Chapter V: Using Chimeras to Identify Determinants of β-lactamase Function

Introduction

The most informative techniques for probing the relationships between protein sequence, structure and function are those that perturb a natural protein sequence to examine the properties of the new protein. Site-directed mutagenesis has become a standard tool for determining if a particular amino acid is necessary for a specific protein property, whether the property is folding, substrate specificity or catalytic activity. However, using mutagenesis alone it is difficult to explore properties that are not specifically tied to one or a few amino acids such as dynamics or allostery. Multiple sequence alignment (MSA) analysis of protein families has allowed the identification of energetic coupling within proteins (Lockless and Ranganathan 1999). However, natural sequences are under selection for additional properties besides the property under investigation and it can be difficult to discern which attributes are responsible for the property of interest.

Recombination of homologous proteins allows construction of proteins that are significantly different from natural proteins. This allows differences between homologous proteins as well as the determinants for a particular protein fold to be examined. Characterization of chimeric proteins in small studies has contributed to understanding product or substrate specificity (Kushiro et al. 1999; Nicot et al. 2002), as well as key elements for folding (Morimotoa and Tamura 2004). However, these data sets are invariably small and conclusions are drawn based on only a few chimeric sequences.

We have created and characterized a large number of chimeras made by recombining distantly related β-lactamases TEM-1, PSE-4 and SED-1. By examining the functional chimeras we can explore which portions of SED-1 are key contributors to the altered substrate specificity that corresponds with extended-spectrum activity. Mutagenesis studies and analysis of multiple sequence alignments and several crystal structures have generated many hypotheses about the sequence determinants of this altered substrate specificity, but few concrete answers.

We have previously observed that the functional chimeras cluster into a few areas of the possible sequence space (Figure V-1). However, examination of the functional chimeras alone does not provide enough information to determine if this is due to our sparse sampling of the theoretical library, or whether some areas are not compatible with functional lactamases. In addition to the many functional lactamase chimeras, we have also generated and characterized a large number of nonfunctional chimeras. Using both sets of sequences we can determine whether the clusters of sequences we observe are caused by inherent limitations of the protein fold, or by our sparse sampling theoretical library.

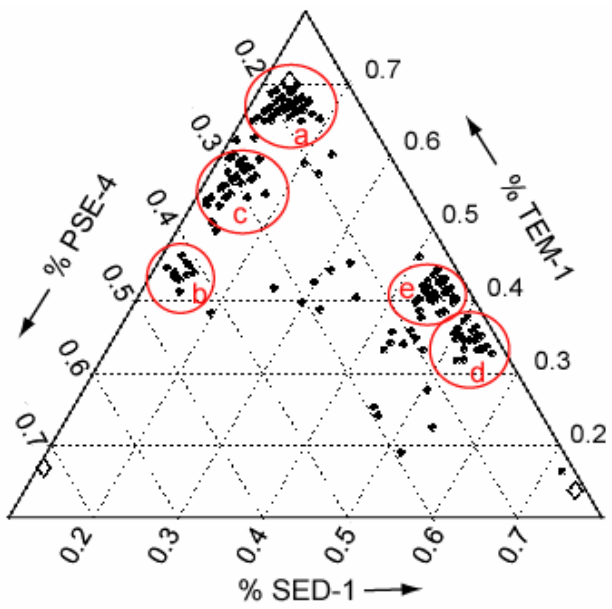

Figure V-1. Ternary diagram showing lactamase chimeras that display ampicillin resistance. The position of each point is determined by the relative similarity of the chimera to each of the parents. To establish the location of a point on the ternary diagram the number of amino acids a chimera shares with each parent sequence is determined. The similarity of the chimera to each parent is then normalized by dividing by the sum of the similarities to each parent. The chimeras fall into five main clusters (a, b, c, d, and e) which can be described by which parents the four largest sequence blocks (1, 3, 7, and 8) are inherited from. The ternary diagram represents compositional space. However sequences clustered on the ternary diagram tend to be clustered based on sequence identity as well. Cefotaxime resistant chimeras fall into cluster d.

Results and Discussion

Determinates of Cephalosporin Resistance

SED-1 is an extended spectrum CTX-M type lactamase that has significant activity toward cefotaxime, while TEM-1 and PSE-4 do not. CTX-M type lactamases are a class of extended-spectrum lactamases that have recently been isolated that are not similar to extended-spectrum TEM-1 variant and they do not simply widen the active site to alter substrate specificity (Orencia et al. 2001). Instead the source of the altered substrate specificity appears to originate from many sequence changes, and remains difficult to pinpoint. To identify which portions of the protein are critical for cefotaxime

resistance we screened characterized chimeras for resistance to cefotaxime. Twenty sequences displayed >10-fold higher cefotaxime resistance (0.1 μg/ml) compared to PSE-4 and TEM-1 and were considered positive on cefotaxime (Table V-1). The cefotaxime resistant sequences appear similar upon inspection; they share blocks 1, 7, and 8 with SED-1. These sequences are nearly all (18 of 20) from one of the clusters of sequences identified in Chapter IV (Figure V-1). All of them conferred resistance to ampicillin, usually to high levels of ampicillin (>1,000 μg/ml). Of the proteins conferring resistance to ampicillin, 83% of those with blocks 1, 7 and 8 from SED-1 confer resistance to high levels of cefotaxime. Those chimeras with blocks 1, 7 and 8 from SED-1 that do not show resistance to cefotaxime have low ampicillin resistance (<250 ug/ml), which indicates they may suffer from marginal stability or poor expression, rather than lack the ability to hydrolyze cefotaxime (Table V-1). In addition, there are four sequences which share this block pattern that did not confer resistance to either ampicillin or cefotaxime. These sequences are likely unfolded or not expressed (Table V-1).

Table V-1. Characterized Chimeras Inheriting Blocks 1, 7, and 8 from SED-1.

Cefotaxime resistant chimeras								CTX	AMP
S	P	T	P	T	S	S	S	0.2	1,000
S	P	T	P	T	S	S	S	0.2	2,000
S	P	T	S	S	T	S	S	0.2	1,000
S	P	T	S	T	S	S	S	2	1,000
S	P	T	S	T	T	S	S	1	2,000
S	P	T	T	S	S	S	S	2	1,000
S	P	T	T	T	S	S	S	0.2	2,000
S	P	T	T	T	T	S	S	0.2	2,000
S	S	S	S	S	S	S	S	>50	2,000
S	S	S	S	T	T	S	S	0.2	1,000
S	T	T	P	T	T	S	S	1	1,000
S	T	T	S	P	P	S	S	5	2,000
S	T	T	S	S	S	S	S	1	2,000
S	T	T	S	T	S	S	S	2	2,000
S	T	T	S	T	T	S	S	1	2,000
S	T	T	T	P	S	S	S	10	2,000
S	T	T	T	S	S	S	S	0.2	2,000
S	T	T	T	T	P	S	S	10	2,000
S	T	T	T	T	S	S	S	5	1,000
S	T	T	T	T	T	S	S	0.2	2,000
Ampicillin resistant, cefotaxime sensitive chimeras									
S	S	T	P	T	T	S	S	<0.1	100
S	S	T	P	T	S	S	S	<0.1	100
S	T	P	T	S	T	S	S	<0.1	250
S	P	T	T	T	P	S	S	<0.1	50
Ampicillin and cefotaxime sensitive chimeras									
S	S	T	P	P	T	S	S	<0.1	<10
S	S	T	S	T	P	S	S	<0.1	<10
S	T	S	P	T	S	S	S	<0.1	<10
S	T	P	T	T	T	S	S	<0.1	<10

Chimera sequences are represented by the parent each block is inherited from P for PSE-4, S for SED-1, and T for TEM-1. MICs for cefotaxime (CTX) and ampicillin (AMP) are given in μg/mL.

Based on various crystal structures of CTX-M type lactamases, the extended-spectrum activity cannot be attributed to active-site widening (Ibuka et al. 1999; Shimamura et al. 2002; Chen et al. 2005). Instead it is credited to several different factors including specific amino acid interactions with the substrate. Asn104 and Ser237 are

residues conserved in CTX-M type lactamases, but not in their narrow-spectrum relatives. Specific interactions between these residues and the carboxylate group and the acylamide side-chain of cefotaxime have been reported (Shimamura et al. 2002). These specific interactions are hypothesized to bind the substrate tightly into the active site. Another factor hypothesized to allow efficient cefotaxime hydrolysis is the position of the ω-loop. The ω-loop has significant affects on substrate specificity when altered in TEM-1 and PSE-4 (Petrosino and Palzkill 1996; Therrien et al. 1998; Sanschagrin et al. 2000). In CTX-M type lactamases there are fewer hydrogen bonds both within the ω-loop (residues 160-181) and between the ω-loop and the third strand of the β-sheet, β3 (residues 229-238). The altered hydrogen bonding pattern results in a change in the position of the ω-loop compared with TEM-1 (Shimamura et al. 2002). However, it also indicates that β3 is less restricted by hydrogen bonds (Ibuka et al. 1999). The third major hypothesis is that movements of β3 allow larger substrates to be accommodated by the active site (Chen et al. 2005). Comparison of anisotropic temperature factors for several CTX-M crystal structures shows that in broader spectrum CTX-M variants there is increased mobility of β3 (Chen et al. 2005). An engineered disulfide to restrict the movement of β3 can reduce the rate of cefotaxime hydrolysis of CTX-M type lactamase TOHO-1 (Shimizu-Ibuka et al. 2004), further supporting the importance of β3 movement for efficient cefotaxime hydrolysis.

Our results suggest that SED-1 blocks 1, 7 and 8 (residues 1-64 and 190-290) contain the necessary components to confer cefotaxime hydrolysis. The ω-loop is composed of blocks 4 and 5, and from this work it appears that they can originate from any of the parents. Additionally, inheriting the ω-loop from SED-1 does not confer

resistance to cefotaxime. The substitutions responsible for disrupting the hydrogen bonds between the ω-loop and β3 occur in both the ω-loop and within β3. Interactions between Asn104 and cefotaxime also do not appear critical for cefotaxime hydrolysis. Asn104 is found in block 3 which is inherited from TEM-1 in most of the chimeras identified.

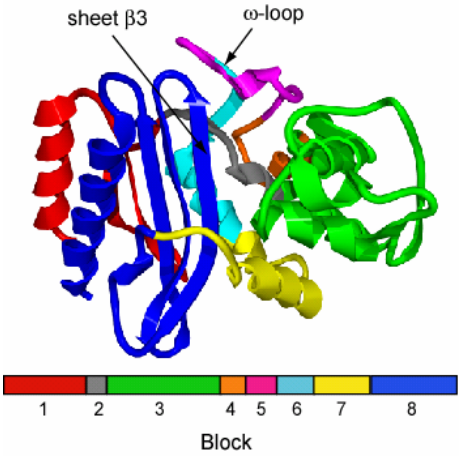

Figure V-2. Structure of TEM-1 with the exchangeable sequence blocks indicated by different colors. The sheet β3 and the ω-loop proposed to be important for cefotaxime resistance are marked.

The remaining hypothesized determinates of CTX-M extended-spectrum substrate specificity, strand β3 including Ser 237, are within block 8. The apparent necessity of inheriting block 7 from SED-1, which is an α-helix that packs against the β-sheet, for cefotaxime hydrolysis is somewhat more surprising. No amino acids within block 7 are near the active site. It is possible that block 7 is constraining the movement of β3 when inherited from TEM-1 or PSE-4. Unfortunately, there is no structural information available for SED-1, so it is difficult to determine what role that particular sequence block is playing. Block 1 may or may not be directly involved with the differences in substrate specificity. As will be discussed shortly, its presence may be necessary for forming a folded protein in conjunction with block 8.

Logistic Regression Analysis of Multiple Sequence Alignment

We previously observed that most functional chimeras cluster into five areas of sequence space (Chapter IV) (Figure V-1). While the sequences cluster on a ternary diagram used to represent the composition of the chimeras, these sequences also cluster based on pairwise sequence identity. To probe whether this clustering is due the sparse sampling of theoretical library, or whether it indicates that some regions of sequence space are unlikely to yield functional lactamases, we examined the entire dataset of 664 functional and nonfunctional chimeras (Appendix III). These data cannot be evaluated by eye like the smaller cefotaxime resistant data set not only because there are many more chimeras, but also because the characterized library is not a random sampling of the theoretical library due to biases introduced during construction. Therefore it is necessary to use an analysis methodology that compares the folded and unfolded chimeras, rather than an analysis method that implicitly assumes an even distribution of possible sequences.

Due to the binary nature of our data (1 for functional, 0 for nonfunctional), we can use logistic regression, an analog of linear regression, to analyze the data. Using logistic regression we fit the folding data to an energy model containing one-body ($\varepsilon_1(i.x)$) and two-body terms ($\varepsilon_2(i.x, j.y)$), that correspond to intra- and inter-block contributions to chimera folding (Equation (V-1)).

$$E = \varepsilon_0 + \sum_i \varepsilon_1(i.x) + \sum_i \sum_{j>i} \varepsilon_2(i.x, j.y) \qquad \text{(V-1)}$$

This method has previously been used to accurately infer interactions from an alignment containing folded and unfolded cytochrome P450s (Otey et al. 2006). The intra-block terms correspond to interactions between the amino acids and the solvent or the main

chain atoms as well as interactions between conserved residues. The inter-block terms correspond to pairwise interactions between the blocks.

Logistic Regression Analysis (LRA) of the β-lactamase data (Appendix III, entire data set including extra positive chimeras), the results of which are shown in Figure V-3, identified five variables as strongly significant (blocks 1, 2, 3, 8, 1-8) (($p << 10^{-6}$)) and three others as marginally significant (5, 1-7, 2-8) ($p \sim= 10^{-4}$) (Figure V-3). When the p-values of blocks 1 and 8 were recalculated relative to a model that includes pair 1-8, their significance diminished considerably (p=0.5 and $4x10^{-3}$ respectively) indicating that the pairwise interaction is the important determinant for folding rather than individual one-body terms. Blocks 2, 3, and the interaction between blocks 1 and 8 remained significant after the second round of p-value testing. The remaining block identities do not seem to have a significant impact on whether a chimera functions.

Inter-Block Interaction between 1 and 8 is Important for Function

The interaction of blocks 1 and 8 is the most significant determinant for retaining functionality (ampicillin resistance) according to the energy model derived by LRA. The diagonal entries corresponding to wild-type interactions are the most favorable (Table V-2), indicating that chimeras inheriting the blocks from the same parent are more likely to function than those inheriting the blocks from different parents. Additionally, chimeras inheriting block 1 from PSE-4 and block 8 from TEM-1 are more likely to function than any other mismatched pairing of the blocks. The importance of the interaction between blocks 1 and 8 was observed in previous experiments where the N- and C-terminal

fragments of the β-lactamase were almost always found from the same parent in functional chimeras (Hiraga and Arnold 2003; Meyer et al. 2003).

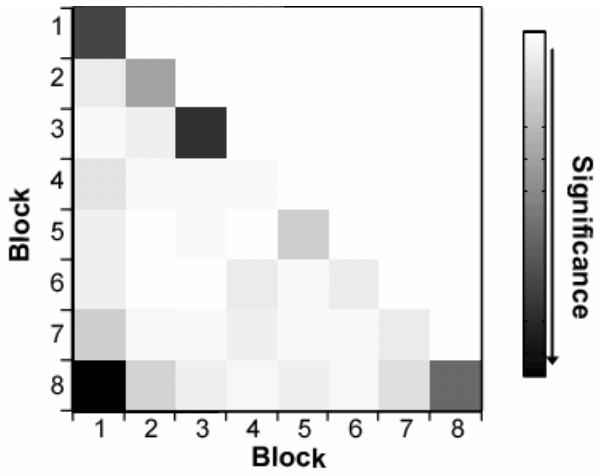

Figure V-3. Logistic regression analysis of functional and nonfunctional chimeras shows that some individual blocks (diagonal), or pairs of blocks are more significant than other for determining whether a chimera will function. The significant terms affecting chimera function are the interaction between blocks 1 and 8, and blocks 2 and 3.

Table V-2. Energies Assigned to Important Interactions by Logistic Regression Analysis

Two-body Terms	Parent at Block 8		
Parent at Block 1	PSE-4	SED-1	TEM-1
PSE-4	-1.2	1.7	-0.5
SED-1	-0.1	-2.8	2.8
TEM-1	1.3	1	-2.3
One-body Terms	Parent		
Block	PSE-4	SED-1	TEM-1
2	-0.5	1.1	-0.5
3	-0.6	1.1	-1.7

An energy value is assigned to each possible parent combination for the pairwise interaction between blocks 1 and 8, and to each parent for blocks 2 and 3. A more negative energy value indicates the block is more likely to be found in functional chimeras.

Blocks 1 and 8 together form almost half the protein; they also have the largest number of structurally contacting residues between two blocks (Table V-3). Block 1 is the most diverse: here the parents share on average only 25% sequence identity. Block 8 is not as diverse as block 1, but SED-1 is significantly more diverged from TEM-1 and PSE-4 than they are to each other in block 8 (Table V-4). The differences between SED-1

and TEM-1 or PSE-4 in block 8 account for most of the increased divergence between them.

Table V-3. Residue-Residue Contacts between Block Pairs and Within Each Block

Block	1	2	3	4	5	6	7	8
1	108	5	0	1	2	19	2	41
2	-	14	15	1	10	8	0	22
3	-	-	258	15	15	2	28	4
4	-	-	-	30	2	9	5	0
5	-	-	-	-	32	15	0	8
6	-	-	-	-	-	32	10	12
7	-	-	-	-	-	-	69	25
8	-	-	-	-	-	-	-	221

Residue-residue contacts between block pairs and within each block (diagonal entries). A residue-residue contact is defined as two amino acids that have any heavy atom, excluding the main chain N and O, within 4.5 Å.

Table V-4. Length and Sequence Identity between Each Pair of Parent Proteins for Each Block

		Sequence Identity		
Block	Length (aa)	PSE-4/TEM-1	TEM-1/SED-1	PSE-4/SED-1
1	40	28%	23%	28%
2	8	75%	63%	50%
3	76	38%	37%	39%
4	11	45%	36%	45%
5	15	60%	80%	60%
6	15	67%	73%	60%
7	27	30%	37%	26%
8	73	51%	30%	32%

The N- and C-termini of the lactamases have diverged significantly so that there are many substitutions in these regions. Analysis of these areas in a multiple sequence alignment shows that they are widely variable: many alignments in fact truncate the N-terminal helix because the sequence identity is nearly undetectable and the start of the mature protein is often uncertain (Bateman et al. 2004). Yet, inheriting residues at the N-

and C-termini from the same parent is almost essential to maintaining a functional protein. Using two different algorithms, Statistical Coupling Analysis (SCA) (Lockless and Ranganathan 1999) and McLachlan Based Substitution Correlation (Gobel et al. 1994; Olmea et al. 1999), to examine the evolutionary covariation of amino acids at the N- and C-terminal helices shows few or no significant interactions (2 or 5 in the top 1% (248) of possible interactions). This is surprising given our results and may indicate that strict covariation is not necessary. The detrimental effect of altering these residues has previously been shown in TEM-1. In a site-saturation study of TEM-1, 18 of the 30 residues which are variable in multiple sequence alignments of class A lactamases and invariable in TEM-1 are within blocks 1 and 8 (Huang et al. 1996). Fourteen of these residues are variable among the three parents studied here. Changing any one of these amino acids could potentially cause the protein to not function correctly. Despite the diversity of sequences at the N- and C-termini of lactamases and the apparent lack of covariation at the individual amino acid level, maintaining the contacts between these two blocks is nearly essential to maintaining a functional lactamase.

Intra-Block Interactions at Blocks 2 and 3 are Important for Function

In addition to the critical interaction of blocks 1 and 8, there are two intra-block variables that are important for determining chimera function. The more significant of these is block 3, where TEM-1 is favored in functional chimeras (Table V-2). TEM-1 at block 3 is found in more of the characterized chimeras than the other two parents due to the biased construction of the library, where 61% inherit this block from TEM-1 and only 17% inherit it from PSE-4 and 21% from SED-1. Because the LGA analysis takes into

account both functional and nonfunctional chimeras in determining the important contributions to chimera folding, this bias only affects whether or not we detect all of the significant variables, not the significance of the variables we do observe.

Block 3 is the largest segment and has the most internal structural contacts, both absolute number and per amino acid (Tables V-3 and V-4). Understanding why block 3 is so strongly favored is difficult due to its large size (76 amino acids) and the small number of functional chimeras (13) that do not have TEM-1 at block 3. There is a disulfide bond within block 3 in TEM-1 (Cys 77 to Cys 123). While this disulfide is not found in SED-1 (the residues are Ala and Ser), it is also present in PSE-4.

The second intra-block term that affects which chimeras are functional is the identity of the parent at block 2. At block 2 SED-1 is disfavored (only 20 of the 143 chimeras with SED-1 at block 2 are functional). In contrast to block 3, block 2 is the smallest segment, with only 8 amino acids, and incorporates at most 4 amino acid changes because the remaining 4 amino acids are conserved in all three parents (Figure V-4). Block 2 contains the active site residues Ser70 and Lys73, and altering any amino acids within it may have a large impact on the activity. SED-1 contains an Ala at position 67 and a Ser at position 72; these positions are Pro and Phe, respectively, in the other parents. In a site-saturation study of TEM-1, Pro67 was found to be invariable, despite the Ala found at this position in multiple sequence alignments of β-lactamases (Huang et al. 1996). In the same study, Phe72 allowed some variation. However, Ser was not one of the identified amino acids. Block 2 is a much more tractable target than block 3 for analyzing the basis for effects of one-body terms on chimera function.

Figure V-4. Sequence alignment of the three parents for block 2 shows only four differences between SED-1 and PSE-4 or TEM-1.

Biophysical Analysis of Block 2

To further investigate the significance of block 2, we examined all 20 functional chimeras that inherited this block from SED-1. The MICs of these chimeras are significantly lower than the MICs of the remaining 162 functional chimeras (Figure V-5). Chimeras with block 2 from SED-1 are not only less likely to confer resistance to ampicillin, but when they do they are impaired.

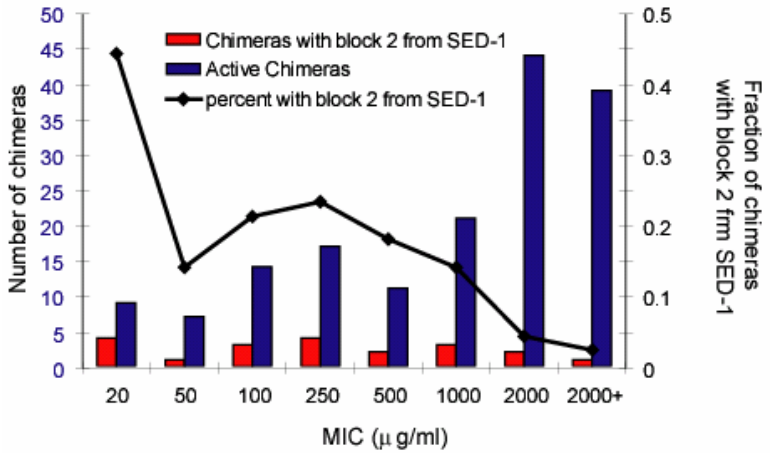

Figure V-5. Characterized functional chimeras with SED-1 at block 2 show less ampicillin resistance compared to the library as a whole.

We also identified sets of characterized functional chimeras that differ by only the parent at block 2 (Table V-5). Examination of all seven sets identified shows that the chimeras with block 2 from SED-1 always have a lower MIC than the same chimera inheriting block 2 from either TEM-1 or PSE-4. In some cases it is not a large difference (only 2-fold), but in many cases the effect is >10-fold (Table V-5).

Table V-5. Characterized Sets of Chimeras Differing Only by Block 2.

Chimera									MIC	T_m (°C)	K_m (μM)	k_{cat} (s^{-1})
1	S	T	T	P	T	T	S	S	4000	52	340 ± 20	450±160
2	S	P	T	P	T	T	S	S	4000	49	16 ± 1.5	280±60
3	S	S	T	P	T	T	S	S	2000	50	25 ± 4	160±39
4	T	T	T	P	T	P	T	T	4000	55	81 ± 1	1400±285
5	T	P	T	P	T	P	T	T	4000	50	12 ± 2.5	14±3
6	T	S	T	P	T	P	T	T	1000	45	25 ± 1.5	40±11
7	T	T	T	T	S	S	T	T	4000	55	168 ± 9	2900 ± 316
8	T	P	T	T	S	S	T	T	2250	49	5.5 ± 1	90 ± 22
9	T	S	T	T	S	S	T	T	1000	48	24 ± 4	60 ± 20
10	T	T	T	P	T	T	T	T	4000			
11	T	P	T	P	T	T	T	T	4000			
12	T	S	T	P	T	T	T	T	212			
13	T	T	T	P	S	T	T	T	4000			
14	T	P	T	P	S	T	T	T	4000			
15	T	S	T	P	S	T	T	T	20			
16	T	T	T	P	P	P	T	T	4000			
17	T	P	T	P	P	P	T	T	4000			
18	T	S	T	P	P	P	T	T	40			
19	T	P	T	T	T	T	T	T	4000			
20	T	S	T	T	T	T	T	T	200			
TEM-1	T	T	T	T	T	T	T	T	4000	55	268 + 49	700 ± 20
SED-1	S	S	S	S	S	S	S	S	4000	55	42 + 4.5	1050 ± 110
PSE-4	P	P	P	P	P	P	P	P	4000	55*		

Chimera sequences are represented by the parents each block is inherited from: P for PSE-4, S for SED-1, and T for TEM-1. Ampicillin MICs (μg/mL) were redetermined to increase fidelity and are not directly comparable with previously reported MICs. For those chimeras for which they were determined the T_m (°C), K_m (μM) and k_{cat} (s^{-1}) have been listed. *For PSE-4 the thermostability is as reported in the literature (Savoie et al. 2000).

The expression level for each of the 21 chimeric proteins was optimized to allow for purification. Analysis of periplasmic extracts shows that chimeras with SED-1 at block 2 have significantly less protein present in the periplasm (Figure V-6) than chimeras with a different parent at block 2. This indicates that a large part of the depressed MIC associated with SED-1 at block 2 may be due to low stability, poor

expression or inadequate transport to the periplasm. While these experiments were performed under high expression conditions, experiments under the screening conditions gave similar results based upon an activity assay performed with cell lysate.

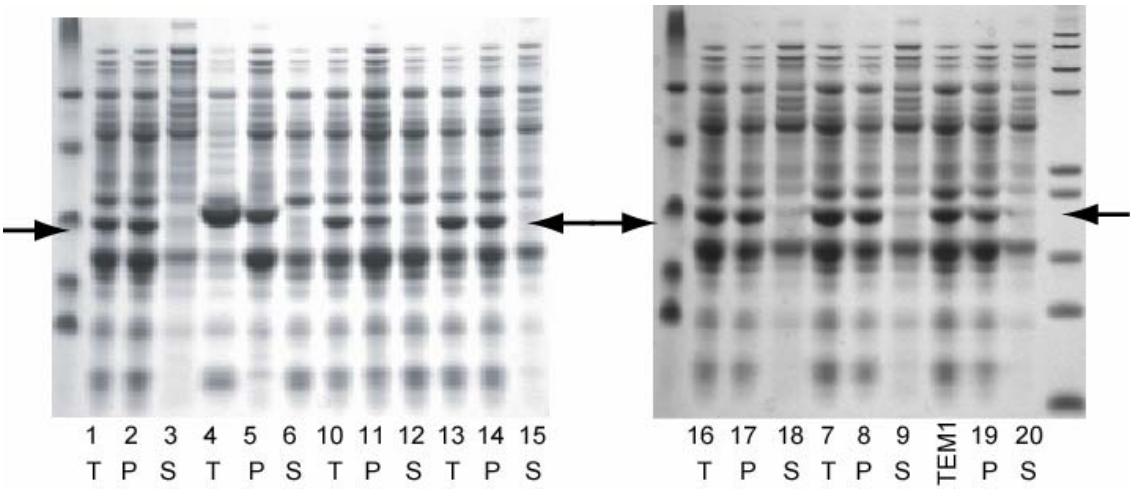

Figure V-6. Periplasmic extracts from chimeras grown under high protein expression conditions. The lactamase is ~30 KD and a band corresponding to that size (marked) is present in all chimeras that do not have SED-1 at block 2. The numbers correspond to chimeras listed on Table V-5, and the letter beneath indicates the identity of the parent at block 2: P for PSE-4, S for SED-1 and T for TEM-1.

The native signal sequence for each parent is included as part of the N-terminal block, and all of the parents are exported correctly to the periplasm. In the past it has been observed that a mutant lactamase can fail to reach the periplasm and become trapped in the cytoplasm when it is partially unfolded (Sideraki et al. 2001). Analysis of cell lysates and periplasmic extracts indicates that there is not a significant difference in activity between the whole cell lysates and the periplasmic extracts under normal expression conditions. However, there may be misfolded or inactive protein present in the cytoplasm. Western blots of whole cell lysates from cells grown under normal expression conditions using an antibody to TEM-1 show a similar pattern to the periplasmic extracts. Only six of the seven sets of chimeras can be examined in this way because the antibody

does not cross react with SED-1 sufficiently to detect chimeras (1, 2, and 3 from Table V-5) where blocks 1 and 8 are from TEM-1.

For three of the sets of chimeras, the members were purified to >95% purity. Kinetic studies were performed with the purified enzymes (Table V-5). The values of K_m and k_{cat} obtained for TEM-1 hydrolysis of ampicillin are consistent values reported in the literature (Schroeder et al. 2002). The chimeras are distinct from one another, but show no clear trend with regard to the identity of block 2. Circular dichroism spectra for TEM-1 and SED-1 and the chimeras were similar to that of PSE-4 (Savoie et al. 2000). T_m apparent was determined for each protein by observing ellipticity at 222 nm during a thermal melt from 1 to 99 °C. The transitions showed cooperative unfolding, and the T_ms agreed well with those determined from activity assays on cell lysates. These studies show that while there is variation in the thermostabilities of the enzymes, none of them are lower than ~45 °C (Table V-5). Thus, these proteins are probably not sufficiently unstable to cause the effect we observe at 37 °C. This is consistent with the fact that lowering the growth temperature to 20 from 37 °C does not have a large impact on the protein expression level. Additionally, adding the well-characterized M182T thermostabilizing mutation (Huang and Palzkill 1997) to four of the chimeras with SED-1 at block 2 (chimeras 3, 12, 15 and 20 from Table V-5) had no effect on the protein expression level, further indicating that the proteins are likely sufficiently stable to be expressed at 37 °C, although they do not accumulate in the cell.

SED-1 is strongly disfavored at position 2, and this effect appears consistent throughout characterized chimeras, even those that are similar to SED-1. The effect is most likely due to decreased expression of these chimeras compared to the parent

proteins. All chimeras with block 2 from SED-1 examined show very low levels of periplasmic protein, and the MICs of the remaining chimeras are consistent with low to no expression. While the expression levels are depressed, the thermostabilities of the purified proteins are not sufficiently low to cause the observed lack of expression. This result ties back to the GFP folding screen conducted in Chapter IV where several chimeras showed ampicillin resistance, but no significant GFP signal. Chimeras that confer resistance to ampicillin, and are stable when purified, do not necessarily accumulate in the cell. It does not take very much of an active lactamase to confer resistance to ampicillin, especially in our low stringency screen. The chimeric proteins which confer resistance but do not accumulate may aggregate or be broken down by the cell for some other reason besides low thermostability.

Despite the effect of block 2 on chimeric proteins, SED-1 itself is well expressed and the codons present in this block are not particularly rare in *E. coli*. It is possible the bias against block 2 has its origins at the mRNA level. However, because the mRNAs are large there is unlikely to be a specific change in RNA folding due to alteration of just four codons. It is also possible that there is a deleterious interaction between block 2 and some other region of the protein that causes the effect. We have not isolated any such interaction with these analyses, but the data are limited. Very few functional chimeras contain block 2, making identification of such an interaction more challenging.

Conclusions

Analysis of chimeric β-lactamases has allowed us to narrow the possible regions of sequence responsible for CTX-M lactamase altered substrate specificity. Nonconserved residues within the ω-loop are not likely contributing to altered substrate specificity because they can be inherited from proteins which do not share this property. Altered substrate specificity is also not likely tied solely to specific amino acid interactions with the substrate because a sequence block distant from the active site is necessary to confer altered substrate specificity.

We have also used the functional and nonfunctional chimeric β-lactamases to inform us about which regions of sequence space might be populated with additional lactamases. All of the clusters observed in Chapter IV (Figure V-1) fall within the areas of sequence space that are compatible with the one-body and two-body terms identified as favorable for producing a functional chimera by LRA. They all have TEM-1 at block 3, and the N- and C-termini originate either from the same parent, or PSE-4 is at the N-terminus and TEM-1 at the C-terminus. However, there are other smaller clusters of functional chimeras that were not originally detected that are likely underrepresented only because more chimeras were not characterized, not because those areas are incompatible with functional chimeras. Examining the critical interactions found in the chimeric β-lactamases shows how the regions of sequence space that functional chimeras populate are limited by specific pairwise interactions. Additionally, portions of sequence that do not appear to interact strongly with other parts of the protein can limit which chimeras function. Why exactly these portions of sequence are so deleterious or

advantageous is still not clear. However, these sequence portions are not necessarily thermostability limiting.

Methods

Cefotaxime Activity Screen

All unique sequences were inoculated from glycerol stocks and used to inoculate 96 deep-well plates which were grown to saturation as for the ampicillin resistance assays in Chapter IV. Aliquots were transferred onto LB agar plates containing various concentrations of cefotaxime (0.05 to 1 μg/mL) similarly to the ampicillin assay previously described (Chapter IV).

Logistic Regression Analysis

Logistic regression assumes that the probability of a chimera folding decreases with energy E according to the sigmoidal relationship

$$Pf = \frac{1}{1+e^{E}}, \tag{V-2}$$

We defined E as the sum of one- and two-body terms in Equation (VI-1). The significance of each term was calculated relative to a reference model that included only the one-body terms using the maximum likelihood test (Endelman et al. 2005). The individual one-body terms were removed from the model and the increase in deviance D measured.

$$D = -2\sum_{i=N_f+1}^{N} E_i + 2\sum_{i=1}^{N} \ln(1+e^{E}) \tag{V-3}$$

The magnitude of this increase follows the chi-square distribution with two degrees of freedom, which was used to calculate a p-value for each one-body term. The significance of two-body terms was determined by recording the decrease in the deviance and the p-value determined from the chi-square distribution with four degrees of freedom. The algorithm MINOS through the NEOS server was used for optimization. The GAMS input file necessary for this computation can be found in Appendix I.

Sequence Analysis

Evolutionary covariation between amino acids was examined using both Statistical Coupling Analysis and McLachlan Based Substitution Correlation. Java code for both of these algorithms was downloaded from http://www.afodor.net/ (Fodor and Aldrich 2004b, 2004a), and the full PFAM lactamase superfamily alignment used for calculation (Bateman et al. 2004).

Protein Purification

With the exception of A1A3, A1H6, and A2A4, proteins were expressed in 2x 1L cultures of TB +35 μg/mL Chl grown to saturation at 37 °C, 250 rpm. The remaining proteins were expressed in 6 x 500 ml TB cultures with the addition of 20 ng/mL anhydrotetracycline (inducer) to maintain a high expression level and grown 48 hours at 25 °C. The cells were pelleted (8000 xg, 8 min, 4C) and the periplasmic proteins isolated through osmotic shock. The cells were resuspended gently in 200 mL 30% sucrose, 30 mM Tris pH 8.0, 1 mM EDTA and allowed to incubate at room temperature for 10 minutes before repeating centrifugation as above. The supernatant was removed

completely and the cells were resuspended in 200 mL ice cold water with shaking and vortexing. Following a 10 minute incubation on ice the cells were centrifuged to pellet the cells (30 min, 15,000g, 4 °C). The supernatant was removed as the periplasmic extract and either dialyzed overnight at 4 °C against 20 mM Tris pH 8.0, or 1 M Tris pH 8.0 was added to a final concentration of 20 mM. SED-1 was dialyzed overnight to 20mM HEPES pH 7.0.

The buffered periplasmic extract was applied to a Q FF HP (Amersham) column and washed with 20 column volumes of 20 mM Tris pH 8.0. The protein was eluted in a gradient of 0-200 mM NaCL over 12 column volumes. Fractions were tested for lactamase activity using nitrocefin and purity assessed through SDS gel electrophoresis. The purest fractions were collected and concentrated using a Millipore Centriprep (YM10) to 0.7 mL and then applied to an S-100 gel filtration column (Amersham) run at 0.15 mL/min in 30 mM Tris pH 8.0. Fractions were tested for activity as above and purity verified through SDS gel electrophoresis.

SED-1 has an isoelectric point around 8 and therefore was purified using cation exchange chromatograhpy. SED-1 periplasmic extract was buffered in 20 mM HEPES pH 7.0, applied to 3 SP FF (Amersham) columns in series and washed with 20 column volumes of 20 mM HEPES pH 7.0. The protein was eluted in a gradient from 0 to 200 mM NaCl over 12 column volumes. The fractions were assayed for activity using nitrocefin and the purity of active was verified by gel electrophoresis. No gel filtration was necessary to obtain >95% purity.

MIC Determination

The MICs for Figure V-5 were determined in Chapter IV during the high-throughput screening. However, MIC determination for sets of chimeras in Table V-4 was repeated to ensure higher fidelity. These MIC's were determined through liquid culture dilutions rather than the plate assay. 500 uL cultures of LB with 35 ug/mL chloramphenicol were grown in deep-well 96-well plates at 37 °C for 18 hours and then diluted 1:1000. 10 uL was used to inoculate 96 well culture plates containing 90 uL of LB with varying concentrations of ampicillin (0, 2.5, 5, 10, 25, 50, 100, 250, 500, 1000, 2500, 5000, 10000 μg/mL). The cultures were grown for 18 hours at 37 °C and the OD_{590} measured. Cultures with an $OD_{590} > 0.1$ were considered grown.

Site-Directed Mutagenesis

The M182T mutation was introduced into chimeras 3, 12, 15 and 20 using quick-change mutagenesis with the following primer and its reverse complement: 5'-CGT GAC ACC ACG ACC CCT GTA GCA ATG G. The altered codon is underlined. The genes were sequenced in both directions to verify correct incorporation of the mutation and no additional mutations.

CD Spectroscopy

Purified proteins were diluted to 30 μM in KPO_4 buffer pH 7.0. Protein concentrations were determined by measuring absorbance at 280 nm. To verify the presence of a folded lactamase, circular dichroism wavelength scans from 200 to 250 nm at 1 nm increments with a 5 second averaging time were performed on a JASCO model J-

600 spectropolarimeter. To determine the apparent T_m, ellipticity was monitored at 222 nm during a thermal denaturation from 0 to 99 °C. The step size was 1 °C, the equilibration time 2 minutes, and the signal averaging time 30 seconds.

Catalytic Activity Assays

Enzyme kinetics were measured at 25 °C in 100 mM KPO_4 buffer pH 7.0. Degradation of ampicillin was measured by UV-Vis at 232 nm, ε_{232} for ampicillin is 912 $cm^{-1}M^{-1}$. Ampicillin concentrations between 10 and 500 μM were tested and protein concentrations ranged between 0.25 and 12 μM as appropriate to record a linear initial rate. Rate constants were fit using a Hanes-Woolf plot ([S]/v vs. [S]).

Rates were measured to compare whole cell lysates and periplasmic extracts using the chromagenic substrate nitrocefin at 25 °C in 100 mM KPO_4 buffer pH 7.0, 50 mg/mL of nitrocefin. Nitrocefin degradation to form a red product (ε_{486}=20,500) was measured by observing at 468 nm. A similar assay was performed at varying temperatures to estimate the T_m for comparison with CD measurements.