Protein Science (1998), 7:2200–2209. Cambridge University Press. Printed in the USA.

Delineation of an evolutionary salvage pathway by compensatory mutations of a defective lysozyme

MILAN JUCOVIC AND ANTHONY R. POTEETE

Department of Molecular Genetics and Microbiology, University of Massachusetts Medical Center, 55 Lake Avenue North, Worcester, Massachusetts 01655

(RECEIVED May 21, 1998; ACCEPTED June 25, 1998)

Abstract

Model-free approaches (random mutagenesis, DNA shuffling) in combination with more "rational," three-dimensional information-guided randomization have been used for directed evolution of lysozyme activity in a defective T4 lysozyme mutant. A specialized lysozyme cloning vector phage, derived from phage λ, depends upon T4 lysozyme function for its ability to form plaques. The substitution W138P in T4 lysozyme totally abolishes its plaque-forming ability. Compensating mutations in W138P T4 lysozyme after sequential random mutagenesis of the whole gene as well as after targeted randomization of residues in the vicinity of Trp138 were selected. In a second stage, these mutations were randomly recombined by the recombinatorial PCR method of DNA shuffling. Shuffled and selected W138P T4 lysozyme variants provide the hybrid λ phage with sufficient lysozyme activity to produce normal-size plaques, even at elevated temperature (42 °C). The individual mutations with the highest compensatory information for W138P repair are the substitutions A146F and A146M, selected after targeted randomization of three residues in the neighborhood of Trp138 by combinatorial mutagenesis. The best evolved W138P T4 lysozymes, however, accumulated mutations originating from both randomly mutagenized as well as target-randomized variants.

Keywords: bacteriophage T4 lysozyme; directed evolution; DNA shuffling; second-site reversion

The lysozyme of bacteriophage T4 is especially suitable for exploring the relationship between protein structure and function, as well as between protein sequence and three-dimensional conformation. T4 lysozyme is a small, globular monomeric protein consisting of 164 amino acid residues, which can be reversibly unfolded in vitro. Its enzymatic activity facilitates lysis of T4 phage-infected cells, thereby releasing replicated phage particles (Tsugita et al., 1968). The crystal structure of T4 lysozyme has been solved at high resolution (Remington et al., 1978; Weaver & Matthews, 1987). Many mutant variants of T4 lysozyme have been found to form isomorphous crystals of high quality, and thus, it has been possible to closely examine the structural effects of numerous amino acid substitutions in this protein. Much of what we know about the structure of mutant proteins is based on the extensive studies of more than 200 T4 lysozyme mutants by Matthews (1995).

T4 lysozyme has additionally provided an informative experimental system in which to explore questions of protein evolution, both natural and directed. It is a member of a group of bacteriophage-encoded lysozymes that are related by sequence homology. For example, systematic substitution of amino acid residues in T4 lysozyme permitted an experimental assessment of the relative functional significance of conserved and nonconserved residues (Poteete et al., 1992). A number of mutant variants of T4 lysozyme displaying increased stability have been described as well (reviewed by Matthews, 1995).

Members of protein families are thought to have evolved from a common ancestor by a process of divergence. Divergent evolution is envisioned as a gradual accumulation of neutral mutations, or of mutations that subtly improve the function of a protein in the particular organism or niche in which it comes to reside. Such a process tends to preserve critical parts of the protein: conserved amino acid residues, whose loss or alteration most frequently results in loss of protein function due to either decreased stability of the folded structure, or to direct interference with a catalytic function.

The foregoing considerations raise the question of what would happen if evolution were constrained to go through a pathway that involved the loss of function due to substitution of a structurally critical amino acid residue, and in which simple reversion to regain the original critical residue was blocked. In our previous studies, we have screened second-site revertants of deleterious single amino acid substitutions in T4 lysozyme with the aim of finding variants with significantly altered structures. In a number of cases, the revertants could be classified according to their effects: increased expression, greater global stability, surface charge changes, or spe-

Reprint requests to: A.R. Poteete, Department of Molecular Genetics and Microbiology, University of Massachusetts Medical Center, 55 Lake Avenue North, Worcester, Massachusetts 01655; e-mail: tpoteete@ummed.edu.

cific structural alterations (Knight et al., 1987; Poteete et al., 1991, 1997; Bouvier & Poteete, 1996). In all of these cases, only the first step of evolutionary salvage was detected. Generally, we could not find second-site revertants of the primary mutants with the worst structural defects. The present study employs new methods to track multistep restoration of function to a severely damaged lysozyme variant.

Results and discussion

Experimental system

We have previously described the specialized lysozyme cloning vector phage λ cI857 Sam7 R::Tc736 (Bouvier & Poteete, 1996). The *R* gene of phage lambda, which encodes transglycosylase, a lysozyme-like function required for host cell lysis, is replaced by the tetracycline resistance-conferring *tetA* gene from pBR322. As a consequence of this genetic change, the phage is defective in forming plaques on a normal host unless it acquires an active lysozyme gene. The restriction enzyme SfiI cuts phage DNA on either side of the *tetA* substitution, generating three fragments: left and right arms, and a fragment containing the substitution sequences. The mixture is ligated with a population of Sfi-ended DNA fragments representing a library of mutagenized variants of a defective T4 lysozyme gene, and packaged into phage lambda capsids in vitro. Only those particles containing chromosomes with functional lysozyme genes can form plaques.

Our previous use of this system involved the following steps: (1) introducing a primary (inactivating) mutation into T4 lysozyme gene, (2) random mutagenesis of T4 lysozyme gene by error-prone PCR, and (3) selection of second-site revertants by visual screening of plaque size. This selection system proved to be very sensitive for detection of low residual activity of lysozyme mutants. Only a portion of mutations labeled as "deleterious" in P22 phage (Rennell et al., 1991) can keep this classification after testing in the lambda phage system (M. Jucovic, C. Lewis, & A.R. Poteete, unpubl. obs.). Evidently, either λ-mediated lysis of *Escherichia coli* requires less lysozyme than P22-mediated lysis of *Salmonella typhimurium*, or else the lysozyme gene is expressed at a higher levels by the λ late promoter than by its P22 counterpart. These observations suggest that 3% of wild-type activity is an upper limit of the sensitivity of the lambda plaque-forming assay. A lower limit is suggested by observations reported below. The specific activity of the W138P mutant lysozyme, deficient in plaque-forming ability in the lambda phage, is 0.3% of that of wild-type. Thus, the edge of detectable lysozyme activity in the lambda phage system should be somewhere within this range (>0.3% and <3% of wild-type lysozyme activity).

Because the system is more sensitive for detection of residual lysozyme activity, we reasoned we might be able to start evolution from a severely damaged protein.

The main limitation of this biological system is the limited range of selective pressure that can be applied. It is likely that a severely damaged protein will require accumulation of several beneficial mutations to cross the selection barrier. The selection barrier, a step from plaque-forming inability to ability, can be adjusted by incubation temperature and/or concentration of top agar. Lysozymes with enhanced activity or thermal stability can be selected at higher temperature (42 °C) and with the usual concentration (0.9%) of top agar, while less active and/or stable variants can be selected at 37 °C with a lower concentration of the top agar. Because the growth of *E. coli* is slowed or stopped by temperatures above 42 °C, we cannot increase the selection pressure for compensating mutations by increasing temperature above this temperature. Under these circumstances, only a visual screen for plaque size or lysoplate assay for halo size formation (Streisinger et al., 1961) can be used for distinguishing further evolved progeny from their parents. This, however, requires plating the phage library at a reasonably low density and thus restricts the size of the screenable mutant lysozyme library.

Rationale for fixing T4 lysozyme pathology

An unbiased genetic approach has been shown to be often more informative about relationships between amino acid sequence and protein structure than a model-based approach; has often yielded unexpected insight into the enzyme's catalytic machinery (Kim et al., 1996; Warren & Benkovic, 1997); and significantly contributed to engineering proteins with novel properties (Shao & Arnold, 1996; Zhang et al., 1997). Although this random approach often makes the most sense, it can access only a small fraction of total protein sequence space due to the limited transformation efficiency of *E. coli* or the limited packaging efficiency of lambda phage. Full randomization of amino acid sequence is, therefore, limited in the best case to less than eight residues (Kast & Hilvert, 1997). As a result, any structural information about the protein will help to target sequence randomization experiments.

A model-free random mutagenesis in combination with a more rational approach, guided by three-dimensional structural information has been used in this work for evolution of an inactive T4 lysozyme variant into a functional protein (Fig. 1). We selected compensating mutations for a deleterious mutation in T4 lysozyme by sequential random mutagenesis as well as by randomization of residues close to the site of primary, inactivating mutation. Compensatory mutations, selected from these independent experiments, were recombined by the DNA shuffling method (Stemmer, 1994) in an attempt to select the best variants.

Position 138 is a candidate for deleterious substitution

Trp138 resides on a very short α-helix (residues 137–141) and apparently anchors two larger helices (residues 126–134 and 143–155; Fig. 2). Trp138 is almost completely buried in the wall of the active site cleft, whereas other residues of this helix Arg137, Tyr139, Asn140, and Gln141 are exposed to solvent. Elwell and Schellman (1979) have carried out a detailed thermodynamic study of the stability of W138Y mutant. They found that Trp138 has a crucial role in maintaining T4 lysozyme in a stable, active form. The W138Y replacement significantly decreased the thermal stability of the enzyme. Temperature-sensitive mutations in this position, W138Y and W138G, have also been reported by Alber et al. (1987). According to conservation scores and mutational studies, Trp138 is among 10 most critical residues in T4 lysozyme (Poteete et al., 1992). In the hybrid P22 phage system, only Tyr, Phe, and Leu are tolerated at this position (Rennell et al., 1991). In the course of this study, we have found (see below) that other substitutions in this position (W138A, W138G, W138T, and W138S) still produce sufficient lysozyme activity for the λ lysozyme cloning vector to form a plaque. Mutation W138P, however, totally abolished plaque formation. Mutant W138P thus appeared to be a good candidate for directed evolution of lysozyme variants.

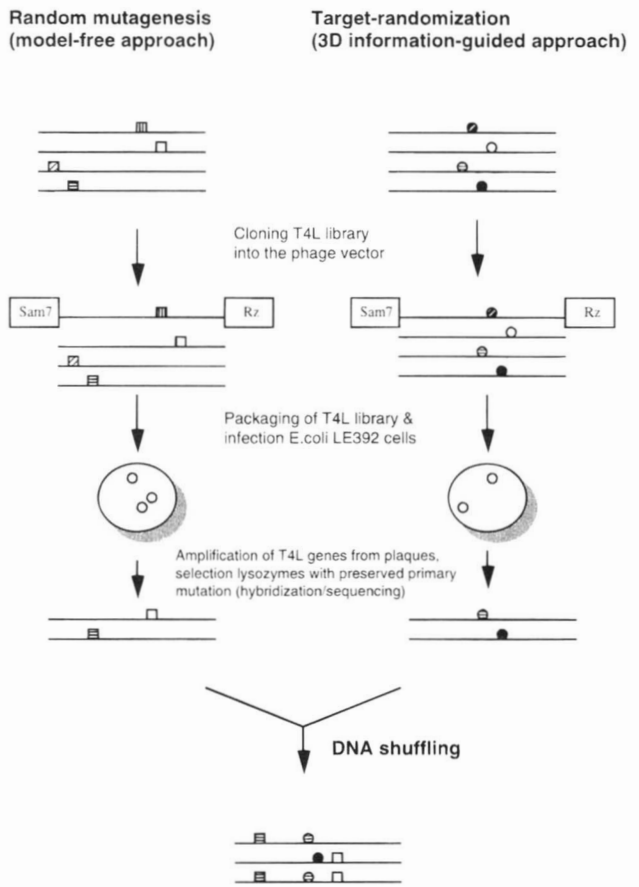

Fig. 1. Strategy for evolution of defective T4 lysozyme mutants. Suppressor mutations for the deleterious substitution W138P have been selected either in a model-free manner by random mutagenesis or by targeted randomization of the neighborhood of the primary-site mutation. Selected mutations have been randomly recombined by DNA shuffling. In this PCR-like reaction, recombination inside the population of mutagenized gene is promoted by cleaving the initial population of molecules into random fragments by DNase I (Stemmer, 1994).

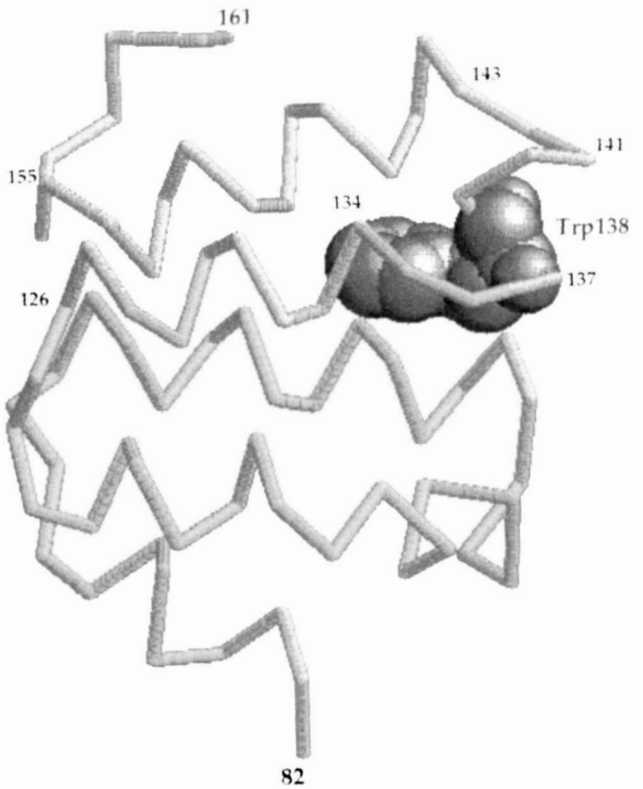

Fig. 2. Trp138 (space filled) and part of polypeptide backbone (residues 82–161) of T4 lysozyme. Trp138 resides on a very short helix (residues 137–141) and anchors two larger helices (residues 126–134 and 143–155). The side chain of Trp138 is almost completely buried in the hydrophobic core of the carboxy-terminal lobe.

Random mutagenesis of lysozyme W138P: First generation of suppressor mutations

The W138P lysozyme gene was subjected to error-prone PCR under mutagenic conditions where the concentration of one nucleotide (dGTP) was limiting and Mg^{2+} was replaced by Mn^{2+}. After subcloning of the PCR product into lambda phage, packaging, and plating, potential revertants were selected at 37 °C. Sequencing of lysozymes from a few candidates indicated that most of these plaque-forming phages were T4 lysozyme primary-site pseudorevertants. The following reversions were found at position 138: P138A, P138G, P138S, and P138T.

To find second-site revertants in the background of primary-site pseudorevertants, lysozyme DNAs were individually amplified from restreaked plaques, dot blotted, and hybridized with an oligonucleotide probe carrying the primary W138P mutation. Only plaques that hybridized to the probe—presumptively retaining the primary mutation—were sequenced. This additional hybridization step significantly reduced sequencing of unwanted background, although it is possible (but unlikely) that we sacrificed putative second-site revertants positioned very close to the primary site mutation, within the oligonucleotide probe sequence. Only 5% of all selected revertants hybridized with the primary-site mutagenic oligonucleotide, suggesting that more than 90% of all putative revertants are primary-site pseudorevertants. Four second-site revertants were found: A93E, T151A, A41V, and the double mutation T151A/L39I, all in combination with the primary W138P mutation.

A93E was selected under stringent conditions (37 °C, 0.9% top agar), while the two other suppressors were selected at 37 °C with 0.45% top agar. Because the plaque size is influenced by, among other factors, the concentration of agar (Sambrook et al., 1989), it is possible that lower concentrations of top agar permit phages with less healthy lysozyme genes to form plaques. Once selected, however, all of these revertants were able to form tiny plaques at 37 °C with 0.9% top agar.

Sequential random mutagenesis: Second generation of suppressor mutations

The best second-site revertant from the first generation, W138P/A93E, was subjected to another round of random mutagenesis. None of the first-generation revertants was able to form plaques at 42 °C, so the second generation of revertants with improved lysozyme function were selected at this temperature. The following sets of evolved lysozymes were obtained at 42 °C: A93E+E108V, A93E+T151A, A93E+E108V+T109A (all in combination with W138P). All these phages form more healthy plaques at 37 °C than

the first generation of revertants, and tiny plaques at 42 °C. New mutations (E108V, T109A), as well as a combination of mutations from the first generation (A93E+T151A), were selected in the second generation at 42 °C (Fig. 3). Mutations identified by sequencing T4 lysozyme genes encoding the selected first and second generation variants are summarized in Table 1.

Selected compensatory mutations are global suppressors

All mutations selected in a model-free manner are remote from the site of primary mutation (Fig. 4), and most of them have been already found to suppress mutations at more than one primary site (Poteete et al., 1997). The A93E substitution, for example, improves the function of lysozymes bearing nine different substitutions at position 156. Ala93 is on the surface of the T4 lysozyme and its substitution by Glu will alter the surface charge of the protein. T4 lysozyme has a strong positive charge, and its diminishing by the A93E substitution may increase mobility of lysozyme among the negatively charged bacteria in the plaque (Dao-pin et al., 1991). The T151A substitution effectively suppress different mutations at position 129, and E108V suppresses two substitutions at position 99 (Poteete et al., 1997). Mutation L39I has been selected in combination with T151A, as well as mutation T109A in combination with A93E+E108V. We have not tested these mutations separately. Mutation A41V has been identified by in vivo screening as a substitution, which enhances the thermostability of T4 lysozyme (Pjura et al., 1993). This substitution in combination with six other stabilizing point mutations increased the protein's melting temperature by up to 8.3 °C at pH 5.4 (Zhang et al., 1995).

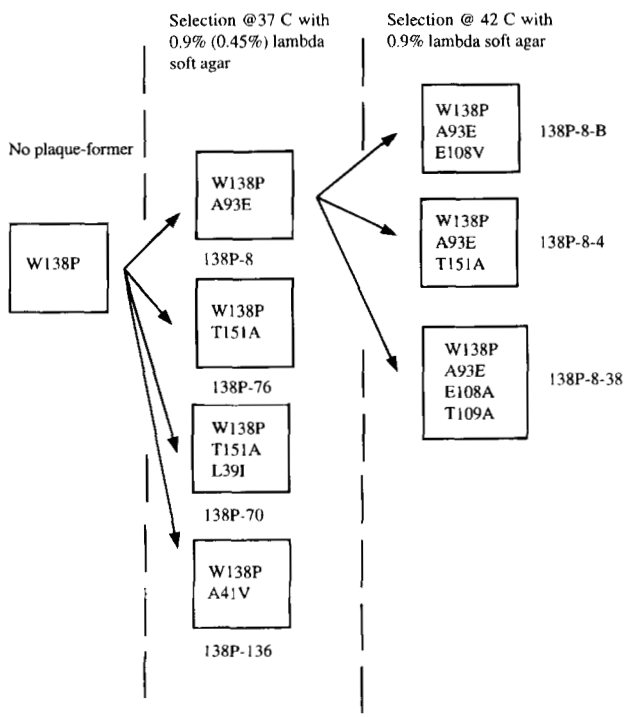

Fig. 3. Evolution of W138P lysozymes by sequential random mutagenesis. The best first-generation mutant selected at 37 °C, W138P/A93E, was subjected to another round of random mutagenesis and improved progeny variants were selected at 42 °C.

Table 1. *Substitutions selected after random mutagenesis*[a]

Phage	Wild-type codon	Mutant codon	Amino acid substitution	Generation
138P[b]				0
138P-8	GCG	GAG	A93E	1
138P-76	ACA	GCA	T151A	1
138P-136	GCT	GTT	A41V	1
138P-70	ACA	GCA	T151A	1
	CTT	ATT	L39I	
138P-8-4	GCG	GAG	A93E	2
	ACA	GCA	T151A	
138P-8-38	GCG	GAG	A93E	2
	GAA	GTA	E108V	
	ACC	GCC	T109A	
138P-8-B	GCG	GAG	A93E	2
	GAA	GTA	E108V	

[a]DNA and amino acid substitutions resulting from sequential random mutagenesis and selection for plaque-forming ability at 37 °C (generation 1) and 42 °C (generation 2).

[b]138P is the lambda lysozyme cloning vector phage bearing the W138P-substituted T4 lysozyme gene. All other phages have additional mutation(s) in the T4 lysozyme gene.

Structural information as a guide to randomization experiments

Certain amino acid residues within 5.0 Å of W138 (Fig. 5) were selected for randomizing mutagenesis, either individually or in combinations. Positions mutagenized were 105, 133, 105+133, 102+106, and 146+149+150.

Simultaneous randomization of positions 146+149+150 yielded compensatory mutations of W138P (Table 2). Several independent phage isolates with the secondary mutation A146F or A146M in the W138P T4 lysozyme background have been selected; one isolate possessed an additional mutation V149I, the other one mutation K147N. Because position 147 has not been randomized, mutation K147N selected in the background of A146F mutation must have been introduced in the second-stage PCR reaction amplifying the spliced, primary PCR products. As the side chain of K147 is almost completely solvent exposed, the substitution Lys → Asn in this position probably plays a negligible role in W138P suppression. Because no substitutions in position 150, and only one substitution in position 149 (V149I), have been found after randomization of three residues (146, 149, and 150), wild-type residues in these positions in combination with A146F or A146M mutation presumably have high information content in a final compensatory performance. All of these T4 lysozyme mutants provide the lambda phage with sufficient lysozyme activity to form normal-size plaques at 37 °C and small plaques at 42 °C.

How can A146F and A146M exert their suppressive effect for the detrimental W138P substitution? Knowledge of the T4 lysozyme structure and extensive structural studies of many mutants from Matthews' laboratory permit us to rationalize the effects of selected suppressing substitutions. Generally, a proline residue cannot fit in to an α-helix without distorting it as well as destabilizing it due to its lack of a backbone amide for hydrogen bonding. The inability to adopt many of the backbone conformations available to other residues makes Pro the second most frequently unacceptable residue in T4 lysozyme (Rennell et al., 1991). Wild-type Trp138

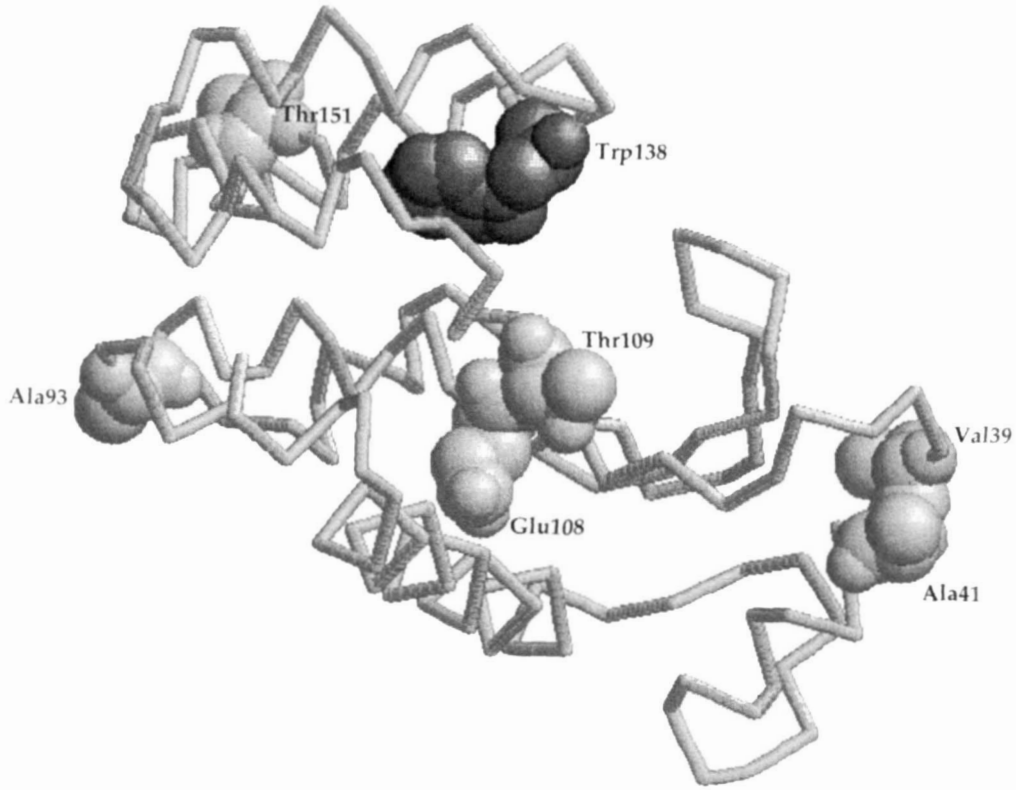

Fig. 4. Remote positions for suppression of W138P mutation. Specific substitutions in positions 41, 93, and 151 provide W138P T4 lysozymes with plaque-forming ability at 37 °C. Substitutions in positions 108 and 108+109 further improve the function of the double T4 lysozyme mutant (W138P/A93E). Also, the combination of the first generation mutations A93E and T151A results in a more active T4 lysozyme, able to form tiny plaques at 42 °C.

resides on a very short helix (137–141) that forms a bridge between two larger helices (126–133 and 143–155; Fig. 2). Its bulky side chain is almost completely buried inside the carboxy-terminal lobe. Ala 146 is within 5 Å of Trp138, and its methyl group is packed against the indole ring of Trp138 (Fig. 6). Replacing of Trp138 by Pro apparently deforms the small helix and forms a cavity that can be partially filled by side chains of phenylalanine or methionine replacing Ala146 in our evolved mutants. This interpretation is consistent with recent thermodynamic and structural studies of compensatory changes in "size-switch" L121A/A129M core repacking variants of T4 lysozyme (Baldwin et al., 1996).

Directed evolution by random recombination

Eight improved W138P mutant variants (Fig. 7), selected either after random mutagenesis or targeted randomization, were chosen as the starting material for further recombinatorial directed evolution by DNA shuffling. This method has proved to be very useful for recombination of homologous DNA sequences during in vitro molecular evolution (Stemmer, 1994; Crameri et al., 1996, 1997, 1998; Moore et al., 1997; Zhang et al., 1997; Zhao & Arnold, 1997a, 1997b). A mixture of equimolar amounts of W138P-bearing revertant T4 lysozyme genes amplified directly from selected phage plaques was shuffled by random fragmentation with DNaseI and reassembled in primerless PCR. After dilution, the reassembled product was amplified in regular PCR with primers. The recombined library was subcloned into the λ phage cloning vector, packaged, and plated. Because some of the parental mutant lysozymes selected after sequential random mutagenesis and target randomization are able to form small plaques at 42 °C, we used a plaque-size visual screen to identify phages with improved lysozyme function at this temperature. Phages that retained the primary, inactivating mutation in T4 lysozyme gene were selected from a population of large plaque-formers by hybridization with the primary mutagenic oligonucleotide.

Six mutants with improved lysozyme function have been selected after DNA shuffling (Fig. 7). All six are progeny of two parental genes; in each case one parental gene was previously selected in a model-free manner after random mutagenesis, the other one after targeted randomization.

It appears that the most important information for fixing the W138P defect originates from the cavity filling "small-to-large" A146F or A146M substitutions, respectively. Surface mutations, however, apparently contribute to salvaging W138P in a nonspecific manner. From a statistical point of view, surface mutations (L39I, A41V, T109A, and T151A) can be expected to be incorporated into a progeny sequence with a probability of 1/8 (one out of eight shuffled sequences has one of these substitutions); and mutations (A93E and E108V) with a probability of 1/4 (two out of eight shuffled sequences carry these substitutions). The probability of generating the wild-type residue at positions 39, 41, 109, and 151 is thus 7/8; at mutation sites 93 and 108 wild-type residues are

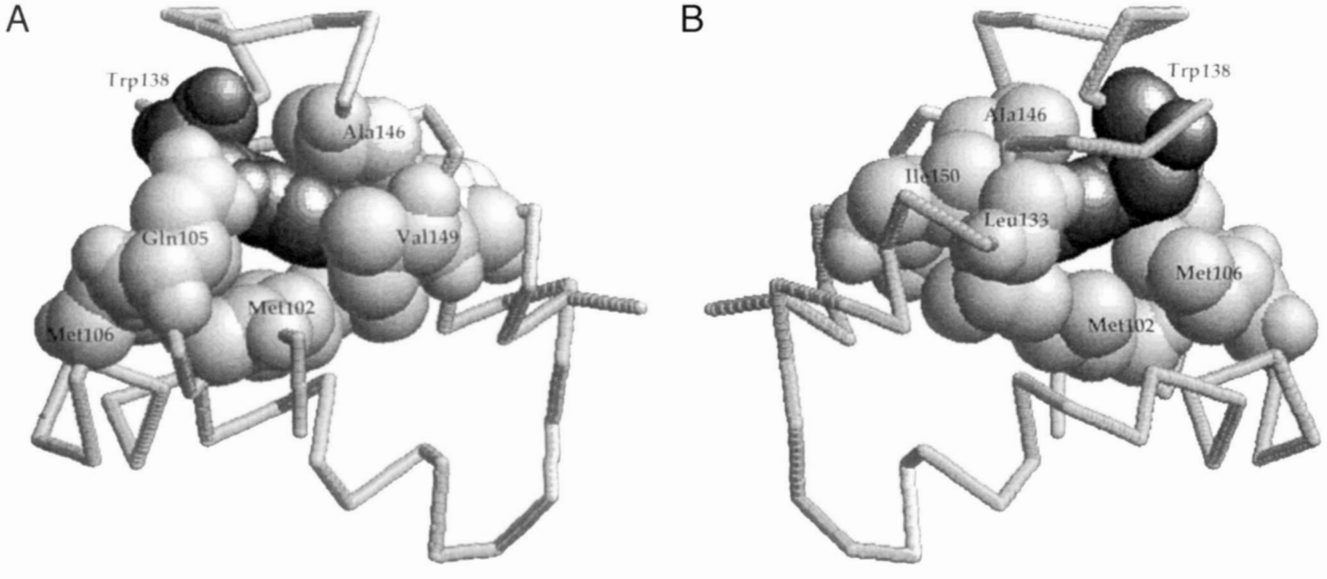

Fig. 5. (**A**) Front and (**B**) back presentation of Trp138 and its environment. The following positions have been randomized individually or in combinations: 105, 133, 105+133, 102+106, and 146+149+150. Simultaneous randomization of three positions (146, 149, 150) yielded compensatory mutations of W138P.

incorporated into shuffled progeny with probability of 3/4. Although there is a presumptive statistical preference for the presence of wild-type residues in the progeny, all variants selected after DNA shuffling adopted at least one of these surface, global-stabilizing substitutions. One of the evolved lysozymes acquired a new K48Q mutation after DNA shuffling. The absence of L39I and T151A mutations in the shuffled progeny is not surprising. Mutation T151A appeared in the first generation of suppressor mutations after random mutagenesis. In another phage isolate from this lineage (phage 138P-70; see Fig. 3), the T151A mutation is coupled with L39I. It is, therefore, likely that L39I is a silent substitution having neither a beneficial nor a damaging effect in W138P T4 lysozyme, and thus has not been selected after DNA shuffling. Appearance of the T151A suppressor mutation in a shuffled generation is possible, but unlikely. Substitutions A146F and A146M have clearly higher compensatory potential for the deleterious W138P substitution than T151A (A146F and A146M provide T4 lysozyme with small plaque-forming ability at 42 °C, while T151A and all selected first-generation randomly mutagenized variants do not); and because these mutation sites (146 and 151) are only 12 nucleotides apart, their recombination in a shuffling reaction is rather limited.

The ability of evolved W138P T4 lysozymes to lyse bacterial cells was examined by the use of a plasmid-based expression system. W138P T4 lysozyme gene and its evolved variants were subcloned into pDR739 (Poteete et al., 1997) in place of the wild-type gene. *E. coli* W3110 F′ *lacI*^{Q} (pDR739) cells infected by lysis-defective lambda phage (Ram5) provide sufficient "leaky" expression of T4 lysozyme from the $P_{A1/04/03}$ promoter (Lanzer & Bujard, 1988) to form large, healthy plaques at 42 °C. The primary T4 lysozyme mutant (W138P) is not able to support cell lysis under this condition, while all shuffled W138P T4 lysozyme vari-

Table 2. *Revertants selected after targeted randomization*[a]

	145	146	147	148	149	150	151
Wild-type	R	A	K	R	V	I	T
A146F	—	F	—	—	—	—	—
A146+K147N	—	F	N	—	—	—	—
A146M	—	M	—	—	—	—	—
A146M+V149I	—	M	—	—	I	—	—

[a]Amino acid residues 145–151 of wild-type T4 lysozyme and suppressor mutations for W138P mutant selected after target randomization. Dashes indicate amino acid identity with wild-type residues.

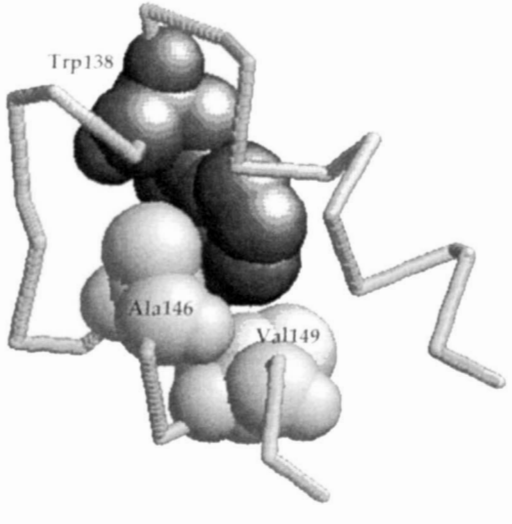

Fig. 6. Trp138 and positions 146, 149, where compensatory mutations of W138P have been found.

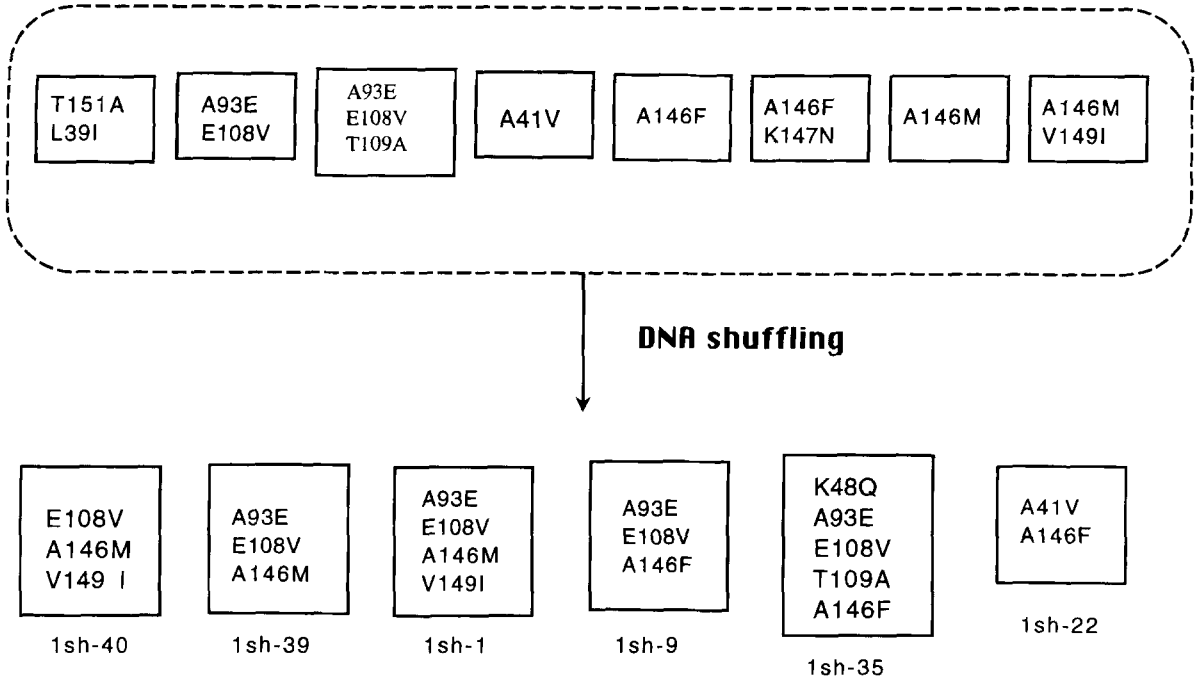

Fig. 7. Recombination of beneficial mutations by DNA shuffling.

ants produce enough activity to form medium-size plaques at 42 °C after λ Ram5 infection. Because structural lysozyme genes were subcloned into the expression vector without any upstream sequences from the phages in which they were selected, there is no doubt that selected W138P variant lysozymes are more functional than their primary mutant ancestor, and that the better performance of the evolved proteins is not caused by higher expression of the lysozyme gene due to possible extragenic mutations.

Improvement of enzyme function: Determination of relative activity of evolved lysozyme in vitro

The specific activity of evolved lysozyme variants is higher than the specific activity of their ancestor, W138P T4 lysozyme. The evolved W138P variant A93E/E108V/A146F (presumptively the best one according to plaque size), primary W138P mutant, and cysteine-less "wild-type" T4 lysozymes were purified after overexpression in *E. coli* cells.

Two methods were used to measure the relative activities of mutant proteins: (1) turbidometric assay, based on decreasing of absorbance of chloroform-treated cells, and (2) visual spot assay, using lysis indicator plates and several dilutions of mutant and reference (wild-type) protein.

The relative activity of the evolved W138P T4 lysozyme (A93E/E108V/A146F), determined by turbidometric assay and spot assay, was approximately 5–6% of the wild-type. The primary lysozyme mutant (W138P) was found to be about 15 times less active than its evolved progeny (~0.3% of the wild-type specific activity). Because activity measurements have been performed at room temperature and the W138P mutation is likely to affect the thermal stability of T4 lysozyme, we could expect an even larger improvement factor (the ratio between activity of evolved protein and source protein) of evolved enzymes at higher temperatures.

We determined relative thermostability of mutant lysozymes by the loss of activity after 10 min of incubation at 65 °C: more than 80% of enzyme activity was lost for W138P T4 lysozyme, about 70% for wild-type, and 35% for evolved W138P variant. The thermal stability of the evolved lysozyme variant thus appears to be higher not only than its source, W138P T4 lysozyme, but even higher than the stability of wild-type protein.

Conclusion

Proteins have evolved for many years into structures whose complexity is still not well understood. Although simple polypeptide chains can fold into functional conformations within a short time, only a relatively low information content sequence–structure dictionary is available so far. If the information for proper folding has been mounted into proteins during evolution, it is reasonable to expect that this process can be mimicked by disrupting the highly ordered protein structure, randomizing the polypeptide sequence, and selecting for function by applying a series of gradually increasing selections. As this process can be performed relatively easy and repeatedly, we believe that employing of evolutionary strategies to the defective variants of small and simple proteins may provide another route for approaching the protein folding code.

Materials and methods

Strains, plasmids, and phages

Escherichia coli W3110/F′ *lacI*^{Q} (Brent & Ptashne, 1981) was used for transformation and plasmid propagation. *E. coli* LE392 [*F*$^{-}$*hsdR574 supE44 supF58 lacY1* or Δ*(lacIZY)6, galK2, galT22, metB1, trpR55*] was used to propagate bacteriophages. Plasmid

pSB735 (Bouvier & Poteete, 1996) is a pBR322-based plasmid that expresses the T4 lysozyme variant C54T/C97A ("cysteine-less wild-type"; Pjura et al., 1990) at a low level. Plasmid pJM2 is a derivative of pSB735 with the W138P substitution in the T4 lysozyme gene. Plasmid pDR739 expresses cysteine-less wild-type T4 lysozyme at a high level (Poteete et al., 1997). The cloning vector phage λ cI857 Sam7 R::Tc736 (Bouvier & Poteete, 1996) was used for selection of functional lysozyme variants. In vitro packaging of phage DNA was performed with the Packagene system (Promega, Wisconsin), according to the directions of the manufacturer.

Selection of functional T4 lysozyme genes

Lambda cI857 Sam7 R::Tc736 DNA was digested with SfiI, ligated with SfiI-digested polymerase chain reaction (PCR)-generated fragments bearing variants of the T4 lysozyme gene, and packaged in vitro. Plaque-forming hybrids bearing functional lysozyme genes were selected by plating at 37 °C in 0.45% top agar for selection of the first generation of evolved lysozymes; subsequent generation(s) were selected at 42 °C in 0.9% top agar. [High concentrations of top agar impede plaque formation (Sambrook et al., 1989). The selection pressure can be slightly increased (decreased), when necessary, by using a higher (lower) concentration of top agar.]

Mutagenesis

Random mutagenesis was based on the method of Leung et al. (1989), in which the error rate of PCR is increased by use of a suboptimal concentration of one nucleotide and substitution of Mg^{2+} by Mn^{2+}. The following reaction conditions were used: 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 0.5 mM $MnCl_2$, 0.2 mM dATP, dCTP, dTTP, 0.04 mM dGTP, 0.5 mM primers, and 2.5–5 units of AmpliTaq DNA polymerase (Perkin Elmer, Foster City, California) per 100 μL of PCR reaction mixture. Primer sequences were: 5′-TCGACGGCCTTCCTGGCCTAAGCTTCA-3′ (SEB3) and 5′-GATCCGGCCGTTCCGGCCA-3′ (SEB6) for the plasmid-borne T4 lysozyme gene; or 5′-GACAGGCAGACGATGATGCA GAT-3′ (SEB12) and 5′-CGGAGTAGAAGATGGTAGAAATC-3′ (SEB18) for the phage-borne gene.

Site-directed mutagenesis was performed either by the method of recombinant cycles PCR (RCPCR; Jones & Howard, 1990), or by two-stage PCR using splicing by overlap extension (Ho et al., 1989).

Splicing by overlap extension was used for simultaneous randomization of three positions (146, 149, and 150) in T4 lysozyme. Plasmid pJM2 carrying a nonfunctional T4 lysozyme gene (W138P) was used as a template. The following primer combinations were used in the primary PCR: 5′-CCTAATCGCNNSAAACGANNS NNSACAACGTTTAG-3′ (MJ26) + 5′-GAAACGCCTGGTATC TTTATA-3′ (CL3), and 5′-AAACGTTGTSNNSNNTCGTTTSNN GCGATTAGGTG-3′ (MJ27) + 5′-CTCGTCGTTTGGTATGGC TTCATTC-3′ (SEB8), where N means an equal amount (25%) of each nucleotide, and S is 50% each G and C. To produce larger amounts of strands overlapping and extending in the second stage, asymmetric PCR with a 10-fold excess of primers SEB8 and CL3 over MJ26 and MJ27 was used for obtaining primary PCR products.

Primer SEB8 corresponds to a forward sequence on plasmid pSB735, situated upstream of T4 lysozyme gene. Primer CL3 corresponds to a reverse pSB735 sequence downstream of T4 lysozyme gene. Primary PCR products were purified from 1% low melting agarose by the use of the Wizard DNA clean up system (Promega, Madison, Wisconsin), spliced, then extended and amplified in PCR-like reactions without primers. After five cycles, nested primers SEB3 and SEB6 were added and spliced product was amplified in an additional 30 cycles. Second-stage PCR product was directly purified with the Wizard purification system, digested with SfiI, and inserted into SfiI-digested phage lambda vector.

DNA shuffling

Recombination by DNA shuffling was accomplished by the use of a slightly modified version of the method of Stemmer (1994). Mutant T4 lysozyme genes from selected phage plaques were amplified directly from phage plaque suspension by using PCR. Single plaques were picked and resuspended in 150 μL of SM buffer (10 mM Tris-HCl pH 7.5, 10 mM $MgSO_4$, 0.1 mg/mL gelatin). An amount of 40 μL of the plaque suspension was used as a template in 100 μL PCRs using the lambda arm primers SEB12 and SEB18. PCR products were mixed and purified from 1% low melting agarose with the Wizard system (Promega). About 3 μg of purified DNA was digested with 0.6 unit DNaseI (Sigma, St. Louis, Missouri) in 100 μL of 50 mM Tris-HCl, pH 7.5, 1 mM $MgCl_2$, for 5 min at room temperature (time required for optimal digestion, producing an average fragment size of about 100 bp, was determined by gel electrophoresis after 1, 5, 15, and 30 min of digestion). DNaseI was inactivated by boiling for 15 min, and intact T4 lysozyme genes were reassembled by random recombination in a PCR-like reaction without primers. A PCR program of 95 °C for 5 min; 10 cycles of 94 °C for 1 min, 70 °C for 1 min, and 72 °C for 3 min; 10 cycles of 94 °C for 1 min, 60 °C for 1 min, and 72 °C for 3 min; 20 cycles of 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 3 min was used. After 10-fold dilution of this product, reassembled products were amplified in a normal PCR reaction with the nested primers SEB3 and SEB6. PCR product was purified with the Wizard PCR system (Promega), digested with SfiI, and cloned into the phage lambda vector.

Expression and purification of mutant proteins

Mutant lysozyme genes were expressed after replacement of the wild-type lysozyme gene in pDR739 with mutant variants via NdeI and BglII sites, as described previously (Poteete et al., 1997). Bacteria carrying the mutant lysozyme-expressing plasmids were grown at 37 °C to a density of 2×10^8, at which point the temperature was decreased to 30 °C and IPTG (1 mM final concentration) was added. After 2 h of continued aeration, partial cell lysis was observed for wild-type and evolved mutant proteins. The cell pellet was resuspended to 1/10 volume of original culture with lysis buffer (10 mM Tris-HCl, pH 7.2; 10 mM EDTA; 0.1 μg/mL PMSF). After lysis, 20 mM $MgCl_2$ and 0.1 μg/mL of DNaseI were added. The lysate was centrifuged, the supernatant diluted by adding 5–10 volumes of distilled water, and CM-Sepharose was added (1 mL of a slurry, in which the swollen beads accounted for approximately 50% of the volume, for each 100 mL of culture). Lysozyme was allowed to adsorb overnight to the CM-Sepharose with gentle mixing. The slurry was then packed into a small column, and extensively washed with 50 mM Tris-HCl, pH 7.2, 1 mM EDTA. Lysozyme was eluted with 0.15 M NaCl in the same buffer. The resulting protein at this stage is generally >90% pure as judged by SDS-PAGE. Alternatively, lysozymes were purified from culture media by the same procedure without employing a lysis

step. Proteins were concentrated by ultrafiltration using centrifugal concentrators Centricon-10 (for smaller volumes) or Centriprep-10 (for larger volumes) supplied by Amicon (Beverly, Massachusetts).

Enzyme assays

Lysozyme activity was measured by the initial rate of decrease in the turbidity (A450) of chloroform-treated *E. coli* cells at room temperature. Lysozyme substrate was prepared by adding 0.25 mL of chloroform into 50 mL of midlog phase *E. coli* cells grown in LB broth and shaking for 1 h at 37 °C. Harvested cells were resuspended in 50 mL of 50 mM Tris-HCl, pH 7.5. Alternatively, lysis indicator plates (Poteete et al., 1991) were used for determination of enzyme activity. Plates were spotted with samples of serial twofold dilutions of lysozyme preparations, and incubated at room temperature to determine the relative amounts required to give comparable lysis zones. Lysozyme concentrations were calculated from A_{280}, assuming the molar absorption coefficient of wild-type T4 lysozyme to be 25,440 and that of the W138P mutant to be 19,940. Molar absorption coefficients were estimated according to Pace et al. (1995).

Screening of revertant plaques by hybridization

To find second-site revertants in the background of primary-site revertants, lysozyme DNAs were individually amplified from restreaked plaques by primers SEB12 and SEB18. A PCR program of 95 °C for 5 min; 30 cycles of 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 3 min was used. After heat denaturation for 5 min in boiling water bath, samples were rapidly chilled on ice and 5 μL of each sample was dot blotted on Hybond-N membrane (Amersham, Arlington Heights, Illinois), UV crosslinked and hybridized overnight at 37 °C with a labeled oligonucleotide probe (Sambrook et al., 1989) carrying the primary, W138P mutation. The membranes were washed 2 × 30 min at 37 °C with wash buffer containing 30% formamide, 1 × SSC, 1 mM EDTA, 0.1% SDS, air dried, and exposed to film.

Sequencing of mutants

Single plaques were resuspended in 0.15 mL of SM buffer, and 40 μL was used as a template in 100 μL PCR reactions with primers SEB12 and SEB18. PCR products were purified by the use of the Wizard PCR system (Promega, Wisconsin) and sequenced by Dyedeoxy Terminator (Perkin Elmer) cycle sequencing at the Nucleic Acid Facility (University of Massachusetts Medical Center, Worcester, Massachusetts).

Molecular graphics

Molecular graphics program RasMol (Sayle & Milner-White, 1995) was used for visualization of protein structures.

Acknowledgments

This research was supported by grant AI24083 from the National Institutes of Health. We thank Kenan Murphy and Anita Fenton for helpful discussions and for critical reading of manuscript.

References

Alber T, Dao-Pin S, Nye JA, Muchmore DC, Matthews BW. 1987. Temperature-sensitive mutations of bacteriophage T4 lysozyme occur at sites with low mobility and low solvent accessibility in the folded protein. *Biochemistry 26*:3754–3758.

Baldwin E, Xu J, Hajiseyedjavadi O, Baase WA, Matthews BW. 1996. Thermodynamic and structural compensation in "size-switch" core repacking variants of bacteriophage T4 lysozyme. *J Mol Biol 259*:542–559.

Bouvier SE, Poteete AR. 1996. Second-site reversion of a structural defect in bacteriophage T4 lysozyme. *FASEB J 10*:159–163.

Brent R, Ptashne M. 1981. Mechanism of action of the lexA gene product. *Proc Natl Acad Sci USA 78*:4204–4208.

Crameri A, Cwirla S, Stemmer WP. 1996. Construction and evolution of antibody-phage libraries by DNA shuffling. *Nat Med 2*:100–102.

Crameri A, Dawes G, Rodrigues E Jr, Silver S, Stemmer WPC. 1997. Molecular evolution of an arsenate detoxification pathway by DNA shuffling. *Nat Biotechnol 15*:436–438.

Crameri A, Raillard SA, Bermudez E, Stemmer WPC. 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature 391*:288–291.

Dao-pin S, Soderlind E, Baase WA, Wozniak JA, Sauer U, Matthews BW. 1991. Cumulative site-directed charge-change replacements in bacteriophage T4 lysozyme suggest that long-range electrostatic interactions contribute little to protein stability. *J Mol Biol 221*:873–887.

Elwell ML, Schellman JA. 1979. Stability of phage T4 lysozymes. II. Unfolding with guanidinium chloride. *Biochim Biophys Acta 580*:327–338.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using polymerase chain reaction. *Gene 77*:51–59.

Jones DH, Howard BH. 1990. A rapid method for site-specific mutagenesis and directional subcloning by using the polymerase chain reaction to generate recombinant circles. *Biotechniques 8*:178–183.

Kast P, Hilvert D. 1997. 3D structural information as a guide to protein engineering using genetic selection. *Curr Opin Struct Biol 7*:470–479.

Kim B, Hathaway TR, Loeb LA. 1996. Human immunodeficiency virus reverse transcriptase—Functional mutants obtained by random mutagenesis coupled with genetic selection in *Escherichia coli*. *J Biol Chem 271*:4872–4878.

Knight JA, Hardy LW, Rennell D, Herrick D, Poteete AR. 1987. Mutations in an upstream regulatory sequence that increase expression of the bacteriophage T4 lysozyme gene. *J Bacteriol 169*:4630–4636.

Lanzer M, Bujard H. 1988. Promoters largely determine the efficiency of repressor action. *Proc Natl Acad Sci USA 85*:8973–8977.

Leung DW, Chen E, Goeddel DV. 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique 1*:11–15.

Matthews BW. 1995. Studies on protein stability with T4 lysozyme. *Adv Prot Chem 46*:249–278.

Moore JC, Jin HM, Kuchner O, Arnold FH. 1997. Strategies for the in vitro evolution of protein function: Enzyme evolution by random recombination of improved sequences. *J Mol Biol 272*:336–347.

Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. 1995. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci 4*:2411–2423.

Pjura P, Matsumura M, Baase WA, Matthews BW. 1993. Development of an in vivo method to identify mutants of phage T4 lysozyme of enhanced thermostability. *Protein Sci 2*:2217–2225.

Pjura PE, Matsumura M, Wozniak JA, Matthews BW. 1990. Structure of a thermostable disulfide-bridge mutant of phage T4 lysozyme shows that an engineered cross-link in a flexible region does not increase the rigidity of the folded protein. *Biochemistry 29*:2592–2598.

Poteete AR, Dao-Pin S, Nicholson H, Matthews BW. 1991. Second-site revertants of an inactive T4 lysozyme mutant restore activity by re-structuring the active site cleft. *Biochemistry 30*:1425–1432.

Poteete AR, Rennell D, Bouvier SE. 1992. Functional significance of conserved amino acid residues. *Proteins 13*:38–40.

Poteete AR, Rennell D, Bouvier SE, Hardy LW. 1997. Alteration of T4 lysozyme structure by second-site reversion of deleterious mutations. *Protein Sci 6*:2418–2425.

Remington SJ, Anderson WF, Owen J, Ten Eyck LF, Grainger CT, Matthews BW. 1978. Structure of the lysozyme from bacteriophage T4: An electron density map at 2.4 A resolution. *J Mol Biol 118*:81–98.

Rennell D, Bouvier SE, Hardy LW, Poteete AR. 1991. Systematic mutation of bacteriophage T4 lysozyme. *J Mol Biol 222*:67–88.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sayle RA, Milner-White EJ. 1995. RASMOL: Biomolecular graphics for all. *Trends Biochem Sci 20*:374–376.

Shao Z, Arnold FH. 1996. Engineering new functions and altering existing functions. *Curr Opin Struct Biol 6*:513–518.

Stemmer WPC. 1994. Rapid evolution of a protein in vitro by DNA shuffling. *Nature 370*:389–391.

Streisinger G, Mukai F, Dreyer WJ, Miller B, Horiuchi, S. 1961. Mutations affecting the lysozyme of phage T4. *Cold Spring Harbor Symp Quant Biol 26*:25–30.

Tsugita A, Inouye M, Terzaghi E, Streisinger G. 1968. Purification of bacteriophage T4 lysozyme. *J Biol Chem 243*:391–397.

Warren MS, Benkovic SJ. 1997. Combinatorial manipulation of three key active site residues in glycinamide ribonucleotide transformylase. *Protein Eng 10*:63–68.

Weaver LH, Matthews BW. 1987. Structure of bacteriophage T4 lysozyme refined at 1.7 A resolution. *J Mol Biol 193*:189–199.

Zhang JH, Dawes G, Stemmer WPC. 1997. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc Natl Acad Sci USA 94*:4504–4509.

Zhang X, Baase WA, Schoichet BK, Wilson KP, Matthews BW. 1995. Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive. *Protein Eng 8*:1017–1022.

Zhao H, Arnold FH. 1997a. Functional and nonfunctional mutations distinguished by random recombination of homologous genes. *Proc Natl Acad Sci USA 94*:7997–8000.

Zhao H, Arnold FH. 1997b. Combinatorial protein design: Strategies for screening protein libraries. *Curr Opin Struct Biol 7*:480–485.