Nucleic Acids Research, 2005, Vol. 33, No. 9 ***e80***
doi:10.1093/nar/gni077

Triplet nucleotide removal at random positions in a target gene: the tolerance of TEM-1 β-lactamase to an amino acid deletion

D. Dafydd Jones*

School of Biosciences, Biomedical Sciences Building, Cardiff University, Cardiff CF10 3US, UK

Received February 15, 2005; Revised April 1, 2005; Accepted April 20, 2005

ABSTRACT

The deletion of amino acids is one of the evolutionary mechanisms by which nature adapts the function of proteins. A simple method has been developed that mimics this event *in vitro* by introducing a deletion of exactly three nucleotides at random positions in a target gene. The method involved the engineering of the mini-Mu transposon to introduce a recognition sequence for the restriction enzyme MlyI. The new transposon, MuDel, was capable of efficient insertion into a target DNA sequence. To determine the efficacy of the method, the *bla* gene that encodes the TEM-1 β-lactamase was used as the target and a small library containing 22 different sequence variants was created. Of these 22 variants, 8 were identified that conferred resistance to ampicillin on *Escherichia coli*. Each of the TEM-1 variants possessed a distinct ampicillin minimum inhibitory concentration, ranging from 500 to >10 000 μg/ml. Sequence analysis revealed that active TEM-1 variants contained deletions not just in loops but also helices, and included regions known to be involved in catalysis, antibiotic resistance and inhibitor binding. This new technology is transferable to most genes, permitting an extensive analysis of deletion mutations on protein function.

INTRODUCTION

Nature has evolved an impressive myriad of proteins to perform the functions for the fitness of an organism. Apart from point mutations, nucleotide insertion and deletion is utilized by nature during the evolutionary process (1). Selection pressure means that deletion or insertion of three nucleotides or multiples thereof are favoured as they maintain the reading frame of the gene (2). During the process of divergent evolution, many insertion–deletion (indel) mutation events result in the addition or removal of just a single amino acid, mainly in coils and loops that link secondary structures (2–4). Many of these length variations have profound effects on the properties of the protein especially folding, ligand or substrate binding, protein–protein interactions and temperature-dependent activity and stability. For example, the sequence variation of the immunoglobulin variable domains is enhanced by amino acid deletions and insertions (5), and various indel events are observed between the structurally homologous subtilisin serine proteases including in regions known to be important for substrate and calcium binding (6).

The deletion of a single codon is one of the most common forms of indel mutation observed in nature and illustrates its importance to the process of evolution (2). Mimicking such an event *in vitro* would help understand the influence of deletion mutations on protein structure and function, and enhance our ability to improve the properties of proteins for a particular application. Currently, the most common method of introducing deletion mutations is by rational design, which relies on structural information to determine the residues to be deleted and requires separate oligonucleotides for each mutation. The introduction of random mutations throughout a target gene is a powerful method for altering the properties of a protein [see (7,8) for reviews]. Most of the current technologies have focused on the introduction of point mutations leading to an amino acid substitution [(9–11) and references therein]. Some methods have been used to introduce amino acids insertions, e.g. pentapeptide scanning mutagenesis (12) and random insertion and deletion (RID) (13). The RID method has the potential to introduce single amino acid deletions but has currently been applied only to introduce amino acid substitutions or insertions. Furthermore, the procedure is complicated and prone to the introduction of unwanted secondary mutations (13).

A new method is described here that introduces triplet nucleotide deletions at random positions throughout a target gene. To demonstrate the method, triplet nucleotide deletions have been introduced at random into the gene encoding the TEM-1 β-lactamase. TEM-1 is a clinically important protein

*To whom correspondence should be addressed. Tel: +44 29 20874290; Fax: +44 29 20874116; Email: jonesdd@cardiff.ac.uk

as it is one of the main causes of bacterial resistance to β-lactam antibiotics. Many natural variants of TEM-1 exist that have evolved to confer resistance to new, extended-spectrum (ES) β-lactam antibiotics (see www.lahey.org/Studies/temtable.asp and references therein). Although no naturally occurring deletion variants of TEM-1 have been discovered, amino acid deletions have been observed in variants of homologous β-lactamases, such as SHV-1 (14) and *Staphylococcus aureus* PC1 (15). TEM-1 has also been the focus of many protein engineering studies (16), including the random substitution of every amino acid (17), directed evolution (18,19) and pentapeptide scanning mutagenesis (20). Therefore, to complement the existing knowledge concerning amino acid substitutions and insertions, the effect of an amino acid deletion on TEM-1 structure and function was assessed.

The method exploits the properties of the mini-Mu transposon, a DNA element that can be accurately and efficiently inserted into a target DNA sequence *in vitro* using the MuA transposase (21). The reaction has a very low target site preference allowing transposon insertion to occur essentially at any point in a given gene. The mini-Mu transposon was modified close to both its termini to incorporate the recognition sequences for the type IIS restriction enzyme MlyI. Cleavage with the restriction enzyme and religation resulted in the removal of 3 bp at random positions within the *bla* gene. Each sequenced variant contained a single amino acid deletion, and the position of the mutation was shown to have a profound effect on TEM-1 activity.

MATERIALS AND METHODS

Materials

The original *cam*^{r}-containing Mu transposon, MuA transposase and pEntranceposon (Cam^r) plasmid were supplied by Finnzymes Oy. All restriction endonucleases were purchased from NE Biolabs (UK) Ltd. The Extensor Hi-Fidelity PCR enzyme mix and T4 DNA ligase were purchased from Abgene UK. *Taq* DNA polymerase was supplied by Promega Ltd. Oligonucleotides were synthesised by Operon Biotechnologies Inc. Chloramphenicol (Cam) and ampicillin (Amp) were supplied by Melford Laboratories Ltd. The isolation of plasmid DNA from cell cultures was performed using the Wizard® Plus SV kit from Promega Ltd. DNA was isolated from agarose gel or PCR reactions using the Qiaquick Gel extraction or PCR purifications kits, respectively, supplied by Qiagen Ltd.

Construction of MuDel transposon

The MuDel transposon was constructed by PCR with the Extensor Hi-Fidelity PCR enzyme mix using the oligonucleotide DDJdi005 (5′-GCTTAGATCTGActCGGCGCACGAAAAACGCGAAAG-3′; lower case letters signify nucleotides undergoing mutagenesis) as both the forward and reverse primer with 0.1 ng of the original *cam*^{r}-containing Mu transposon acting as template. The 1322 bp product was purified and digested with BglII. The digested transposon was isolated and purified after agarose gel electrophoresis and recloned into BglII-digested pEntranceposon (Cam^r) using T4 DNA ligase. The MuDel transposon was released from the plasmid by digestion with BglII and purified after agarose gel electrophoresis.

Transposition with MuDel

The efficiency of the transposition reaction with MuDel was tested using pUC18 plasmid and the libraries constructed using pNOM (see Supplementary Material for details concerning the construction of pNOM). Target plasmid DNA (100 ng of either pUC18 or pNOM) and either mini-Mu (20 ng) or MuDel (20 ng or 100 ng) were added to reaction buffer [25 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 110 mM NaCl, 0.05% Triton and 10% (v/v) glycerol] containing 0.22 μg of MuA transposase in a final volume of 20 μl. The reactions were left at 30°C for 3 h followed by heat inactivation at 75°C for 10 min. Either 1 or 2 μl of the reaction mixture was used to transform *Escherichia coli* DH5α cells by electroporation (cell competency 3×10^7 colonies formed/μg pUC18 plasmid) and the cells were plated on LB agar containing 20 μg/ml Cam to select for cells containing the *cam*^{r} gene and hence the mini-Mu or MuDel transposon.

Creation of deletion libraries and selection for active *bla* gene products

Transposition was performed as described above using the MuDel transposon and pNOM as the target DNA. After transformation of *E.coli* DH5α with the transposition reaction, 48 colonies that grew on 20 μg/ml Cam were selected. These colonies were replated on separate LB agar plates containing either 100 μg/ml Amp or 20 μg/ml Cam. Of the 48 colonies, 22 grew only on the Cam plate and were deemed to have a disrupted *bla* gene due to transposon insertion in this region and therefore chosen as the members of the BLA^{DEL} library. To confirm the presence of the MuDel transposon, PCR was performed with *Taq* DNA polymerase on each of the 22 colonies using primers DDJdi010 (5′-TCCGCTCATGAGACAATAACCCTG-3′) and DDJdi011 (5′-CTACGGGGTCTGACGCTCAGTG-3′) that flank the *bla* gene. The PCR products were isolated and restriction analysis was performed with MlyI to confirm the diversity of transposon insertion positions. The MuDel-inserted pNOM plasmids were isolated from each of the 22 colonies and equal amounts of each plasmid were pooled and subjected to restriction digestion with MlyI followed by agarose gel electrophoresis. The band corresponding to the linear pNOM minus MuDel was isolated and purified. Intramolecular ligation was performed using T4 DNA ligase and ~10 ng of linear pNOM. The reaction was left at 25°C for 10 min followed by 10 h at 16°C. *E.coli* DH5α cells were transformed by electroporation with 1 μl of the reaction mixture and the cells were plated on LB agar plates containing 15 μg/ml Amp. The plates were left overnight at 37°C and 94 BLA^{DEL} library and 2 pNOM-containing colonies were selected and transferred to 96 deep-well culture plates containing 200 μl of Luria–Bertani (LB) and 15 μg/ml Amp. The cells were grown for 16 h at 37°C with vigorous shaking. Sterile glycerol was added to 10% (v/v) for storage at −80°C.

The TEM-1 β-lactamase activity of each colony was measured *in vivo* by determining the minimum inhibitory concentration (MIC) of Amp that prevents *E.coli* growth. Each colony was transferred to LB agar in Nunc Omnitray plates containing 50, 100, 500, 2500, 5000, 7500 or 10 000 μg/ml Amp using

a 96 pin replicator fork (Boekel Scientific) and left at 37°C for 16 h. Several clones that exhibited MIC at each Amp concentration were subjected to PCR with *Taq* DNA polymerase using primers DDJdi010 and DDJdi011. The PCR products were sequenced to confirm the position of the amino acid deletion.

RESULTS

Design and construction of the MuDel transposon and pNOM

The codon deletion mutagenesis procedure is outlined in Figure 1 and uses a modified version of the mini-Mu transposon. This transposon was chosen as it has an accurate and efficient *in vitro* transposition reaction with a very low target site preference (21).

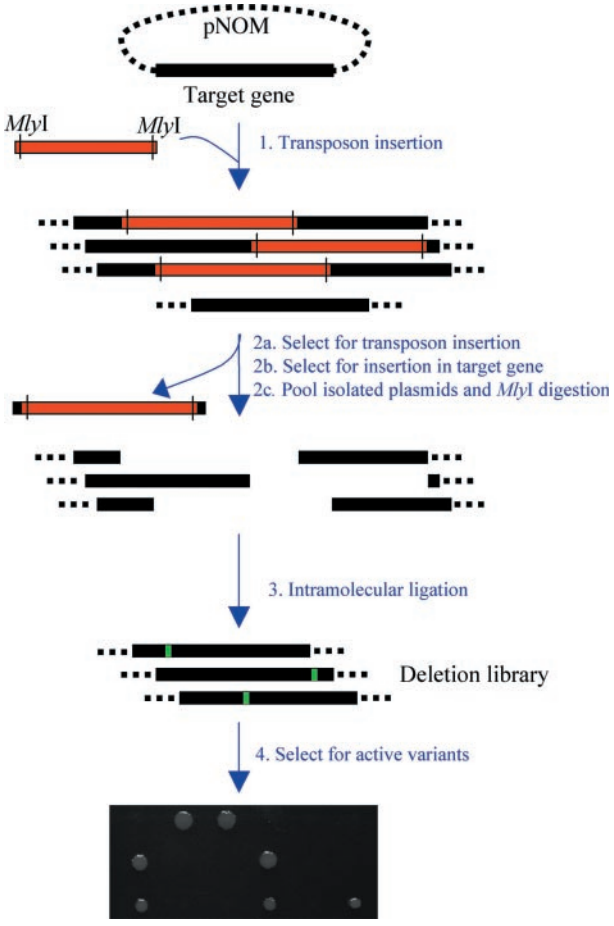

Figure 1. Outline of the triplet nucleotide deletion mutagenesis method. The procedure consists of four main steps. Step 1: The MuDel transposon is inserted into the target plasmid. Step 2: Cells containing a plasmid-integrated MuDel contain the *cam*^{r} gene and so can grow in the presence of Cam. The cells are then duplicate plated and grown in the presence of Amp. Those cells growing only on Cam contain a disrupted *bla* gene due to MuDel insertion and are therefore selected. The plasmids are isolated and pooled, and the transposon is removed by MlyI digestion. Step 3: Intramolecular ligation results in the reformation of the *bla* gene, minus 3 bp. Step 4: The resulting library is grown on LB agar containing various concentrations of Amp to determine the *in vivo* activity. Red blocks represent the transposon, black blocks the *bla* gene, green blocks the deletion point and the thick dashed lines the rest of the plasmid backbone.

The ability to delete nucleotide triplets depends on the transposon insertion mechanism, the position of the two introduced restriction sites and the cutting characteristics of the restriction endonuclease, as outlined in Figure 2. The mini-Mu transposon was engineered to act as a vehicle for the insertion of specific restriction sites into the target gene (Figure 2a). The restriction endonuclease chosen was MlyI (GAGTC(N_5)↓), a type IIS enzyme that cuts 5 bp outside its recognition sequence to generate a blunt end. The MlyI recognition site was placed 1 bp away from the site of transposon insertion, creating MuDel (Figure 2a). The two required point mutations both lie outside the R1 region that is involved in MuA binding (22), thereby minimizing disruption to the protein–DNA interactions that can potentially affect the efficiency of the transposition reaction. The positioning of the MlyI recognition sequences within MuDel should also take into account the transposition mechanism. Transposition of MuDel occurs via a 5 bp staggered cut in the target DNA (23,24) that results in the duplication of these 5 bp (Figure 2b). Digestion of the DNA with MlyI removes the transposon together with four additional base pairs at both termini from the target gene. Intramolecular ligation of the two blunt ends results in the in-frame deletion of three nucleotide base pairs from the target gene (Figure 2b).

The MuDel transposon, containing the *cam*^{r} gene that confers resistance to Cam, was constructed as described in Materials and Methods. To test if the introduced mutations disrupted transposition efficiency, pUC18 was used as the target DNA substrate. Transposition of the wild-type mini-Mu (20 ng) transposon into pUC18 as the target DNA resulted in the growth of ~180 *E.coli* DH5α colonies on 20 μg/ml Cam plates after transformation with 1/10th (2 μl) of the transposition reaction mixture. This translates into a minimum of 1800 transposon insertions into pUC18 per reaction. Replacing

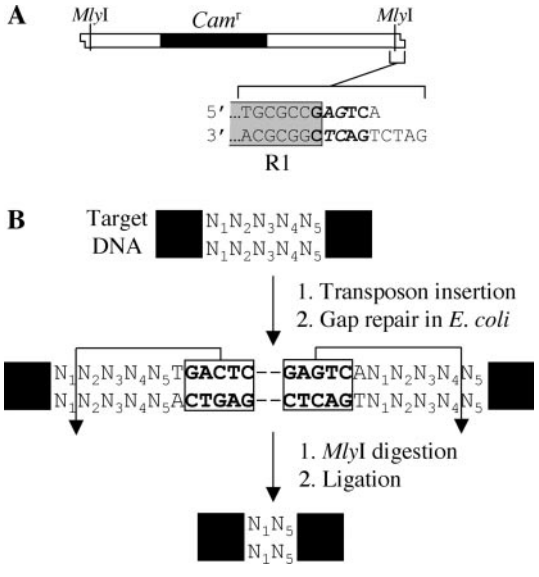

Figure 2. (**A**) Sequence of the engineered MuDel transposon. The letters in italics are those mutagenized compared with the wild-type mini-Mu. (**B**) Mechanism for the introduction of 3 bp deletion. Transposon insertion results in a 5 bp duplication after gap repair in *E.coli* (the 4 bp overhang from the transposon is removed). The cleavage site of MlyI is 5 bp away from the recognition sequence, resulting in the removal of 4 bp of the target gene at both ends. Ligation of the two termini rejoins the gene but minus 3 bp.

the mini-Mu transposon with either 20 or 100 ng of MuDel resulted in the growth of ~200 and 860 colonies, respectively. This translates into a minimum of 2000 and 8600 transposon insertions per reaction, respectively. Therefore, MuDel still acts as an efficient substrate for the transposition reaction to enable the potential sampling of the whole of the plasmid. It must be noted that the number of *E.coli* DH5α cells recovered, containing transposon-inserted pUC18, will benefit greatly from using electrocompetent cells with higher transformation efficiencies than that used here.

For the amino acid deletion method to work, the target DNA or plasmid containing the target gene cannot contain any MlyI restriction sites. The recognition sequence of MlyI is only 5 bp in length, so many plasmids have at least one if not more MlyI restriction sites. For example, pUC18 has four MlyI restriction sites, including one in the *bla* gene encoding TEM-1 β-lactamase, one in the origin of replication (*ori*) and another in the multiple cloning site (MCS). Therefore, an appropriate plasmid was constructed that contained no MlyI sites and a useful MCS, making it a suitable subcloning vector for general use outside of the present context. The new vector, pNOM, was constructed as described in the Supplementary Material. The majority of the plasmid was donated by pUC18, including the *ori* regions and the *bla* gene. The MlyI site present in the *bla* gene was removed by the introduction of a silent mutation so as not to disrupt the primary structure of the TEM-1 β-lactamase. Removal of the MlyI site from the *ori* region was achieved by creating a library in which two of the nucleotides that form the MlyI recognition sequences were randomized, as it was unknown how rational mutations may affect plasmid replication. The new MCS site contains many common restriction sites and the pelB leader sequence for the attachment of a signal sequence.

Plasmids that conferred Amp (100 μg/ml) resistance upon *E.coli* DH5α were selected and purified. Upon restriction endonuclease digestion analysis, it was observed that none of the plasmids was cut with MlyI (data not shown), so proving that pNOM no longer contained recognition sequences for MlyI.

Creation and selection of deletion variants of TEM-1

The general outline of the method is shown in Figure 1. The *bla* gene that encodes TEM-1 β-lactamase was chosen as the target as it is a clinically important enzyme responsible for resistance to some β-lactam antibiotics. It also provides an easy selection method, as active TEM-1 variants will confer resistance to ampicillin on *E.coli*, thereby permitting cell growth. The new vector, pNOM, was used as the source of the *bla* gene and therefore acts as the target DNA for MuDel insertion (step 1, Figure 1). The position of MuDel insertion into pNOM should be distributed evenly throughout the plasmid; therefore, a strategy is required that will select cells containing MuDel inserted into the *bla* gene region (step 2, Figure 1). The successful insertion of MuDel into pNOM will confer resistance to Cam on *E.coli*. After transformation of *E.coli* DH5α and plating on LB agar containing 20 μg/ml Cam, 48 colonies were selected at random and replated on LB agar containing either 20 μg/ml Cam or 100 μg/ml Amp. Those colonies that have MuDel inserted within the *bla* gene region will disrupt TEM-1 expression and thus affect the cells ability to grow in the presence of Amp. Of the 48 colonies, 22 failed to grow on Amp. Screening of the 22 colonies by PCR using primers that flank the *bla* gene confirmed that the MuDel transposon was present within this region (data not shown). The *bla* gene constitutes 861 bp of the 2115 bp pNOM (~41%); therefore, 22 of 48 colonies were close to the expected value if the MuDel insertion is random. It must be noted that insertions in the *ori* region may be detrimental to plasmid replication, which will affect the cells' resistance to Cam, and that MuDel insertions into the *bla* gene may not always result in the loss of TEM-1 activity. The 22 clones containing MuDel within the *bla* gene region were selected to constitute the BLA^{DEL} library.

To determine the potential diversity of the BLA^{DEL} library, the PCR products generated above were subject to restriction analysis with MlyI. Digestion with MlyI will result in the removal of the MuDel transposon and 8 bp of the *bla* gene (1310 bp), generating two fragments of varying length, depending on the MuDel insertion point (Figure 3a). The restriction analysis revealed that the insertion of MuDel occurred randomly and only one transposon was inserted in this region (Figure 3b). Length analysis of the two smaller fragments confirmed that their cumulative size was approximately equal to that of the PCR product minus MuDel.

Equal concentrations of each of the 22 MuDel-inserted pNOM plasmids were pooled and digested with MlyI (steps 2b and 2c, Figure 1). The linear pNOM DNA minus MuDel was isolated and intramolecular ligation reunited the two ends of the same pNOM to reform the *bla* gene minus 3 bp (step 3, Figure 1). The ligation mixture was used to transform *E.coli* DH5α cells that were then grown on LB agar containing 15 μg/ml Amp. Selection with such a low Amp concentration ensured that cells with low TEM-1 activity would survive. Even at such low concentrations of Amp, inactive TEM-1

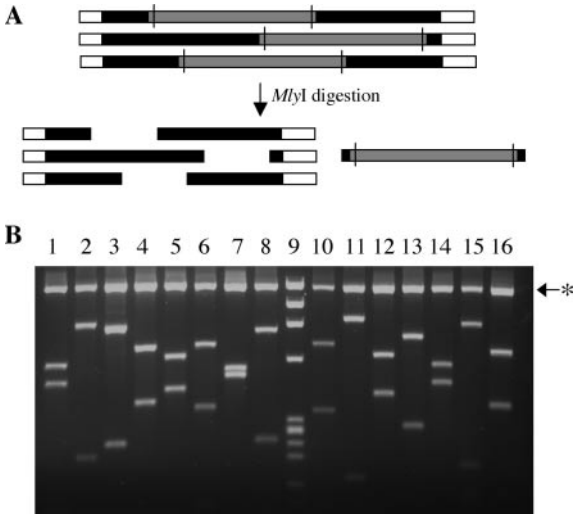

Figure 3. Restriction analysis of the BLA^{DEL} library with MlyI to determine the randomness of transposon insertion. (**A**) Illustration of the restriction analysis procedure. The transposon region is coloured grey, the *bla* region is black and the regions that flank the *bla* gene are white. The vertical lines represent the MlyI recognition sites. (**B**) Restriction analysis of 15 of the 22 BLA^{DEL} library members. Lanes 1–8 and 10–16 represent different members of the BLA^{DEL} library and lane 9 is the ϕ174 DNA-HaeIII molecular weight ladder (from top to bottom 1353, 1078, 872, 603, 310, 281/271, 234, 194 and 118 bp, respectively). The band labelled with an asterisk corresponds to the transposon.

deletion variants will no longer be able to support *E.coli* growth and will thus be removed from the library at this stage. More than 1000 colonies were observed per transformation. A control ligation was performed in the absence of T4 DNA ligase and only 10 colonies were observed indicating that false positives are restricted to <1% of cells on the experimental plate.

Ninety four BLA^{DEL} colonies together with 2 pNOM wild-type colonies were picked at random and grown in 96 well culture plates containing 15 μg/ml Amp. The *in vivo* activity of each variant was estimated by determining the MIC of Amp required to prevent cell growth (step 4, Figure 1). Each colony was plated on LB agar containing 50, 100, 500, 2500, 5000, 7500 or 10 000 μg/ml Amp and left for 16 h at 37°C and colonies capable of growing at each Amp concentration were noted. The MIC of each original colony for Amp is shown in Supplementary Figure S1. All the cells grew at both 50 and 100 μg/ml Amp. Fourteen had an Amp MIC of 500 μg/ml. No variants had an Amp MIC of 2500 μg/ml and 36 had an Amp MIC of 5000 μg/ml. No variants had an Amp MIC of 7500 μg/ml and the growth of 14 variants was inhibited at 10 000 μg/ml Amp. The remaining 32 colonies were still viable at 10 000 μg/ml Amp, including the two wild-type pNOM controls. Such a spread of MIC values indicates that various 3 bp deletion mutations have been incorporated into the *bla* gene and have had a profound effect on the *in vivo* activity of TEM-1.

Sequence analysis of selected TEM-1 variants

The point of insertion of MuDel with respect to a single codon will determine the nature of the deletion. The three possibilities are shown in Supplementary Table S1. One-third of all insertions will create a true deletion of a codon. In the other two-thirds, the 3 bp removed will overlap two codons and may result in a secondary point mutation. The nature of the secondary mutation will vary depending on the surrounding DNA sequence. Owing to the degeneracy of the genetic code, some of the point mutations will be silent while others will result in amino acid substitutions. This in fact mirrors the natural process by which amino acid deletions are introduced into a gene (2).

Several *bla* genes were isolated from clones exhibiting specific Amp MICs and sequenced to confirm the position of the amino acid deletion and if any secondary mutations have occurred. Table 1 shows all the different sequences isolated from active TEM-1 variants of the BLA^{DEL} library. Of the 22 different codon deletion variations present in the BLA^{DEL} library, only 8 were identified from the 30 sequenced clones that could be tolerated by TEM-1. Two of the eight (R83Δ and R275Δ; Δ denotes the residue deleted) were true codon deletions. Three of the sequences (T114Δ, G196Δ and A217Δ) did not generate true codon deletions at the genetic level, but owing to the degeneracy of the genetic code, no amino acid substitutions resulted. Three of the sequences (P62Δ-E63Q, T114Δ-D115N and E177Δ-R178D) contained a substitution mutation as well as a codon deletion. The MuDel insertion point with respect to a single codon is evenly distributed (2:3:3) as expected (Table 1). No other mutations were observed for any of the sequenced variants and no wild-type TEM-1 was detected.

The sequences were spread across the whole length of the primary structure of TEM-1 with varying effects on the *in vivo* activity. Two mutations, T114Δ and T114Δ-D115N, were only separated by a transposon insertion position of 2 bp (Table 1). It is unknown whether this region is more favourable to transposon insertion.

There was a good correlation between the location of the mutation and the Amp MIC (Table 1). The P62Δ-E63Q, R83Δ and E177Δ-R178D TEM-1 sequence variants all imposed an Amp MIC of 500 μg/ml on *E.coli* DH5α. Colonies containing the T114Δ and T114Δ-D115N mutations exhibited an Amp MIC at 5000 μg/ml. Those containing the R275Δ mutation had a relatively high MIC for Amp (5000 μg/ml) even though the deletion takes place within a helix. Growth of colonies containing the A217Δ TEM-1 mutant were inhibited at 10 000 μg/ml, whereas those containing the G196Δ mutation were still capable of growth at 10 000 μg/ml Amp.

DISCUSSION

Modified transposons as a method for generating triplet nucleotide deletions

The introduction of amino acid insertions and deletions is known to be an important contributor to protein evolution (4,25), with removal of a single codon being the most common

Table 1. The relationship between transposon insertion site, deletion mutation produced and Amp MIC

Wild-type sequence[a]	Deletion sequence[b]	Mutation[c]	Amp MIC (μg/ml)	Frequency[d]	Secondary structure
CGC**CCCGA**AGAA	CGCC———AAGAA	**P62Δ-E63Q**	500	3	Loop S2-SB1
TTATCCCGTATT	TTATCC———ATT	**R83Δ**	500	2	H2
CATC**TTACG**GAT	CATCT———GGAT	**T114Δ**	5000	3	Loop H3-SC4
CATCTT**ACGGA**T	CATCTTA———AT	**T114Δ-D115N**	5000	4	Loop H3-SC4
GACG**AGCGTG**AC	GACGA———TGAC	**E177Δ-R178D**	500	1	Ω loop
TTAA**CTGGC**GAA	TTAAC———CGAA	**G196Δ**	>10 000	7	Loop H8-H9
GT**TGCAG**GACCA	GTTG———GACCA	**A217Δ**	10 000	5	Loop H9-H10
GATGA**ACGAA**AT	GATGAA———AAT	**R275Δ**	5000	5	H11
–	–	**None[e]**	>10 000	2	–

No variant exhibited a MIC of 50, 100, 2500 or 7500 μg/ml.

[a]The five base pairs duplicated owing to transposon insertion are shown in underline.

[b]The new codons generated after deletion are underlined.

[c]The residues are numbered using the recommended numbering systems (30). A Δ after an amino acid residue number signifies that the residue has been deleted.

[d]The frequency refers to the number of sequenced *bla* genes with that mutation at that particular Amp MIC value.

[e]Refers to the two controls known to contain wild-type TEM-1 resident on the pNOM plasmid.

of such mutations (2,4). This has been successfully mimicked *in vitro* with a new technique using the mechanism of transposon insertion and restriction digestion with a type IIS endonuclease to generate 3 bp deletions at random positions in a target DNA sequence. The potential of the method has been demonstrated by using the MuDel transposon in conjunction with the newly constructed vector, pNOM, to successfully create a small library of TEM-1 β-lactamase variants containing a single amino acid deletion. From this small test library, at least eight different deletion mutations were identified that are tolerated by TEM-1, with the *in vivo* activity of each variant towards Amp dependent on the mutation present (Table 1).

One issue with this method is the potential for 'hot spots' within the target DNA for transposon insertion. The 5 bp target (duplication) site for Mu insertion has a reported consensus sequence of NPyG/CPuN (21). Of the eight identified deletion mutations, only four contained the 5 bp target site that strictly adhered to this already ambiguous consensus sequence (Table 1). Therefore, present data suggest that the transposon insertion is essentially random allowing deletion mutations to be distributed across the whole of the *bla* gene.

Another issue is the insertion of the transposon outside of the target gene. Of the 48 selected colonies containing MuDel within pNOM, a secondary screen based on the disruption of TEM-1 activity revealed that 22 contained the transposon in the *bla* gene. Using the *bla* gene separated from pNOM can alleviate this problem. It has been demonstrated previously that linear DNA acts as an efficient target for transposon insertion (21). Recloning the *bla* or any other gene back into pNOM after transposition will result in theoretically 100% of Cam tolerant cells containing the transposon in the target gene and remove the requirement for the secondary selection strategy based on the disruption of the target gene within the plasmid. This is particularly important when applying the method to a system in which there is no simple selection or screen available for the gene of interest.

The library appeared to be biased towards colonies with Amp MIC of >10 000 μg/ml, with all the colonies sequenced in this class having the G196Δ mutation in TEM-1. This suggests that 30 out of the 94 colonies contain this TEM-1 variant (Supplementary Figure S1). The MlyI restriction analysis of each member of the BLA^{DEL} library gave only one band pattern (5′ end 636 bp and 3′ end 431 bp) that could correspond to the G196Δ mutation (lane 12 in Figure 1). The subsequent sequencing of this particular PCR product revealed that it would produce the G196Δ mutation (data not shown). The most likely reason is a bias in the colony selection (step 4, Figure 1). The first 10 colonies selected from LB agar plates containing 15 μg/ml Amp (A3–A12 in Supplementary Figure S1) were among the largest, which may be owing to a highly active TEM-1 variant. This is further substantiated by the fact that three of the sequenced colonies in this row, A3, A4 and A7, contained the G196Δ mutation.

A similar reason can be used to explain the low frequency of colonies with the lowest Amp MIC (500 μg/ml), as they constitute only 14 of the 94 colonies screened and yet three different mutations are present in this class. The original colonies from this class were small, even when grown in a medium containing 15 μg/ml Amp. Many of the colonies may not have been visible by eye after just 16 h at 37°C and therefore may have been overlooked in the original selection of the 94 colonies.

Effect of deletion mutations on TEM-1 structure and function

The deletion of a single amino acid has had a major effect on the *in vivo* activity of TEM-1. Of the 22 different codon deletion variations present in the BLA^{DEL} library, only 8 were identified from the 30 sequenced clones that could be tolerated by TEM-1. Furthermore, the nature of the deletion mutation had a profound effect on the *in vivo* activity of TEM-1, as Amp MIC values varied depending on the mutation. The position of the mutations with respect to the tertiary structure of TEM-1 is illustrated in Figure 4 and their position within the secondary structure is shown in Table 1. The most likely place an amino

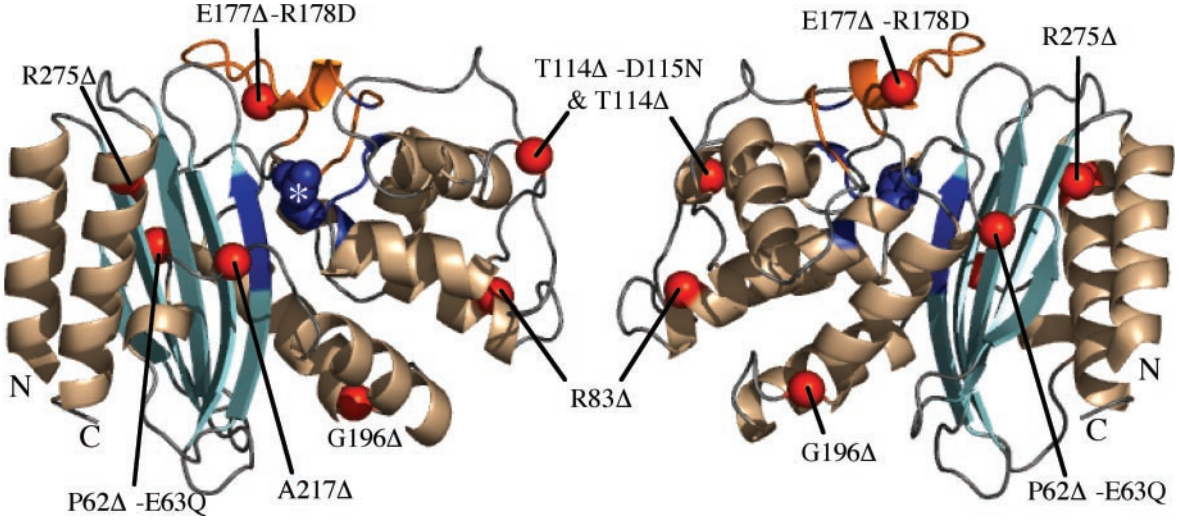

Figure 4. Position of the mutations on the tertiary structure of TEM-1 [PDB code 1BTL (26)]. The mutation sites are shown as red balls. Regions important for catalysis are coloured blue and S70 is shown as a space-fill and labelled with an asterisk. The Ω loop is coloured orange. The images differ by a rotation of ~180° around the *y* axis. The residues are numbered using the recommended numbering systems (30).

acid deletion is thought to be tolerated is in loops and this is generally the case in nature (4). This is substantiated by the fact that six of the eight deletion mutations were within loops (Table 1). Two of these six deletion mutations (E177Δ-R178D and P62Δ-E63Q) are in loops known to be important for TEM-1 structure and function (16,17,20,26,27), and another (A217Δ) removes a buried residue not thought to be tolerant to substitution (17). One of these loop mutations, G196Δ, appears to have little effect on the *in vivo* activity towards Amp as the MIC value is similar to that observed for the wild type (Table 1).

Amino acid deletions within helices and strands would normally be considered detrimental as they can potentially change the phase (or register) of these structures. However, two of the deletion mutations, R83Δ and R275Δ, occur within helical regions indicating that such mutations need not be restricted to loops. Moreover, the amino acid deletions occur in helices known to be important to catalysis and substrate binding (26). Whereas the R83Δ has a major effect on TEM-1 activity (Amp MIC value of 500 μg/ml), the protein can tolerate the deletion of R275 while retaining significant activity (Amp MIC value of 5000 μg/ml).

Two of the mutations, T114Δ and T114Δ-D115N, lie very close to each other owing to the MuDel transposon insertion point being separated by 2 bp (Table 1). The adjacent substitution mutation appears to have no effect on activity, as TEM-1 variants containing the deletion mutation alone or in combination with D115N both exhibit the same Amp MIC value of 5000 μg/ml.

The creation of a small library of deletion variants has illustrated that such mutations can have a profound effect on the activity of TEM-1 towards Amp. Although it has been shown here that amino acid deletion has a largely detrimental effect on TEM-1 activity, under different selection regimes (e.g. in the presence of an ES β-lactam or β-lactamase inhibitor) different characteristics may be observed that introduce beneficial properties. For example, mutation to R275 is known to play a role in the resistance to β-lactamase inhibitors such as clavulanic acid (28,29). Furthermore, the deletion of an amino acid may reveal point mutations not tolerated in the wild-type TEM-1 that affect the properties of the enzyme.

Summary

The transposon-based triplet nucleotide deletion method is a simple and rapid method for creating libraries of protein variants with single amino acids removed at random positions. Amino acid deletion mutagenesis will complement existing mutagenesis strategies for probing protein structure and function, and will aid in the optimization of proteins for biotechnological applications. Although triplet nucleotide removal has been demonstrated here, the method has other applications as the insertion of the transposon and subsequent digestion with MlyI introduces random, single breaks within a given target gene. The intramolecular ligation procedure can be replaced by the insertion of new DNA-encoding peptide sequences such as new codons (e.g. as in alanine scanning mutagenesis or for the introduction of unnatural amino acids) and random codons for directed evolution, epitopes or whole protein domains. These applications are currently being explored.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

The author would like to thank Cardiff University and the WDA Knowledge Exploitation Fund for funding, Dr Corinne Squire and Profs Rob John and Michael Ehrmann for critical reading of this manuscript, and Dr Colin Berry and Prof. John Kay for sharing their facilities. The method outlined in this manuscript is the subject of a patent application (0501189.5). The Open Access publication charges for this article were waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Chothia,C., Gough,J., Vogel,C. and Teichmann,S.A. (2003) Evolution of the protein repertoire. *Science*, **300**, 1701–1703.
2. Taylor,M.S., Ponting,C.P. and Copley,R.R. (2004) Occurrence and consequences of coding sequence insertions and deletions in mammalian genomes. *Genome Res.*, **14**, 555–566.
3. Lesk,A.M. (2001) *Introduction to Protein Architecture*. Oxford University Press, Oxford.
4. Pascarella,S. and Argos,P. (1992) Analysis of insertions/deletions in protein structures. *J. Mol. Biol.*, **224**, 461–471.
5. de Wildt,R.M., van Venrooij,W.J., Winter,G., Hoet,R.M. and Tomlinson,I.M. (1999) Somatic insertions and deletions shape the human antibody repertoire. *J. Mol. Biol.*, **294**, 701–710.
6. Siezen,R.J. and Leunissen,J.A. (1997) Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci.*, **6**, 501–523.
7. Tao,H. and Cornish,V.W. (2002) Milestones in directed enzyme evolution. *Curr. Opin. Chem. Biol.*, **6**, 858–864.
8. Arnold,F.H. (2001) Combinatorial and computational challenges for biocatalyst design. *Nature*, **409**, 253–257.
9. Lutz,S. and Patrick,W.M. (2004) Novel methods for directed evolution of enzymes: quality, not quantity. *Curr. Opin. Biotechnol.*, **15**, 291–297.
10. Neylon,C. (2004) Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res.*, **32**, 1448–1459.
11. Dalby,P.A. (2003) Optimising enzyme function by directed evolution. *Curr. Opin. Struct. Biol.*, **13**, 500–505.
12. Hallet,B., Sherratt,D.J. and Hayes,F. (1997) Pentapeptide scanning mutagenesis: random insertion of a variable five amino acid cassette in a target protein. *Nucleic Acids Res.*, **25**, 1866–1867.
13. Murakami,H., Hohsaka,T. and Sisido,M. (2002) Random insertion and deletion of arbitrary number of bases for codon-based random mutation of DNAs. *Nat. Biotechnol.*, **20**, 76–81.
14. Prinarakis,E.E., Miriagou,V., Tzelepi,E., Gazouli,M. and Tzouvelekis,L.S. (1997) Emergence of an inhibitor-resistant β-lactamase (SHV-10) derived from an SHV-5 variant. *Antimicrob. Agents Chemother.*, **41**, 838–840.
15. Zawadzke,L.E., Smith,T.J. and Herzberg,O. (1995) An engineered *Staphylococcus aureus* PC1 β-lactamase that hydrolyses third-generation cephalosporins. *Protein Eng.*, **8**, 1275–1285.
16. Matagne,A., Lamotte-Brasseur,J. and Frere,J.M. (1998) Catalytic properties of class A β-lactamases: efficiency and diversity. *Biochem. J.*, **330**, 581–598.
17. Huang,W., Petrosino,J., Hirsch,M., Shenkin,P.S. and Palzkill,T. (1996) Amino acid sequence determinants of β-lactamase structure and activity. *J. Mol. Biol.*, **258**, 688–703.
18. Camps,M., Naukkarinen,J., Johnson,B.P. and Loeb,L.A. (2003) Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I. *Proc. Natl Acad. Sci. USA*, **100**, 9727–9732.
19. Stemmer,W.P. (1994) Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature*, **370**, 389–391.

20. Hayes,F., Hallet,B. and Cao,Y. (1997) Insertion mutagenesis as a tool in the modification of protein function. Extended substrate specificity conferred by pentapeptide insertions in the omega-loop of TEM-1 β-lactamase. *J. Biol. Chem.*, **272**, 28833–28836.
21. Haapa,S., Taira,S., Heikkinen,E. and Savilahti,H. (1999) An efficient and accurate integration of mini-Mu transposons *in vitro*: a general methodology for functional genetic analysis and molecular biology applications. *Nucleic Acids Res.*, **27**, 2777–2784.
22. Craigie,R. and Mizuuchi,K. (1987) Transposition of Mu DNA: joining of Mu to target DNA can be uncoupled from cleavage at the ends of Mu. *Cell*, **51**, 493–501.
23. Craig,N.L. (1995) Unity in transposition reactions. *Science*, **270**, 253–254.
24. Mizuuchi,K. (1992) Transpositional recombination: mechanistic insights from studies of Mu and other elements. *Annu. Rev. Biochem.*, **61**, 1011–1051.
25. Soding,J. and Lupas,A.N. (2003) More than the sum of their parts: on the evolution of proteins from peptides. *Bioessays*, **25**, 837–846.
26. Jelsch,C., Mourey,L., Masson,J.M. and Samama,J.P. (1993) Crystal structure of *Escherichia coli* TEM1 β-lactamase at 1.8 Å resolution. *Proteins*, **16**, 364–383.
27. Wang,X., Minasov,G. and Shoichet,B.K. (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J. Mol. Biol.*, **320**, 85–95.
28. Canica,M.M., Barthelemy,M., Gilly,L., Labia,R., Krishnamoorthy,R. and Paul,G. (1997) Properties of IRT-14 (TEM-45), a newly characterized mutant of TEM-type β-lactamases. *Antimicrob. Agents Chemother.*, **41**, 374–378.
29. Henquell,C., Chanal,C., Sirot,D., Labia,R. and Sirot,J. (1995) Molecular characterization of nine different types of mutants among 107 inhibitor-resistant TEM β-lactamases from clinical isolates of *Escherichia coli*. *Antimicrob. Agents Chemother.*, **39**, 427–430.
30. Ambler,R.P., Coulson,A.F., Frere,J.M., Ghuysen,J.M., Joris,B., Forsman,M., Levesque,R.C., Tiraby,G. and Waley,S.G. (1991) A standard numbering scheme for the class A β-lactamases. *Biochem. J.*, **276**, 269–270.