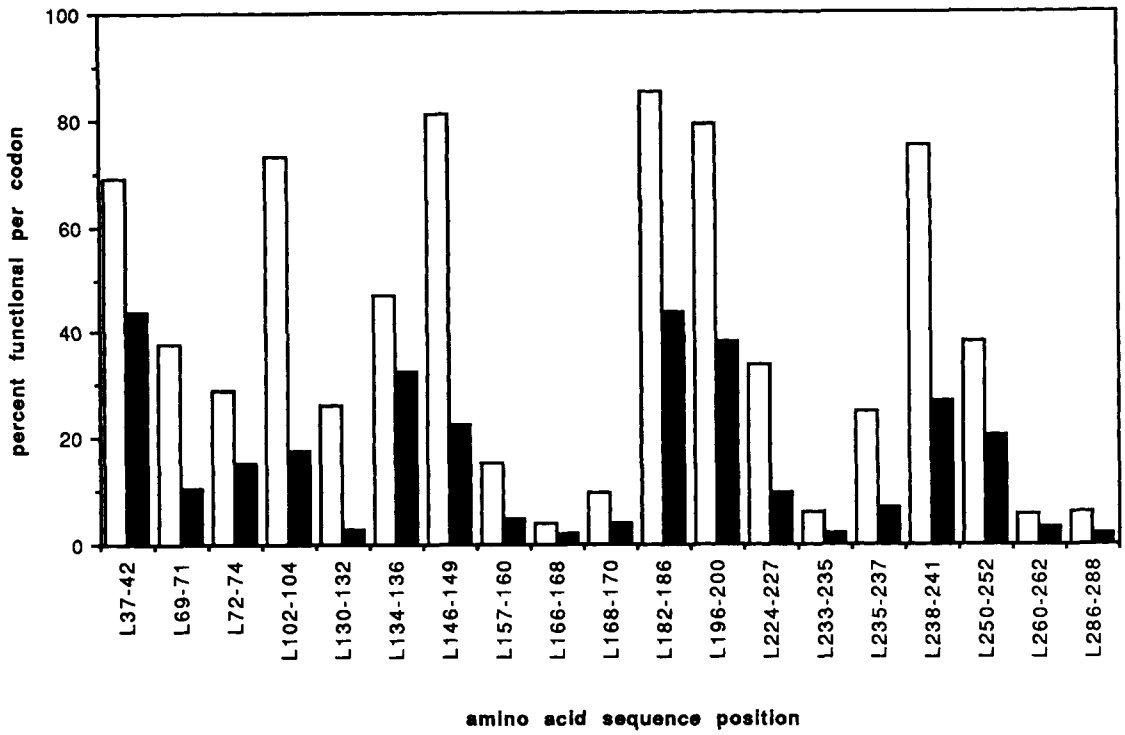

Fig. 5. Percent functionally acceptable random replacement sequences at 10 μg/ml and 1 mg/ml ampicillin. The percentages are normalized to percent functional per codon. Each set of columns represent a different random replacement experiment and is labelled accordingly. The open boxes represent percent functional replacements at 10 μg/ml ampicillin and the solid boxes represent percent functional replacements at 1 mg/ml ampicillin.

β-lactamase, lysine 234 has been postulated to participate in catalysis by acting as an electrostatic anchor for the C3 carboxylate of penicillins.[22] Site-specific mutagenesis results suggest the role of lysine 234 is more complex with a role in establishing an electrostatic environment for substrate binding as well as for stabilizing the transition state.[31] Our results indicate that a hydrogen bond donor group may be all that is required in the position 234 side chain for minimal function which suggests a hydrogen bonding interaction is important for function. Position 235 is a threonine in the homologous *B. licheniformis* enzyme and has been proposed to serve as a hydrogen bond donor to the C3 carboxylate of penicillins.[24] Consistent with this suggestion is the fact that only serine and threonine were found at position 235 among the functional replacements that were sequenced.

It has been suggested, based on crystal structures, that Gly-236 is conserved in class A enzymes because of its proximity to the catalytic Ser70.[54] A β-carbon containing residue at this position would be expected to clash with the serine residue. Consistent with this proposal only glycine is found at this position among the 1 mg/ml ampicillin selected replacements. However, other residues, including a bulky leucine chain in the L235-237 2-33 mutant, do not completely abolish function suggesting the enzyme can partially compensate for the steric clash (Fig. 6).

Positions 237-241 vary among both the class A β-lactamases and the functional replacements suggesting these residues are not critical for binding or catalysis. However, there is a preference for amino acids with small side chains at the Ala237 and Gly238 positions among the 1 mg/ml ampicillin selected replacement mutants. It is interesting to note that a Ser238 substitution has been found in TEM enzymes from clinical isolates with increased hydrolytic activity against the third-generation cephalosporin cefotaxime.[52,53] All of the replacement mutants with serine or cysteine at position 238 have increased activity against cefotaxime (Palzkill and Botstein, unpublished results).

Although positions 240 and 241 vary freely, substitutions at these sites can influence the activity of the enzyme. This is illustrated by mutants L238-241 1-34 and 3-30 (Fig. 6). A single amino acid substitution of arginine for aspartate at position 241 can reduce the activity of the enzyme toward ampicillin. Comparison of mutants 1-24 and 3-20 illustrates the same phenomenon for position 240. These results are interesting in that a Lys240 substitution identified initially in a clinical isolate has been shown to increase the activity of the enzyme towards third generation cephalosporins.[55]

L37-42

	37					42
WT	A	E	D	Q	L	G	A	R
1 mg/ml ampicillin								
C-26	A	A	A	L	L	K	A	R
C-27	A	A	S	A	L	A	A	R
C-29	A	E	-	L	G	N	Q	R
C-31	A	E	R	Y	T	Q	S	R
10 μg/ml ampicillin								
A-3	A	D	S	F	E	E	T	R
A-1	A	A	S	L	M	R	G	R
A-9	A	A	G	G	R	F	G	R
A-8	A	A	-	I	C	G	G	R
M-2	A	E	V	F	T	N	E	R
M-5	A	D	R	N	G	Q	L	R
A-4	A	D	V	Q	N	E	T	R
A-5	A	D	C	I	A	S	G	R
A-6	A	A	V	V	K	G	Q	R
A-2	A	A	S	F	K	G	E	R
M-1	A	A	F	C	H	P	G	R
M-4	A	V	L	S	G	L	T	R
Non-functional								
P-1	A	A	L	P	T	D	K	R
P-14	A	V	W	S	I	D	R	R
P-15	A	V	S	S	E	R	K	R
P-16	A	V	F	E	G	N	R	R
P-17	A	A	S	L	R	G	K	R
P-20	A	A	V	V	R	L	K	R
P-21	A	A	L	R	T	H	K	R

L69-72

	69	71		
WT	M	M	S	T	F
1 mg/ml ampicillin					
1-25	M	M	S	T	F
1-26	M	M	S	T	F
1-14	M	S	S	T	F
1-16	M	S	S	T	F
1-19	M	T	S	T	F
10 μg/ml ampicillin					
1-1	M	M	S	I	F
1-2	M	M	S	G	F
1-4	M	M	S	Q	F
1-6	M	T	S	P	F
1-7	M	T	S	G	F
1-3	M	I	S	T	F
1-9	M	I	S	P	F
1-10	M	I	S	H	F

L72-74

	72	74		
WT	T	F	K	V	L
1 mg/ml ampicillin					
C-34	T	F	K	V	L
C-26	T	F	K	V	L
C-2	T	F	K	L	L
A-33	T	V	K	V	L
10 μg/ml ampicillin					
C-35	T	C	K	V	L
C-43	T	C	K	I	L
C-6	T	S	K	I	L
C-27	T	Q	K	V	L
A-27	T	Q	K	I	L
A-1	T	L	K	V	L
A-24	T	L	K	I	L
A-28	T	Y	K	I	L
A-30	T	T	K	V	L
A-31	T	K	K	V	L
A-32	T	W	K	V	L
A-34	T	V	K	I	L
A-20	T	L	R	V	L
A-29	T	V	R	I	L

L103-105

	103	105		
WT	L	V	E	Y	S
1 mg/ml ampicillin					
2-33	L	V	K	Y	S
4-3	L	L	G	Y	S
6-36	L	I	L	Y	S
6-6	L	V	S	Y	S
2-35	L	L	A	Y	S
1-4	L	A	L	H	S
10 μg/ml ampicillin					
6-3	L	Y	Q	S	S
1-7	L	V	Y	N	S
1-15	L	V	L	G	S
1-2	L	V	Y	N	S
2-28	L	F	R	N	S
3-11	L	T	R	N	S
1-23	L	S	W	P	S
5-18	L	L	Y	H	S
6-31	L	V	H	H	S
3-6	L	L	R	C	S
1-5	L	H	R	H	S
1-9	L	L	F	Y	S
1-12	L	C	R	G	S

L196-200

	196				200	
WT	T	G	E	L	L	T	L
1 mg/ml ampicillin							
C-1	T	A	D	L	I	S	L
C-2	T	A	P	T	F	T	L
C-3	T	G	P	S	V	A	L
C-8	T	D	R	Q	R	A	L
C-16	T	A	S	S	V	T	L
10 μg/ml ampicillin							
A-1	T	A	R	P	Q	T	L
A-2	T	V	Y	H	R	E	L
A-3	T	A	P	G	T	A	L
A-4	T	A	L	G	R	D	L
A-5	T	A	R	S	S	G	L
A-6	T	G	F	R	R	K	L

L233-235

	233		235	
WT	A	D	K	S	G
1 mg/ml ampicillin					
2-1	A	D	K	S	G
2-2	A	D	K	S	G
2-7	A	D	K	S	G
10 μg/ml ampicillin					
1-2	A	D	K	T	G
1-4	A	S	N	T	G
A-4	A	G	N	T	G
A-7	A	V	H	T	G
A-8	A	R	R	T	G
A-9	A	L	N	S	G

L235-237

	235		237	
WT	K	S	G	A	G
1 mg/ml ampicillin					
1-2	K	S	G	A	G
1-3	K	S	G	A	G
1-16	K	S	G	A	G
1-21	K	S	G	A	G
1-4	K	S	G	G	G
1-1	K	S	G	G	G
1-18	K	S	G	G	G
10 μg/ml ampicillin					
C-6	K	S	G	N	G
C-20	K	S	G	S	G
C-7	K	S	G	T	G
C-10	K	S	G	T	G
C-19	K	S	G	T	G
2-33	K	L	L	Q	G
2-34	K	S	S	S	G
2-36	K	S	C	G	G
2-37	K	S	G	K	G
2-38	K	S	S	Q	G

L238-241

	238		241	
WT	A	G	E	R	G
1 mg/ml ampicillin					
3-27	A	G	Y	D	G
3-25	A	G	R	H	G
1-10	A	G	Q	P	G
4-13	A	A	C	A	G
5-21	A	-	C	P	G
1-34	A	S	K	D	G
1-24	A	C	D	S	G
10 μg/ml ampicillin					
3-30	A	S	K	R	G
3-4	A	S	S	P	G
2-21	A	S	E	N	G
4-5	A	S	S	A	G
4-15	A	C	Y	N	G
3-20	A	C	N	S	G
3-1	A	C	C	H	G
1-1	A	L	E	Y	G
1-2	A	L	R	H	G
1-3	A	V	Q	P	G
1-9	A	N	H	C	G

Fig. 6. Amino acid sequences of functional replacement mutants from 8 random libraries. Each library is labelled at top and the wild-type sequence of TEM-1 β-lactamase is listed. The randomized amino acids are in bold type.

CONCLUSIONS

The results of the replacement experiments indicate that TEM-1 β-lactamase is very tolerant of amino acid substitutions. A similar tolerance for amino acid substitutions has been observed for other proteins. The 360 amino acid residue *lac* repressor

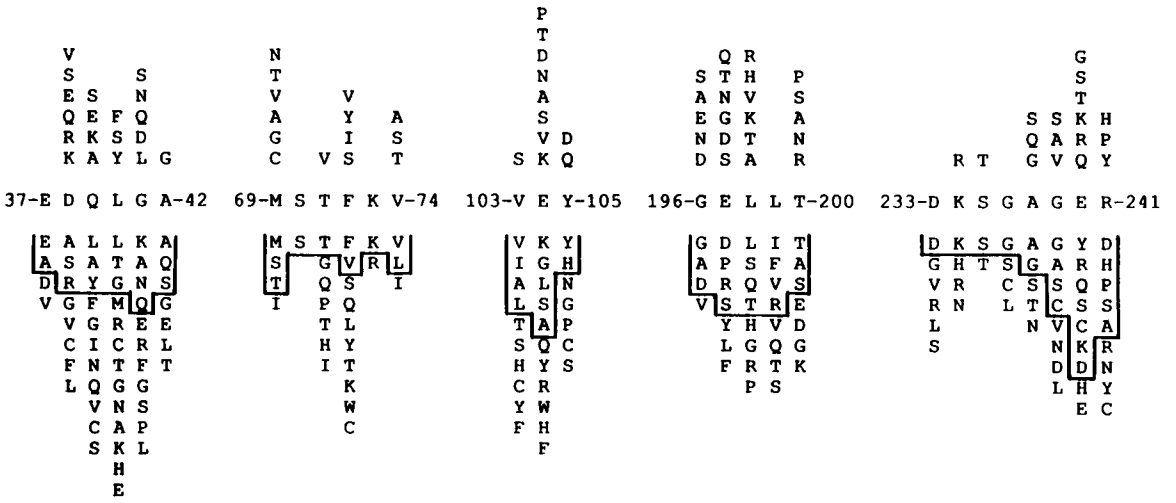

Fig. 7. Amino acid sequence variability of class A β-lactamases versus TEM-1 random replacement mutants. Listed and numbered appropriately are the TEM-1 β-lactamase residues that were randomized and for which sequence data is available. Above the TEM-1 residues are the different amino acids found at the analogous positions in the 14 class A β-lactamases that have been aligned.[28] If no amino acid is present at a position it reflects complete conservation of the position among class A sequences. Below the TEM-1 sequence is the different amino acid residues found for each position among the random replacement mutants whose sequences are listed in Figure 6. Sequences found among 1 mg/ml ampicillin selected replacements are boxed.

has been studied with over 1600 single amino acid substitutions.[6] The DNA binding domain of the repressor is located in the amino-terminal 59 residues and is sensitive to substitutions. However, the remaining 301 amino acid core of the repressor is extremely tolerant of substitutions, with only 30% of the amino acid replacements resulting in a nonfunctional phenotype.[6] Over 30 residues in the N-terminal domain of λ repressor have been examined by randomizing codons and selecting for function.[1,2] Nearly half of the residues accept nine or more different residues, and only a few positions are absolutely conserved. The results with TEM β-lactamase (an enzyme) and the *lac* and λ repressors (DNA binding proteins) suggests that high tolerance of amino acid substitutions may be a general property of proteins.

The tolerance of β-lactamase to substitutions has important implications for antibiotic resistance. The use of β-lactam antibiotics places selective pressure on β-lactamases to increase or change their hydrolytic capacity towards those antibiotics. The tolerance to substitutions allows many modified forms of the enzyme to be sampled for altered properties. Thus whatever antibiotic is used, there is a good chance resistance can develop by mutations in β-lactamase. This point is emphasized by a recent epidemic of plasmid mediated mutant derivatives of TEM and a related class A enzyme, SHV, that hydrolyze third generation cephalosporins.[56] It will be of interest to screen the random libraries described here against a variety of different β-lactam antibiotics.

Sequencing 128 functional replacements from 8 different random libraries revealed a wide variation in the tolerance of individual residue positions to amino acid substitutions. A comparison of the pattern of conserved residues among a set of 14 class A β-lactamases with the conservation of residues among the sequenced functional replacement mutants shows a rough correlation. In general, positions that are conserved by evolution are conserved among the functional replacements (Fig. 7). However, there are exceptions. Specifically, Glu37, Ala42, Val103, and Leu199 are conserved among the class A β-lactamases but not among the functional replacement mutants. This could be due to an insufficient number of class A sequences. As more sequences become available, these positions may become less conserved. Alternatively, the functional selection used here may not be as stringent as those required in nature. For example, if TEM-1 is in a pathogen it will need to function at temperatures higher than 37°C such as occur in patients with a fever. Assaying the functional mutants under different conditions might address this question.

The results presented here illustrate the utility of random replacement mutagenesis. The regions identified as most sensitive to substitution are located either in the active site or in buried positions that probably contribute to the structure of the protein. The method seems particularly suited to identifying important regions in proteins whose three-dimensional structure is not known. Random replacement alone does not distinguish between positions important for function versus positions important for structure, but once important regions are identified further experiments can be performed to distinguish the possibilities.

Because they contain all possible amino acid sub-

stitutions for a given region of a protein, the random libraries generated in the procedure are useful for a number of purposes: (1) DNA sequencing functional replacements gives information on the side chain characteristics important at a position(s), (2) variants within the library may contain substitutions that change the structural or functional properties of the protein such as stability to denaturants or substrate specificity, and (3) the libraries may contain conditional lethal mutations such as cold sensitive or temperature sensitive alleles useful in genetic studies.

ACKNOWLEDGMENTS

We thank members of the Botstein lab for helpful discussions. We also thank Dr. Georjana Barnes for critical reading of the manuscript. The coordinates for the *B. licheniformis* β-lactamase X-ray structure were kindly provided by Dr. James Knox and colleagues. We also thank Genentech, Inc., for the use of facilities and many colleagues at Genentech for helpful advice. A special thanks to Dr. Michael Pleiss for patient instructions on molecular modeling.

REFERENCES

1. Reidhaar-Olson, J.F., Sauer, R.T. Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. Science 241:53–57, 1988.
2. Reidhaar-Olson, J.F., Sauer, R.T. Functionally acceptable substitutions in two α-helical regions of λ repressor. Proteins 7:306–316, 1990.
3. Lim, W.A., Sauer, R.T. Alternative packing arrangements in the hydrophobic core of λ repressor. Nature (London) 339:31–36, 1989.
4. Perutz, M.F., Kendrew, J.C., Watson, H.C. Structure and function of haemoglobin. Some relations between polypeptide chain conformation and amino acid sequence. J. Mol. Biol. 13:669–678, 1965.
5. Bowie, J.U., Reidhaar-Olson, J.F., Lim, W.A., Sauer, R.T. Deciphering the message in protein sequences: tolerance to amino acid substitutions. Science 247:1306–1310, 1990.
6. Kleina, L.G., Miller, J.H. Genetic studies of the *lac* repressor XIII. Extensive amino acid replacements generated by the use of natural and synthetic nonsense suppressors. J. Mol. Biol. 212:295–318, 1990.
7. Shortle, D. Lin, B. Genetic analysis of Staphyloccocal nuclease: identification of three intragenic "global suppressors" of nuclease minus mutations. Genetics 110:539–555, 1985.
8. Dunn, I.S., Cowen, R., Jennings, P.A. Improved peptide function from random mutagenesis over short "windows." Protein Engineer. 3:283–291, 1988.
9. Bushman, F.P.C. Shang, C., Ptashne, M. A single glutamic acid residue plays a key role in the transcriptional activation function of λ repressor. Cell 58:1163–1171, 1989.
10. Oliphant, A.R., Struhl, K. Defining the consensus sequences of *E. coli* promoter elements by random selection. Nucl. Acids Res. 16:7673–7683, 1988.
11. Oliphant, A.R., Struhl, K. An efficient method for generating proteins with altered enzymatic properties: application to β-lactamase. Proc. Natl. Acad. Sci. USA 86:9094–9098, 1989.
12. Dube, D.K., Loeb, L.A. Mutants generated by the insertion of random nucleotides into the active site of the β-lactamase gene. Biochemistry 28:5703–5705, 1989.
13. Datta, N., Kontomichalou, P. Penicillinase synthesis controlled by infectious R factors in Enterobacteriaceae. Nature (London) 208:239–241, 1965.
14. Bush, K. Characterization of β-lactamases. Antimicrob. Agents Chemother. 33:259–263, 1989.
15. Ambler, R.P. The structure of β-lactamases. Phil. Trans. Ry. Soc. London, Ser. B 289:321–331, 1980.
16. Cartwright, S.J., Coulson, F.W. Active site of staphylococcal β-lactamase. Phil. Trans. Ry. Soc. London, Ser. B. 289: 370–372, 1980.
17. Fisher, J., Belasco, J.G., Khosla, S., Knowles, J.R. β-lactamase proceeds via an acyl-enzyme intermediate. Interaction of the *Escherichia coli* RTEM enzyme with cefoxitin. Biochemistry 19:2895–2901, 1980.
18. Sutton, B.J., Artymiuk, P.J., Cordero-Borboa, A.E., Little, C., Phillips, D.C., Waley, S.G. An X-ray crystallographic study of β-lactamase II from *Bacillus cereus* at .35 nm resolution. Biochem. J. 248:181–188, 1987.
19. Joris, B., Ghuysen, J.-M., Dive, G., Renard, A., Dideberg, O., Charlier, P., Frere, J.-M., Kelly, J.A., Boyington, J.C., Moews, P.C., Knox, J.R. The active-site-serine penicillin-recognizing enzymes as members of the *Streptomyces* R-61 DD-peptidase family. Biochem. J. 250:313–324, 1988.
20. Tipper, D.J., Strominger, J.L. Mechanism of action of penicillins. A proposal based on their structural similarity to acyl-D-alanyl-D-alanine. Proc. Natl. Acad. Sci. U.S.A. 54: 1133–1141, 1965.
21. Samraoni, B., Sutton, B.J., Todd, R.J., Artymiuk, P.J., Waley, S.G., Phillips, D.C. Tertiary structural similarity between a class A β-lactamase and a penicillin-sensitive D-alanyl carboxypeptidase-transpeptidase. Nature (London) 320:378–380, 1986.
22. Herzberg, O., Moult, J. Bacterial resistance to β-lactam antibiotics: Crystal structure of β-lactamase from *Staphylococcus aureus* PCl at 2.5 Å resolution. Science 236:694–701, 1987.
23. Dideberg, O., Charlier, P., Wery, J.P., Dehotty, P., Dusart, J., Erpicum, T., Frere, J.-M., Ghuysen, J.-M. The crystal structure of the β-lactamase of *Streptomyces albus* G at 0.3 nm resolution. Biochem. J. 245:911–913, 1987.
24. Moews, P.C., Knox, J.R., Dideberg, O., Charlier, P., Frere, J.-M. β-lactamase of *Bacillus licheniformis* 749/C at 2 Å resolution. Proteins 7:156–171, 1990.
25. Oefner, C., D'Arcy, A., Daly, J.J., Gubernator, K., Heinze, I., Hubschwieler, C., Winkler, F.K. Refined crystal structure of β-lactamase from *Citrobacter freundii* indicates a mechanism for β-lactam hydrolysis. Nature (London) 343: 284–288, 1990.
26. Kelly, J.A., Dideberg, O., Charlier, P., Wery, J.P., Libert, M., Moews, P.C., Knox, J.R., Duez, C., Fraipont, C., Joris, B., Dusart, J., Frere, J.-M., Ghuysen, J.-M. On the origin of bacterial resistance to penicillin: Comparison of a β-lactamase and a penicillin target. Science 231:1429–1431, 1986.
27. Medeiros, A.A. β-lactamases. British Medical Bulletin 40: 18–27, 1984.
28. Juteau, J.-M., Levesque, R.C. Sequence analysis and evolutionary perspectives of ROB-1 β-lactamase. Antimicrob. Microb. Agents Chemother. 34:1354–1359, 1990.
29. Dalbadie-McFarland, G., Cohen, L.W., Riggs, A.D., Morin, C., Itakura, K., Richards, J.H. Oligonucleotide-directed mutagenesis as a general and powerful method for studies of protein function. Proc. Natl. Acad. Sci. U.S.A. 79:6409–6413, 1982.
30. Madgwick, P.J., Waley, S.G. β-lactamase I from *Bacillus cereus*. Structure and site-directed mutagenesis. Biochem. J. 248:657–662, 1987.
31. Ellerby, L.M., Escobar, W.A., Fink, A.L., Mitchinson, C., and Wells, J.A. The role of lysine-234 in β-lactamase catalysis probed by site-directed mutagenesis. Biochemistry 29:5797–5806, 1990.
32. Adachi, H., Ohta, T., Matsuzawa, H. Site-directed mutants, at position 166, of RTEM-1 β-lactamase that form a stable acyl-enzyme intermediate with penicillin. J. Biol. Chem. 266:3186–3191, 1991.
33. Sutcliffe, J.G. Nucleotide sequence of the ampicillin resistance gene of *Escherichia coli* plasmid pBr322. Proc. Natl. Acad. Sci. U.S.A. 75:3737–3741, 1978.
34. Roy, C., Segura, C., Tirado, M., Reig, R., Hermida, M., Teruel, D., Foz, A. Frequency of plasmid-determined β-lactamases in 680 consecutively isolated strains of Enterobacteriaceae. Eur. J. Clin. Microb. 4:146–147, 1985.
35. Cooksey, R., Swenson, J., Clark, N., Gay, E., Thornsberry, C. Patterns and mechanisms of β-lactam resistance among isolates of *Escherichia coli* from hospitals in the United States. Antimicrob. Agents Chemother. 34:739–745, 1990.

36. Cunningham, B.C., Jhurani, P., Ng, P., Wells, J.A. Receptor and antibody epitopes in human growth hormone identified by homolog scanning mutagenesis. Science 243: 1330–1336, 1989.
37. Miller, J.H. "Experiments in Molecular genetics." Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1972.
38. Maniatis, T., Fritsch, E.F., Sambrook, J. "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1982.
39. Garger, S.J., Griffith, O.M. Grill, L.K. Rapid purification of plasmid DNA by a single centrifugation in a two-step cesium chloride-ethidium bromide gradient. Biochem. Biophys. Res. Comm. 117:835–842, 1983.
40. Birnboim, H.C., Doly, J. A rapid alkaline extraction procedure for screening recombinant DNA. Nucl. Acids Res. 7:1513–1523, 1979.
41. Maniatis, T., Fritsch, E.F., Sambrook, J. "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1989.
42. Sanger, F., Nicklen, S., Coulson, A.R. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74:5463–5467, 1977.
43. Heffron, F., So, M., McCarthy, B.J. In vitro mutagenesis of a circular DNA molecule by using synthetic restriction sites. Proc. Natl. Acad. Sci. USA 75:6012–6016, 1978.
44. Kwoh, T.J., Zipser, D., Wigler, M. Mutational analysis of the cloned chicken thymidine kinase gene. J. Mol. Appl. Genet. 2:191–200, 1983.
45. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491, 1988.
46. Heyneker, H.L., Shine, J., Goodman, H.M., Boyer, H.W., Rosenberg, J., Dickerson, R.E., Norang, S.A., Itakura, K., Lin, S.-Y., Riggs, A.D. Synthetic *lac* operator DNA is functional in vivo. Nature 263:748–752, 1976.
47. Szybalski, W. Universal restriction endonucleases: Designing novel cleavage specificities by combining adaptor oligodeoxynucleotide and enzyme moieties. Gene 40:169–173, 1985.
48. Mormeneo, S., Knott, R., Perlman, D. Precise nucleotide sequence modifications with bidirectionally cleaving class-IIS excision linkers. Gene 61:21–30, 1987.
49. Koshland, D., Botstein, D. Secretion of β-lactamase requires the carboxy end of the protein. Cell 28:749–760, 1980.
50. Sigal, I.S., Harwood, B.G., Arentzen, R. Thiol-β-lactamase: replacement of the active-site serine of RTEM β-lactamase by a cysteine residue. Proc. Natl. Acad. Sci. U.S.A. 79: 7157–7160, 1982.
51. Schultz, S.C., Richards, J.H. Site-saturation studies of β-lactamase: Production and characterization of mutant β-lactamases with all possible amino acid substitutions at residue 71. Proc. Natl. Acad. Sci. U.S.A. 83:1588–1592.
52. Collatz, E., VanNhieu, T., Billot-Klein, D., Williamson, R., Gutmann, L. Substitution of serine for arginine in position 162 of TEM-type β-lactamases extends the substrate profile of mutant enzymes TEM-7 and TEM-101 to ceftazidime and aztreonam. Gene 78:349–354, 1989.
53. Sougakoff, W., Petit, A., Goussard, S., Sirot, D., Bure, A., Courvalin, P. Characterization of the plasmid genes blaT4 and blaT5 which encode the broad-spectrum β-lactamases TEM-4 and TEM-5 in Enterobacteriaceae. Gene 78:339–348, 1989.
54. Herzberg, O. Refined crystal structure of β-lactamase from *Staphylococcus aureus* PC1 at 2.0 Å resolution. J. Mol. Biol. 217:701–719, 1991.
55. Sowek, J.A., Singer, S.b., Ohringer, S., Malley, M.F., Dougherty, T.J., Gougontas, J.Z., Bush, K. Substitution of lysine at position 104 or 240 of TEM-1_{pTZ18R} β-lactamase enhances the effect of serine-164 substitution on hydrolysis or affinity for cephalosporins and the monobactam aztreonam. Biochemistry 30:3179–3188, 1991.
56. Bush, K. Excitement in the β-lactamase arena. J. Antimicrob. Chemother. 24:831–840, 1989.