Chapter IV: Construction and Characterization of Site-Directed Recombination Libraries

Introduction

There are several different methods for creating protein chimeras. By far the simplest technique is to take advantage of existing restriction sites to swap portions of DNA sequence. This is often augmented by PCR overlap-extension to generate a small populations of protein chimeras that are used for structure/function experiments (Back and Chappell 1996; Kushiro et al. 1999). However, constructing chimeras individually is time consuming and is not practical for creating a library of chimeras.

To create large numbers of protein chimeras there are variety of techniques that allow proteins to be recombined randomly. These methods fall into two general categories: homology-dependent and homology-independent. The homology-dependent methods include methodologies similar to DNA shuffling that create random gene fragments and reassemble them through PCR. Because these methods are annealing-based they rely on high sequence identity between the parental sequences for successful assembly. This limits them to recombining genes that share more than about 70% sequence identity. The homology-independent methods are capable of recombining distantly- or even nonrelated proteins. However, many of these techniques do not maintain the reading frame and allow inserts and deletions to occur at the recombination sites. Because the reading frame is not maintained, 2/3 of variants are out of frame and therefore not encoding useful proteins.

We have extended site-directed chimera construction to combinatorial assembly of gene fragments using ligation (Hiraga and Arnold 2003). This technique allows

nonhomologous genes to be recombined, but maintains the reading frame. Similarly to constructing individual chimeras using restriction sites, combinatorial ligation utilizes blocks of sequence with specific basepair overhangs. These overhangs are the same in all the parental sequences, but different at each recombination site, and allow specific ligation of the sequence blocks in the correct order. Because the overhangs are the same for each parental sequence, the blocks from different parents are freely exchangeable during construction (Figure IV-1)

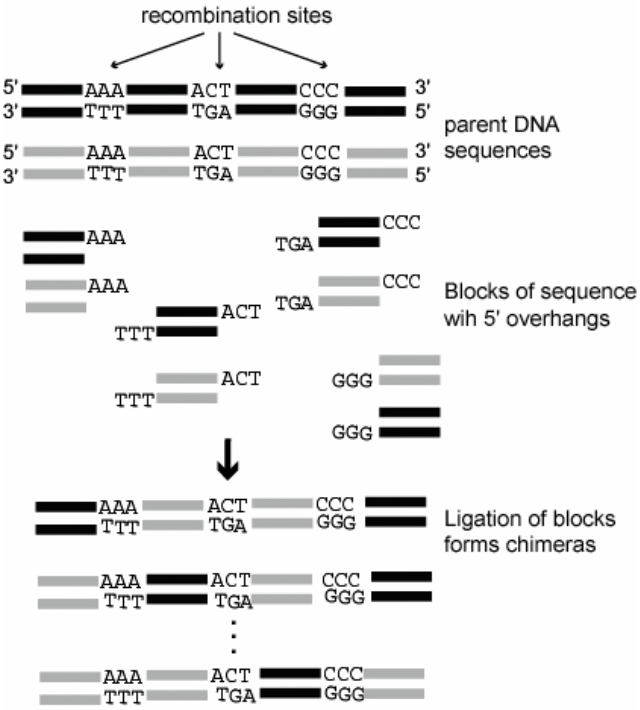

Figure IV-1. Overview of combinatorial gene assembly using sequence blocks of any sequence identity.

There are several potential methods for creating the necessary sequence blocks. As the cost of DNA synthesis has decreased, the easiest method is to simply order the gene fragments as oligonucleotides. These oligonucleotides can be phosphorylated, annealed and used directly without further modification. However, the blocks of sequence must be relatively small (<20-25 amino acids) so that the DNA segments

purchased are not too long. Longer oligonucleotides have a greater probability of incorporating a deletion. This synthetic methodology was used in the construction of the library described in Chapter II (Meyer et al. 2003). Alternatively, the gene fragments can be PCR amplified and separated with tag sequences that can generate any desired overhang when cleaved with a Type IIB restriction enzyme (cleaves outside its recognition site) (Hiraga and Arnold 2003). This methodology permits the incorporation of larger gene fragments, and has the advantage errors can be minimized by cloning and sequencing the PCR amplified gene segments prior to library construction.

In this chapter we describe the construction and characterization of the lactamase libraries designed in Chapter III. They were constructed using the two different variations on combinatorial ligation described above. Once a library is constructed, information about the chimeras must be obtained in a high-throughput manner so that a large number of chimeras can be assessed. The goals of this work include exploring what altered substrate specificities might be obtained in chimeric proteins, as well as investigating the properties of folded and functional chimeras. To meet these goals it is necessary to obtain sequence, function, and folding information for a large number of the chimeras created. High-throughput techniques exist for gathering much of this information. However, several methods required adaptation for this particular system.

Results

Construction of the RandE:APST lactamase library

The RandE:APST lactamase library was designed in Chapter III using random enumeration. It consists of four parents (TEM-1, PSE-4, AST-1 and SED-1) with nine

exchangeable fragments. The N- and C-termini are fixed to the same parent, and there are eight blocks between them. This library was constructed through sequential ligation of purchased oligonucleotides. All of the library blocks with the exception of the N- and C-termini are less than 21 amino acids. Since the N- and C-termini are always fixed to the same parent, a separate plasmid containing the termini was constructed for each parent (Figure IV-2A). Thus to construct the complete library, a set of reconstructed library blocks (2-9) must be ligated into each parental plasmid. Between the termini there is a cassette that contains a stop codon in each frame to prevent any translation before the complete library is added. This cassette is removed to construct the final libraries.

The purchased oligonucleotides were phosphorylated, annealed, and subsequently ligated in a sequential scheme, outlined in Figure IV-2B. There was some difficulty obtaining sufficient material for ligation into the final plasmid, which was remedied by PCR amplification of the reconstructed blocks prior to the final ligation. A second difficultly was encountered with one of the parent proteins, AST-1. AST-1 was originally cloned with a GTG start codon (Poirel et al. 2001). In our hands, the clones failed to confer resistance in our expression system. Additionally, AST-1 required extreme PCR conditions compared with the other parents (see methods) which might result in biases in the final library. As a result, AST-1 was dropped from the library. The library actually constructed therefore consisted only of parents TEM-1, SED-1 and PSE-4 and was known as RandE:PST. As discussed in Chapter III, this library is significantly smaller than the designed library (3^9= 19,653 vs. 4^9=262,144). The trade-off between diversity and fraction folded, however, remained similar to that of the original library (Table III-1).

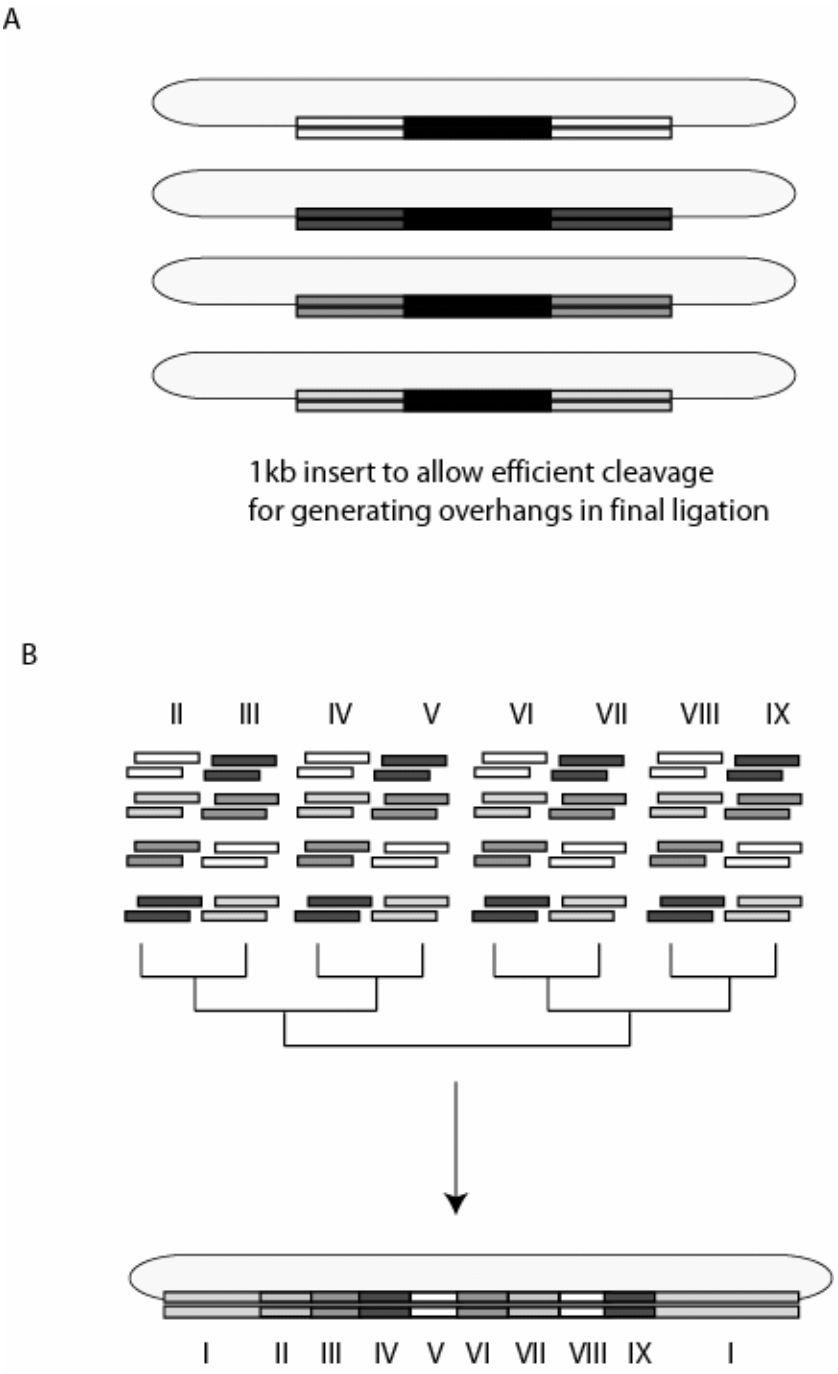

Figure IV-2. Overview of construction methodology for RandE:APST, and RandE:PST libraries. A: The N- and C-termini for each of the four parents were placed into a plasmid with a stop cassette between them. B: Annealed and phosphorylated oligonucleotides were ligated in series (see methods) to generate a full-length insert. This insert is ligated between the N- and C-termini from each parent to generate the full-size library.

To examine the quality of the RandE:PST library, 20 chimeras were completely sequenced. This sequencing revealed 2.5-3 basepair mutations per chimera and 0.6 deletions per chimera on average. These events were spread throughout the oligonucleotide-derived portion of the library. Due to the high deletion rate, a large proportion (~40%) of the library was out of frame. The oligonucleotides used for construction were on average 53.4 bp, not significantly longer than those commonly used for many molecular biology applications. However, to construct a perfect chimera with no mutations or deletions, 16 such oligonucleotides must all be perfect. The nucleotides purchased were cartridge-purified, which is not sufficient for this application. This library was not characterized further due to its high deletion rate.

Construction of RASPP:PST lactamase library

The RASPP:PST library described in Chapter III (Table III-1) designed using RASPP for three parents (TEM-1, PSE-4 and SED-1) with eight exchangeable fragments was constructed using Sequence Independent Site-Directed Chimeragenesis (SISDC) (Hiraga and Arnold 2003); an overview is in Figure IV-3. This method involves PCR amplifying the gene fragments to insert sequence tags. These tags are later removed using a type IIB restriction endonuclease (BsaX1) to generate the specific basepair overhangs necessary for ligation. The only problem encountered with this methodology is that one of the gene fragments was small (<30 bp) and was consistently lost during one phase of the procedure (see methods). To remedy this, oligonucleotides corresponding to the gene fragments were purchased and added to the ligation reactions as phosphorylated and annealed gene fragments. The oligonucleotides were short (24 bp) and PAGE purified.

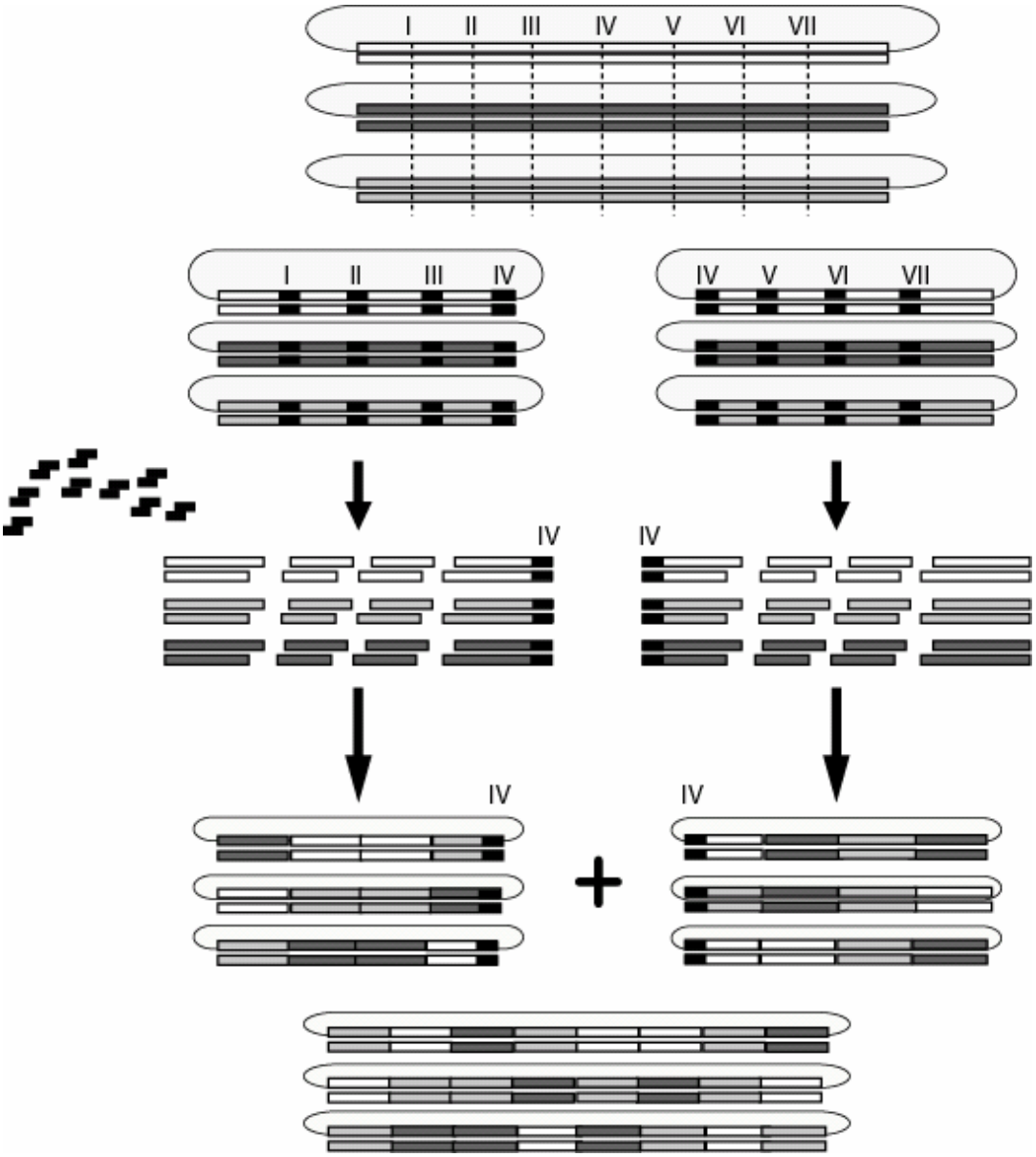

Figure IV-3. Overview of construction methodology for the RASPP:PST library. Tag sequences that will allow specific overhangs to be generated are inserted into the genes using overlap-extension PCR. The tag-inserted genes are cut with a type IIB restriction enzyme to expose the DNA fragments with desired overhangs, and the tag sequences are removed. The DNA fragments are then ligated together to form two minilibraries which are cloned individually. Finally, the two mini-libraries are then ligated to form full-length genes. Sequences cloned and transformed into *E. coli* are shown with the plasmid backbone.

To assess the quality of the library, 48 randomly chosen chimeras were sequenced. The rates of single-base mutation and basepair deletion are much lower than those observed for the RandE:PST library. However, with 2 mutations and 10 deletions affecting 11 of 48 chimeras, deletions are still prevalent. Of the 10 deletions, 3 were found at segment junctions and the remaining 7 were found in regions within PCR primers used after the block assembly during construction, usually at the N-terminus. None of the deletions or mutations was found within the small block added as an oligonucleotide. Additionally no deletions were detected while sequencing half-length chimeras generated during the construction procedure (see methods).

The high rate of single basepair deletion observed in 19% the full-length chimeras may occur because producing no protein (frame shift in first few amino acids) is more favorable than producing large amounts of unfolded protein. Thus chimeras with deletions are slightly favored over those without under nonselective conditions. We have previously observed that expression of lactamase chimeras under nonselective conditions can affect fragment biases in the library, presumably because some fragments are potentially deleterious (Hiraga and Arnold 2003). As a consequence of the deletion rate, there are potentially a large number of false negatives (up to 4% of characterized chimeras). However, 23% of the functionally characterized chimeras were observed multiple independent times. Any contradictions in functionality assignments were explicitly examined, further reducing the number of erroneous functionality assignments. Anecdotally, several chimeras that confer resistance to ampicillin contain deletions in the first few amino acids. This may occur because the first 24-30 amino acids comprise a

periplasmic targeting sequence and a downstream ATG permits sufficient expression to confer resistance.

High-Throughput Sequencing

The DNA sequences of 811 randomly chosen chimeras were determined by DNA probe hybridization, to obtain 553 unique sequences (Meinhold et al. 2003). To assess the error rate in the hybridization, 48 randomly chosen chimeras were sequenced. The probe hybridization is accurate, with 47 of 48 sequences correctly assessed. Examining the composition of the characterized sequences on a ternary diagram shows that the characterized library does not have equal representation of the different parents (Figure IV-4A). In particular, few chimeras similar to PSE-4 and many similar to TEM-1 were characterized. The proportion of the different parents at each position shows that PSE-4 was severely underrepresented at block 8 (Figure IV-4B). This discrepancy is due to an error in construction; a restriction site in block 8 from PSE-4 was used in the construction process. Chimeras that do contain PSE-4 at block 8 are a result of incomplete cleavage of the site. The library properties do not change significantly if all chimeras containing PSE-4 at block 8 are omitted ($<m>=67 \pm 36$ and $<E>=43 \pm 21$ for all chimeras lacking PSE-4 at block 8 vs. $<m>=66 \pm 21$ and $<E>=45 \pm 15$ for all chimeras in library). However, the library size is reduced by 1/3. Additionally TEM-1 is favored at most positions, block 3 most strongly. Because one parent (TEM-1) is more likely to be found at all of the blocks whose frequencies were determined from the original DNA mixing (3, 6, and 7), it is possible that the unequal representation of the different parents in the characterized library is due to improper quantification of the DNA during the initial

stages of construction. It is also possible that some clones do not survive to the characterization stage because the expressed protein is deleterious. Examination of the *E* and *m* distributions of the characterized chimeras shows that the characterized library has roughly the same distributions as the theoretical library despite its biases (<*E*> = 44 ± 17 and <*m*> =66 ± 22 for theoretical library, <E>=45 ± 15and <m>=66 ± 21, for characterized library, Figure IV-5)

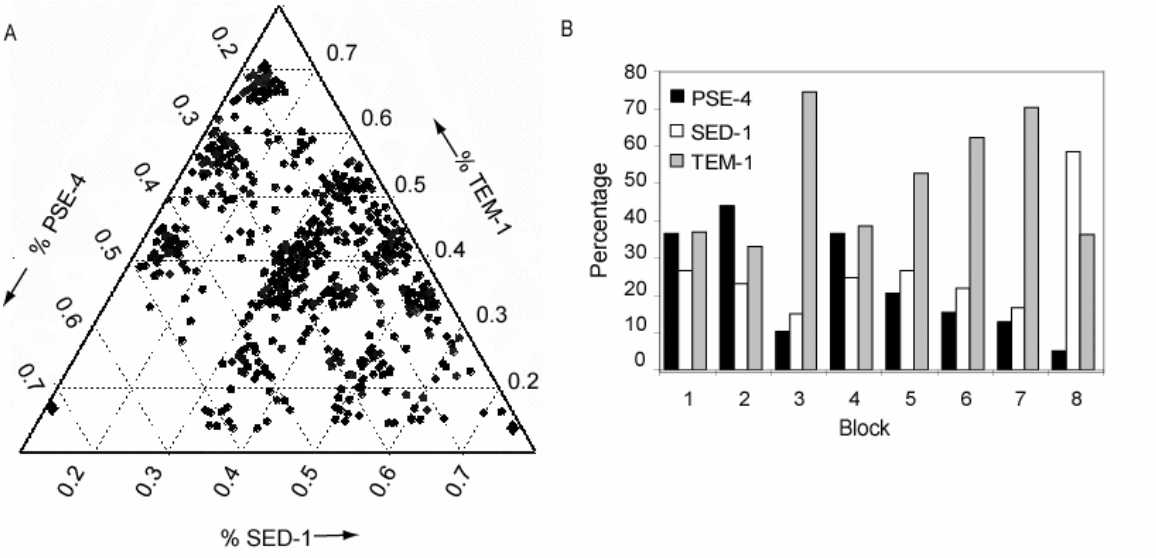

Figure IV-4. A: Ternary diagram showing the compositions of the 553 characterized chimeras. Characterized chimeras do not evenly populate the available sequence space, but are biased toward some areas. The position of each point is determined by the relative similarity of the chimera to each of the parents. To establish the location of a point on the ternary diagram the number of amino acids a chimera shares with each parent sequence is determined. Positions where there is no variation among the three parents are not included. Including such positions does not change the qualitative representation but merely shrinks the diagram into a smaller spread of space. The similarity of the chimera to each parent is then normalized by dividing by the sum of the similarities to each parent. B: The proportion of each parent protein at each block of the library. A perfectly balanced library would have 33% of each parent at each position.

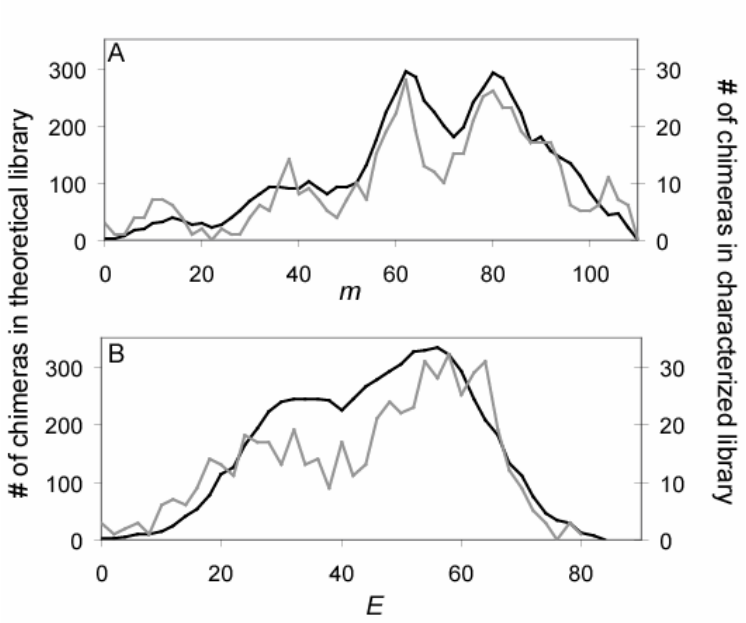

Figure IV-5. Distribution of chimeras with respect to *m* (A) and *E* (B) in the theoretical (black line) and characterized (gray line) libraries shows that the characterized library has approximately the same distribution as the theoretical library.

Evolution of New Function

One of the goals of this project is to determine whether site-directed recombination can generate chimeras with new functionality. We searched for chimeras with resistance to extended-spectrum antibiotics using functional selections. Unfortunately, there were two confounding factors in this process. First, SED-1 in our hands was significantly more resistant to most cephalosporins than originally described in the literature (Petrella et al. 2001); this has since been reexamined by the authors (Petrella et al. 2004). Second, most β-lactam based antibiotics are cell density dependent. This property makes it very easy to isolate false positives. The lactamase parents were tested against 11 different antibiotics (Table IV-1). For many of these, SED-1 displayed significantly more activity than TEM-1 or PSE-4. The RASPP:PST library was tested using antibiotics to which the lactamases displayed relatively low resistance (see methods). Typically the antibiotic concentration was lowered to the point where false

positives were isolated due to the density dependency of the antibiotic. No chimeras with significantly improved resistance to any of these antibiotics were isolated.

Table IV-1. MICs (μg/mL) of TEM-1, PSE-4, and SED-1 on β-lactam and Cephalosporin Antibiotics

antibiotic	TEM-1	PSE-4	SED-1
ampicillin	>2000	>2000	>2000
cefamandole	>2000	500	>2000
cephalothin	2000	1000	>2000
ceftazidime	<1	<1	200
cefoxitin	100	200	200
cefoperazone	>2000	1000	>2000
cefotaxime	1	20	>2000
ceftriaxone	2	40	>2000
cefsulodin	1000	500	>2000
carbenicillin	>2000	>2000	>2000
moxalactam	1	2	10
aztreonam	1	2	>50

It is somewhat surprising that no chimeras with increased resistance to cephalosporins were identified. Previous studies with TEM-1, and the profusion of natural TEM-1 variants, indicate that it is relatively easy to obtain extended-spectrum activity in TEM-1. However, many of the single mutations introduced to give TEM-1 extended-spectrum activity are not incorporated by our recombination. Additionally, because SED-1 already confers a higher level of resistance to many of these antibiotics, it is more difficult to identify variants with increased resistance. SED-1 limits the number of possible substrates, and the baseline antibiotic concentrations used in the selections are much higher than they would be for PSE-4 and TEM-1 alone (Table IV-1). It may be that recombination is not a good strategy for improving functions, but rather is more

exploratory for finding different functions. It is also possible that the sequences used in this work do not have the "right" sequence diversity within them to increase activity toward cephalosporins.

Folding

There are several different high-throughput methodologies in the literature that purport to detect a correctly folded protein (Maxwell et al. 1999; Waldo et al. 1999; Philipps et al. 2003). Most of them actually measure the amount of soluble protein in the cell because this corresponds well with properly folded protein. One potential complication in using many of these methods is that β-lactamases are exported to the periplasm and have an N-terminal signal sequence. We chose to implement one of these methods by fusing GFP to the N-terminus of the lactamases. Good signal differences between positive and negative controls were achieved under high-throughput conditions (Figure IV-6). However, when a library of clones was examined, the distribution of values obtained made it difficult to assess where the line between folded and unfolded should be drawn. Additionally there were several sequences which displayed very low fluorescence (less than negative control) but retained resistance to ampicillin (Figure IV-7), indicating chimeras may still be capable of catalysis, but not accumulate large quantities of protein in the cell. This observation indicates that any folding assay based on measuring soluble protein may not accurately describe those proteins that are folded enough to maintain catalytic activity but do not accumulate in the cell. Therefore, rather than continue to pursue a folding assay, we concentrated on measuring base line catalytic function.

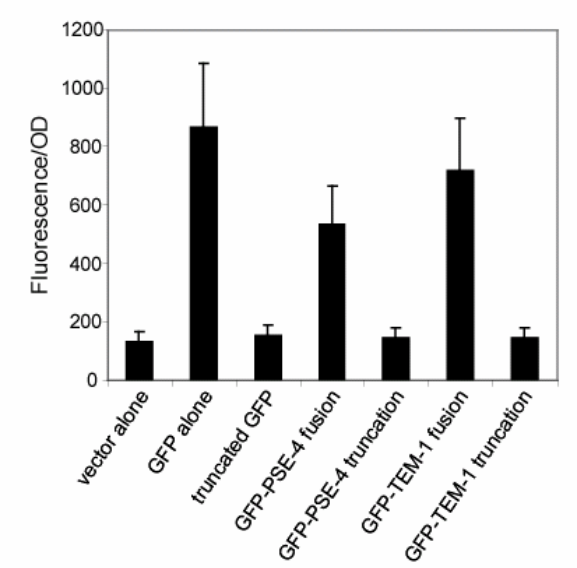

Figure IV-6. Fluorescence measurements of control strains used for GFP folding screen showing a good difference between positive and negative controls.

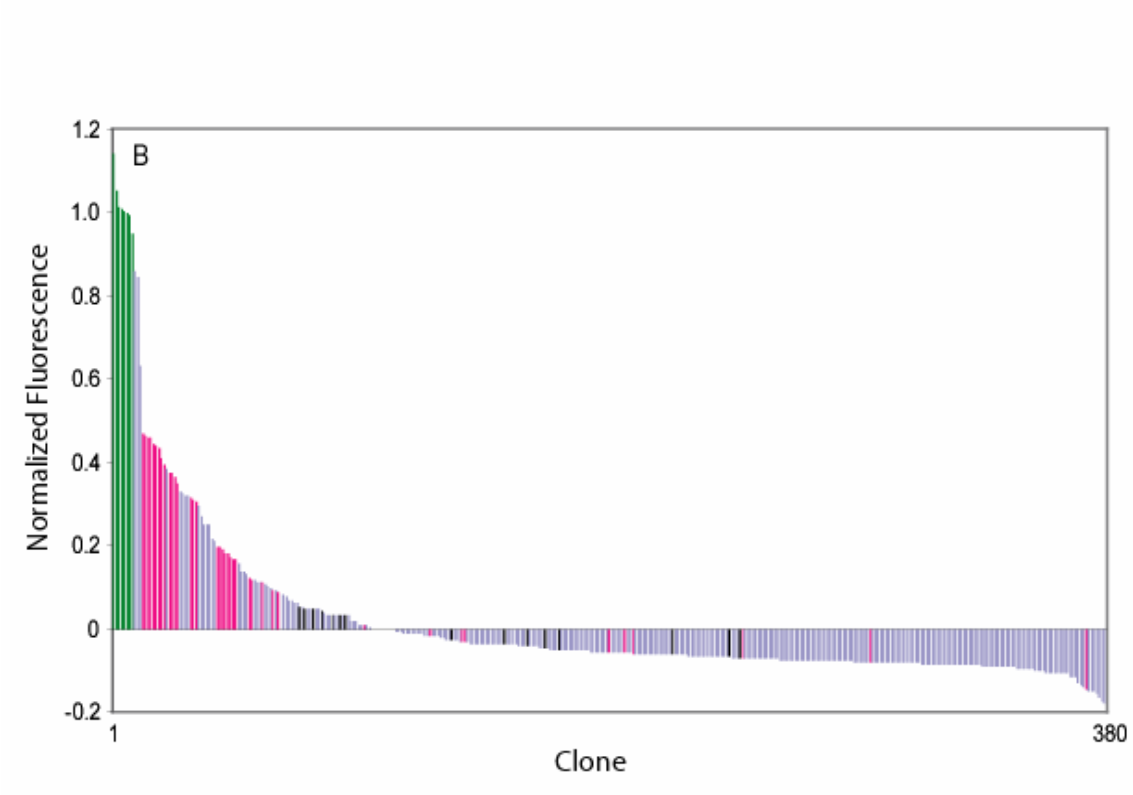

Figure IV-7. Normalized fluorescence from a library of different clones in the GFP screening system. Green bars are positive controls, black bars are negative controls, and pink bars are ampicillin resistant clones. All values are normalized between zero (average of negative controls) and one (average of positive controls). The distribution of values obtained for chimeras makes it difficult to draw a line between "folded" and "unfolded." Additionally some chimeras displaying less fluorescence than negative controls conferred resistance to ampicillin.

Retention of Function

Because the folding assays were unable to distinguish folded and unfolded chimeras, a low stringency functional screen was used to assess which chimeras retained basic catalytic function, and thus a folded structure. Chimeras in the RASPP:PST library were screened for a function shared by all three parents, the ability to confer ampicillin resistance. The screen was conducted at very low stringency (>500x lower concentration of ampicillin than the wild-type MIC) to capture chimeras with even very minimal activity. Of the 554 unique sequences tested, 20% (111) conferred resistance to ampicillin and are considered functional lactamases. An additional 51 unique functional lactamase sequences were obtained by selecting functional clones prior to probe hybridization sequencing, giving a total of 162 functional lactamases. A complete listing of all chimera sequences and their functionality status can be found in Appendix III. Of the functional chimeras, 51% conferred a MIC of 2,000 μg/ml ampicillin or greater, indicating approximately wild-type activity (~5,000 μg/ml for all three parents). 10% of chimeras displayed a MIC of 50 or below, indicating weak activity**.** Chimeras that did not confer resistance to ampicillin may be not folded, may not be well expressed, may be folded but not catalytically active, or may have a combination of these properties. Because the screen was very low stringency, chimeras that are well-expressed, folded proteins with any catalytic activity are likely to have been identified.

Discussion

Examining the naïve data set of 553 chimeras of which 111 (20%) are folded shows that, like the previous lactamase chimera library (Meyer et al. 2003), chimeras

with low E are more likely to function than chimeras with high E (Figure IV-8). Unlike the library described in Chapter II, the decline in probability of retaining function with respect to E is not exponential (Meyer et al. 2003), but instead is more reminiscent of the sigmoidal function originally described by Voigt et al. (2002). The difference in the form of P_f is not surprising. This is a designed library, where the distribution of chimeras is skewed toward those that are likely to fold. The distribution of chimeras in the library affects the form P_f with respect to E can take when measured with any given library. This is also a more accurate estimation of P_f because all the chimeras used in the calculation have been explicitly observed. In the previous work (Chapter II) we observed only functional chimeras directly and assumed that most other were nonfunctional. This assumption can lead to a less accurate description of P_f. Chimera probability of functioning decreases with increasing m (Figure IV-8). However, there is not a simple function that describes the behavior P_f with respect to m. The slightly bimodal behavior is a result of the underlying bimodal distribution of m of the chimeras in the characterized library (Figure IV-5A).

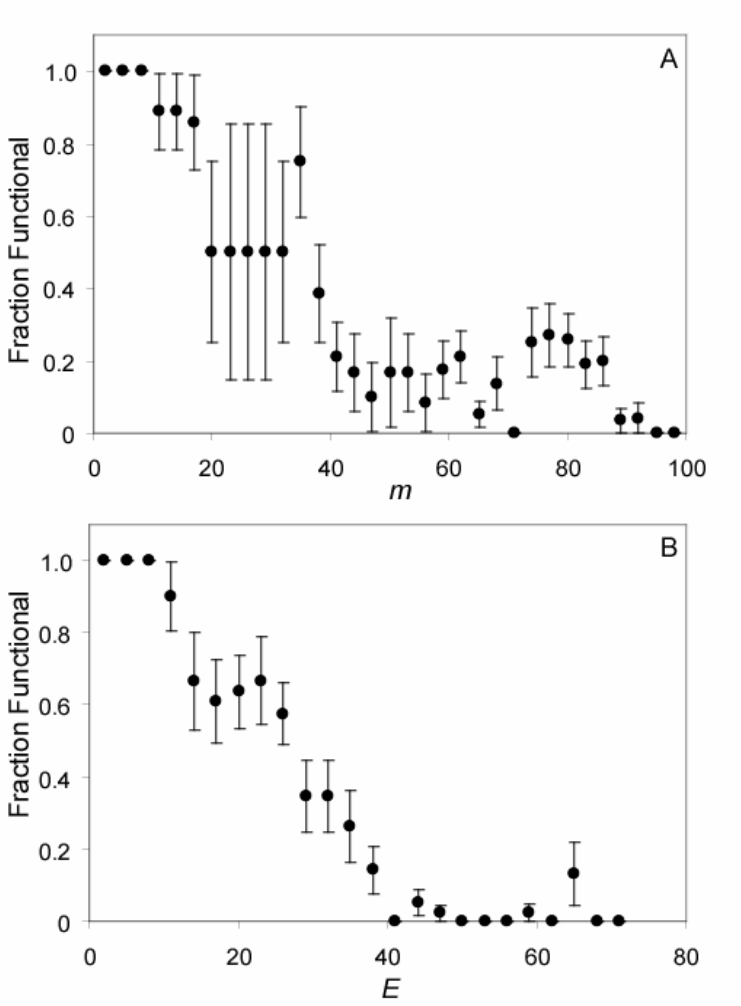

Figure IV-8. A: Fraction of functional chimeras with respect to *m.* B: Fraction of functional chimeras with respect to *m.*
Chimeras were put into bins of 3 *E* or *m* to ensure representation in all bins and the fraction functional *(*$F_{functional}$*)* is = $N_{functional}/N$ where N is the total number in the bin and $N_{functional}$ is the number functional. The error was approximated by the standard error on an estimate of the binomial proportion (sqrt (($F_{functional}$*(1-$F_{functional}$)/N))).

The 162 functional chimeric lactamases span a range of mutation levels compared to the parental proteins, and contain up to 80 mutations from the closest parent. One-third of active chimeras displayed ≤75% sequence identity to any known lactamase. Five concentrated clusters of sequences account for 77% of the functional chimeras (Figure IV-9). Within these clusters, sequences share on average 95% identity. The clusters result from several different factors including uneven block sizes, sparse sampling of the theoretical library, and favorable or unfavorable block interactions. It is likely that there are other such clusters which are not observed due to differences between the characterized and theoretical library.

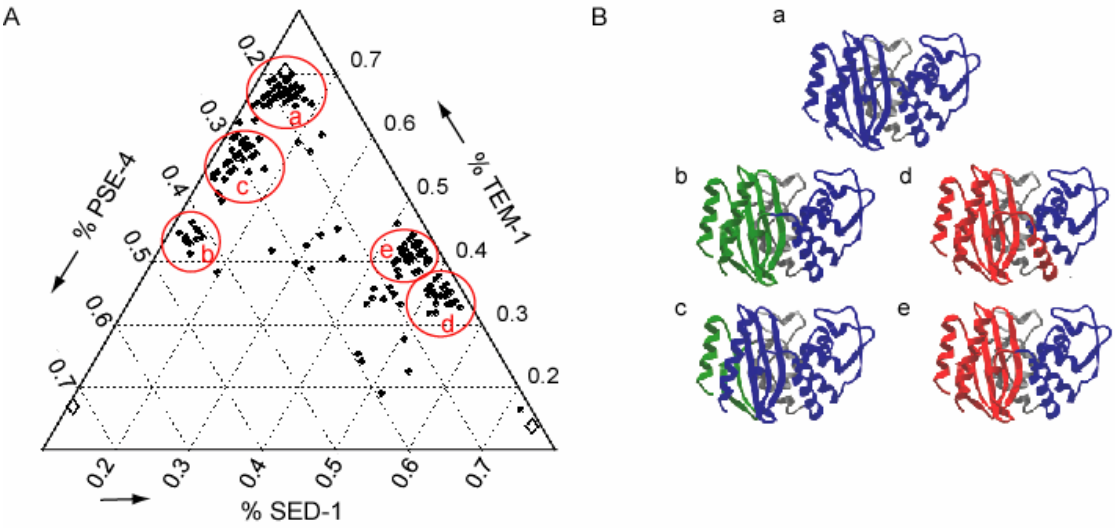

Figure IV-9. Ternary diagram similar to the one in Figure IV-4 representing only chimeras that display ampicillin resistance. The chimeras fall into five main clusters which can be described by which parents the four largest sequence blocks (1, 3, 7, and 8) are inherited from. The structures show which parents each cluster inherits its major blocks from: green ribbons indicate PSE-4, blue TEM-1 and red SED-1, gray blocks are variable within each cluster.

Chimeras in the first cluster (Figure IV-9, cluster a) have all the large blocks (1, 3, 7, and 8) from TEM-1. On average these chimeras differ from TEM-1 by only 12 mutations. Chimeras in the second cluster (Figure IV-9, cluster b) have the N- and C-termini from PSE-4, and blocks 3 and 7 from TEM-1. These chimeras are significantly different from both TEM-1 and PSE-4, and an average of 74 mutations from the closest parent. Chimeras in the third cluster (Figure IV-9, cluster c) have the N-terminus from PSE-4 and the remaining large blocks (3, 7, and 8) from TEM-1. These chimeras have on average 37 mutations to TEM-1. The last two clusters are sequences comprised mostly of TEM-1 and SED-1. Chimeras in the fourth cluster (Figure IV-9, cluster d) of these have the N- and C-termini as well as block 7 from SED-1 and block 3 from TEM-1. These chimeras have on average 60 mutations to SED-1. The fifth cluster (Figure IV-9, cluster e) has the N-and C-terminus from SED-1 and blocks 3 and 7 from TEM-1. These

sequences are the most distant from any of the parents, with an average 78 mutations to the closest parent.

There are 1785 lactamase sequences in the PFAM database for protein families (Bateman et al. 2004), at least 450 of which are class A lactamases by phylogenetic analysis. New lactamase sequences continue to be identified. However, many of the characterized lactamases are minor variations of a few very prevalent sequences. For example, there are over 100 characterized variants of TEM-1, differing from TEM-1 by only a few amino acids (Jacoby and Bush 2005). The lactamase skeleton seems relatively tolerant to mutation. 220 of 263 positions in TEM-1 accept at least one other amino acid when mutated in isolation (Huang et al. 1996), and several other experiments indicate that PSE-4 and TEM-1 can easily tolerate minor modifications (Petrosino and Palzkill 1996; Matagne et al. 1998; Sanschagrin et al. 2000; Osuna et al. 2002).

The clusters of functional lactamases observed here represent regions of sequence space that are not populated by known natural β-lactamases. Not including the cluster most similar to TEM-1, the cluster consensus sequences range from 72% to 80% identity to any natural lactamase. However, these areas appear relatively densely populated with catalytically active and folded proteins. While functional lactamases cluster toward some areas of sequence space, it is not known based on these simple observations whether they cannot occupy others.

Methods

All enzymes used were purchased from New England Biolabs and all chemicals were purchased from Sigma unless otherwise indicated.

Construction of RandE:APST and RandE:PST Libraries (purchased oligonucleotides)

Four parent plasmids were constructed by PCR amplifying the N- and C-termini from each parent separately and combining them by overlap extension PCR with a cassette that consisted of stop codons in each frame and an out of frame segment from P4501A2. TEM-1, SED-1 and PSE-4 all amplified under standard PCR conditions. AST-1 required the addition of 5% DMSO as well as a denaturation temperature of 98 °C. These constructs were placed into the expression plasmid pProTet E.333 (Clontech). To release the correct overhangs, the final plasmids were digested with SapI, and treated with alkaline phosphatase. Doubly cut plasmid was separated from linearized plasmid by agarose gel electrophoresis. The desired product was recovered from the gel and purified using a Zymocan DNA gel recovery kit.

Oligonucleotides that compose the smaller sequence blocks for the library described in Chapter III were ordered from Invitrogen as cartridge purified stocks (Table AII-1). Complementary oligonucleotides were annealed (1.25 mM each primer, 50 mM NaCl, heat to 95 °C for 2 min and ramp 1 °C/s to 4 °C) and then phosphorylated (10U T4 polynucleotide kinase, 37 °C, 1 hr in T4 ligase buffer). The annealed and phosphorylated oligonucleotides from all parents for each sequence block were mixed and ligated together in pairs (2-3, 4-5, 6-7, 8-9) using T4 ligase at 16 °C for 5 hours. Following ligation, the correct product was isolated by gel electrophoresis and purified using a

Zymoclean DNA gel recovery kit. The ligation process was repeated with the products until full length inserts (8 pieces total) were obtained. This product was PCR amplified using *Pfu* polymerase (Stratagene) with primers from all possible parent pairs (Table AII-2). The final product was cut with Sap1 to generate the overhangs and ligated as above into the four parent plasmids. The number of clones obtained for each parent plasmid varied between 5,000 and 13,000.

Construction of RASPP:PST Library (SISDC)

Each fragment was PCR amplified (*Pfu* Ultra, Stratagene) with primers to introduce a tag sequence at all internal junctions (see Table AII-3) for primers and sequence fragments). All tag sequences contained BsaX1 and Nde1 restriction sites, as well as a unique region for each junction. Additionally each junction was designed to have a unique 3bp sequence found in all three parents which is released upon BsaX1 cleavage (See Table AII-3). There is no tag sequence between segment 4 and 5, but rather a Sap1 restriction site accompanied by a Pst1 restriction site (fragment 4) or Sal1 restriction site (fragment 5) was inserted. The first and last four segments of each parent were separately reconstructed using overlap extension from PCR amplified segments, each resulting in a half-length gene product with tag sequences inserted at the junctions. The N-terminal half was cloned into pBC (SK+) with Sal1 and Pst1 sites on the forward and reverse primers, and the C-terminal half with Xho1 and Pst1 sites. The primer sequences for PCR reactions for tag insertion can be found on Table III-1.

The DNA sequence of each half-library parent was confirmed. The DNA for each half library was mixed in equal proportions based on spectrophotometric quantification,

then cut with Sal1/Pst1 (front) or Xho1/Pst1 (back), dephosphorylated, and the insert purified by agar gel electrophoresis. This insert was then cut with BsaX1 to remove the tag sequences and generate 3 bp overhangs. Following column purification (Zymogen Clean and Concentrate) to remove the tag sequences, fragments for block 2 were added as annealed and phosphorylated oligonucleotides, and the mixture was ligated using T4 ligase for 5 hours at 16°C. The ligation was column purified, cut with Nde1 and BsaX1 to remove any incompletely cleaved tag sequences, and then PCR amplified with 9 primer sets (see Table AII-4) to generate the complete library. The PCR reactions were mixed in equal proportions based on agarose gel quantification. The N-terminal half-library was cut with Xho1/HindIII and ligated into pBC cut with the same enzymes. The C-terminal half-library was cut with Sal1/Pst1 and ligated into pProTet (Clontech) cut with the same enzymes. The resulting DNA was transformed into DH5αPRO to prevent expression from the Tet promoter on pProTet.

A few thousand clones were obtained for each half-library, sufficiently higher than the expected complexity (81 sequences) of each half. These colonies were pooled, and the DNA was purified from them (Qiagen midi-prep). The N-terminal half-library was removed from pBC using Kpn1 and Sap1 restriction sites and inserted into the C-terminal half-library cut with the same two enzymes to reconstruct full-length genes. This ligation was transformed into XL-l Blue (Stratagene) and the clones used directly for analysis.

High Throughput Sequencing

To determine the block sequence of the chimeras, ~1100 clones were picked and grown overnight to saturation in 384-well plates containing 70 µl of LB + 35 µg/ml chloramphenicol. Each plate contained four samplings from each parent as well as the expression plasmid (pProTet) containing no lactamase insert. The plates were stamped onto N+ Hybond membranes (Amersham) layered onto 2% agar LB plates and allowed to grow at 37 °C for 18 hours. The membranes were removed from the plates, the cells lysed according to Meinhold et al. (2003) and the DNA attached to the membrane through UV cross-linking. The membranes were dried and stored at 25 °C for up to 1 month.

Probes for used for hybridization are listed in Table AII-5 and were labeled with DIG-dUTP using Roche DIG Oligonucleotide 3'-End labeling Kit, 2nd Generation according to the manufacturer's instructions. Hybridization was performed at 58 °C, and the stringency washes carried out at 53 °C in 2x, 1x or 0.5 x SSC +0.1% SDS depending on the probe (Table AII-5). Probes were detected using a Roche DIG nucleic acid detection kit according to the manufacturer's instructions and visualized using Kodak MRX film.

New Antibiotics

The MICs of 11 different antibiotics was determined for TEM-1, SED-1 and PSE-4 by spotting saturated culture onto an agar plate containing the antibiotic and 35 ug/mL chloramphenicol (Table IV-1). The MIC was measured as concentration of antibiotic which prevented visible growth. These conditions simulate antibiotic screening rather

than selection. For antibiotic library selections (moxalactam, ceftazidime, and cefoxitin) the MIC determined above was used as the starting concentration of antibiotic. The concentration was progressively decreased until colonies were observed on a negative control plate. To search for chimeras with increased resistance to these antibiotics, library plasmid DNA was transformed into XL-1 Blue (Stratagene) according to the manufacturer's instructions. 200 uL of cells was spread on each 100 x 15 mm plate LB agar plate containing antibiotic, and 10 uL was spread on a nonselective plate to determine the approximate number of colonies obtained. The plates were grown for 18 hours at 37 °C. Colonies were restreaked onto plates with the same antibiotic concentration to verify resistance.

Folding

The GFP folding assay was implemented similarly to Waldo et al. (Waldo et al. 1999). Briefly GFPuv (Clontech, from pGFPuv) was placed N-terminal to the lactamase in pProTet. The signal sequences were removed to residue 24 for PSE-4 and 26 for TEM-1. SED-1 was never tested in the folding assay because the assay failed to distinguish between folded and unfolded chimeras cleanly. For PSE-4 the linker was Gly-Ser-Ala-Gly-Ser-Ala-Asn-Ala-Ser-Gly, an additional Ser-Gly was added directly before TEM-1. An NsiI restriction site was incorporated within the linker. To place a library into the expression system, a negative control protein was removed and the PCR amplified library incorporated using NsiI and PstI. Chimeras expressed in BL-21 were grown in deep-well plates containing 1 mL M9 medium with 35 ug/mL chloramphenicol at 30 °C, 220 rpm, 80% humidity for 18 hours. The plates were then centrifuged to pellet cells and stored at

4 °C for 24 hours. Cells were rinsed with 500 uL PBS and then resuspended in 300 uL PBS. OD_{600} and fluorescence (excite 395 nm, emit 509 nm) were measured.

Ampicillin Activity Screen

To screen for chimera function, deep-well 96-well plates containing 500 μl of LB medium with 35 μg/ml chloramphenicol were inoculated from the 384-well plates used for hybridization and allowed to grow at 37 °C for 18 hours 220 rpm 80% humidity. Approximately 2 μl aliquots of each culture were transferred to LB agar plates containing varying concentrations of ampicillin (0, 5, 10, 25, 50, 100, 250, 500, 1,000, 2,000 μg/mL) using the 96-well stamp and allowed to grow for 18 hours. Duplicate plates were generated at each concentration. After 18 hours the plates were observed for growth. Chimeras growing at concentrations of ampicillin 10 μg/ml or greater were considered positive. XL-1 containing pPro with no lactamase insert survive to 5 μg/ml ampicillin in this assay. The concentration of ampicillin necessary to prevent growth was recorded as the MIC. Chimeras that grew on the 2,000 μg/mL plates are recorded at 2,000+.