Rapid Evolution of Reversible Denaturation and Elevated Melting Temperature in a Microbial Haloalkane Dehalogenase

Kevin A. Gray,[a,*] Toby H. Richardson,[a] Keith Kretz,[a] Jay M. Short,[a] Flash Bartnek,[a] Ryan Knowles,[a] Lynn Kan,[b] Paul E. Swanson,[b] Dan E. Robertson[a]

[a] Diversa Corporation, 4955 Director's Place, San Diego, CA 92121, USA
Fax: (+1) 858–526–5661, e-mail: kgray@diversa.com
[b] The Dow Chemical Company, Biotechnology Capability, 5501 Oberlin Dr., San Diego, CA 92121, USA

Received May 17, 2001; Accepted July 17, 2001

Abstract: Haloalkane dehalogenases have the potential for use in high-value biocatalytic processes to convert haloalkanes into epoxides via intermediate haloalcohols. Initial bioreactor studies probing the hydrolysis of 1,2,3-trichloropropane by immobilized wild-type dehalogenase isolated from *Rhodococcus rhodochrous* demonstrated, however, that productivity was too low to realize a commercially viable process. A strategy to increase enzyme performance was undertaken to increase the reaction temperature, however it was determined that the wild-type enzyme was not stable for long periods of time at elevated temperatures. The accelerated laboratory evolution technique of Gene Site Saturation Mutagenesis (GSSM^{TM}) was used to create a clonal enzyme library comprising all single site sequence variants of the *Rhodococcus* enzyme. Using high throughput screening techniques and rapid kinetics assays, this library was probed for improvements in thermostability and for the ability of the enzyme to undergo a fully reversible cycle of thermal denaturation-renaturation. Eight single site mutants were discovered that had considerable effects on these aspects of the dehalogenase phenotype. Compared to the parental dehalogenase ($t_{1/2}$ = 11 minutes at 55 °C) single site variants have half-lives ranging from 300 minutes to 2700 minutes. Combinations of these mutations dramatically improved the half-life demonstrating the enhancing effects of mutational additivity. Combining five of the mutations into a single protein (Dhla5) improved the half-life to 29,000 min and a combination of all eight single-site mutations (Dhla8) increased the half-life by another factor of ten. Thus, the final Dhla8 protein was 30,000 times more stable than the parent molecule as measured by its ability to refold after denaturation at high temperature. Kinetic analysis showed that the improvement in thermal stability associated with Dhla5 did not negatively affect the rate of catalysis at ambient temperature, and allowed a significant increase in rate with no deactivation at 55 °C. Differential scanning calorimetry demonstrated that mutational combinations in both Dhla5 and Dhla8 led to an 8 °C increase in T_m and substantiated that partial reversibility (Dhla5) and full reversibility of Dhla8. Thermal denaturation of Dhla8 was fully reversible upon scanning up to 90 °C. Bioreactor studies showed that improved thermal stability of Dhla5 and Dhla8 correlated qualitatively with increased productivity when haloalkane hydrolysis was conducted using immobilized forms of these evolved enzymes under high temperature conditions.

Keywords: calorimetry; dehalogenation; enzyme catalysis; halogens; molecular evolution

Abbreviations: TCP: 1,2,3-trichloropropane; DCH: 2,3-dichloropropanol; GSSM^{TM}: Gene Site Saturation Mutagenesis; DSC: differential scanning calorimetry; Dhla5: 5-amino acid substituted dehalogenase: Dhla8, 8-amino acid substituted dehalogenase

  1615-4150/01/34306–7-607–617 $ 17.50-.50/0

Introduction

A dogma of evolutionary theory states that all information required for the ultimate expression of a particular enzyme's complex phenotype resides in its amino acid sequence. Nature randomly and continually creates changes in the enzyme sequence and tests these variants for fitness with the result of providing an efficient catalyst within a complex ensemble of cellular reactions. Recently developed technologies for rapid gene evolution have demonstrated that an enzyme may be removed from the thermodynamic and kinetic constraints imposed by the cellular metabolic balance and evolved to new phenotypic optima by randomly and rapidly searching the side-chain variants clustered around an initial amino acid sequence. In an impressive display of phenotypic plasticity the techniques for rapid evolution have been used to create robust, specific enzyme catalysts suitable for industrial applications. The following study was undertaken to optimize an enzyme catalyst for a predefined bioreactor-based transformation. The starting phenotype of this enzyme included certain characteristics, e.g., turnover number, product inhibition, which required optimization to ensure ultimate economic productivity. To first address turnover, the option of raising reactor temperature and leveraging the Arrhenius effect on catalytic rate was chosen. Laboratory evolution techniques were used in conjunction with high throughput assays to search for increased thermal reversibility ($t_{1/2}$) and elevation of enzyme melting temperature (T_m).

Commercial manufacture of epoxides through alkene halogenation/hydrolysis protocols can result in the generation of polyhalogenated alkanes as undesirable side-products. In order to improve the economics of these processes the recovery and conversion to the corresponding halohydrins has been considered.[1] The halohydrins can subsequently be transformed into the desired epoxides. Haloalkane dehalogenases (Dhla) catalyze the hydrolysis of a variety of halogenated compounds into alcohols with release of the corresponding hydrogen halide (Equation 1).

$$\text{R-X} + \text{H}_2\text{O} \xrightarrow{\text{Dhla}} \text{R-OH} + \text{HX} \qquad (1)$$

Two general classes of dehalogenases have been identified by sequence homology; those bearing amino acid identity to the *Xanthobacter autotrophicus* enzyme and those more closely related to the enzymes from *Rhodococcus rhodochrous* and *Sphingomonas paucimobilis*.[2,3,4] These two sequence classes have been further divided into at least three substrate specificity classes. A substantial amount of mechanistic and structural information is available on the haloalkane dehalogenases from *Rhodococcus, Sphingomonas* and *Xanthobacter*.[5–15] All are members of a superfamily of proteins containing an α/β hydrolase fold. The overall structures of the three enzymes are very similar and each contains a triad of catalytic residues (D117, H283, and E141 in *Rhodococcus*) that forms the active site with the carboxylate of the aspartic acid residue acting as the nucleophile that displaces the halide group.[6] Dehalogenases have a distinctive cap domain shielding the catalytic triad, however, crystallographic data indicate that the *Rhodococcus* and *Xanthobacter* enzymes diverge considerably in this region. Kinetic analyses have shown that the enzymes also differ in inhibition patterns and substrate specificity and it has been postulated that the cap domain plays a major role in the substrate specificity of the different enzymes. The *Rhodococcus* enzyme has been shown to have higher activity towards multiply halogenated long-chain alkanes relative to the *Xanthobacter* enzyme[5,16]. For instance, *Rhodococcus* dehalogenase is the first enzyme that has been reported to have significant activity on 1,2,3-trichloropropane (TCP); however, the rate of this transformation was substantially slower than rates observed with analogous brominated substrates.[16]

In order to develop an efficient biocatalytic procedure for the conversion of halogenated alkanes to halohydrin products, we sought to improve the performance of the *Rhodococcus* enzyme in a bioreactor. Performance in a reactor is a combination of reaction rate and stability. A convenient method for increasing reaction rate is to increase the temperature of the reaction. According to the Arrhenius equation, the rate of an enzymatic reaction will approximately double for every 10 °C increase in temperature. However, most enzymes are not stable at elevated temperatures, and tend to be irreversibly inactivated at these temperatures. Mutagenesis has been used to improve the thermostability of mesophilic enzymes;[17–21] however, there is very little consensus on what types of interactions are important to achieve a thermally stable protein. Recently, there has been a tremendous amount of effort placed on understanding structural features that confer stability on naturally thermophilic proteins (for a review, see ref.[22]) with the hope that a more directed approach can be used to develop thermal stable derivatives of mesophilic proteins.

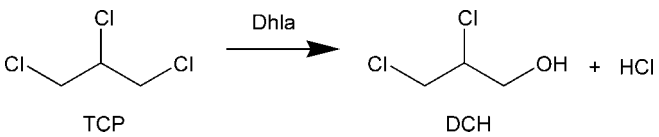
In the present study we have used a novel, highly efficient mutagenesis technique referred to as Gene Site Saturation Mutagenesis (GSSM™) in combination with high throughput screening methods in an effort to develop extremely thermostable derivatives of the *Rhodococcus* dehalogenase. GSSM™ can be used to randomize each successive codon in a gene, thereby generating pools of variants containing every amino acid substitution at every single position within a pro-

tein. Error prone PCR has generally been used to introduce nucleotide changes in many genes, but statistically only single nucleotide substitutions in any particular codon can be screened. The result of this single random nucleotide substitution is that, assuming a perfect distribution of nucleotide substitutions and sampling, an average of 5 – 6 amino acid changes may be sampled at any particular site in the protein. GSSM™, in contrast, theoretically allows all single site mutants to be sampled. It is important to note that, as part of our high-throughput screening program, we demanded that the more stable variants retain high activity at room temperature, thus eliminating the possibility that we would develop a protein that requires high temperature for activity.

Results

Thermal Stability of the Parental Dehalogenase

In the present study, we have examined the effectiveness of our directed evolution strategy GSSM to produce improved enzymes for hydrolysis of the representative polyhaloalkane, 1,2,3-trichloropropane, to the corresponding alcohol 2,3-dichloropropanol (Equation 2).



(2)

In particular, we have attempted to increase the thermal stability of the dehalogenase in order to avail ourselves of the increase in reaction rate and productivity coincident with an increased reaction temperature. Preliminary mutagenesis work (unpublished data) using standard error prone PCR methods resulted in a two-site mutant (G3D, C187F) of the wild-type *Rhodococcus* dehalogenase. This two-site mutant has about a four-fold improvement in k_{cat} relative to the wild-type dehalogenase and was used as the starting parental clone for the evolution studies described below.

Thermostability of the parental dehalogenase was measured by assaying activity at elevated temperatures. Figure 1 shows time courses of the dehalogenase reaction at 35 °C and 55 °C. The initial rate at 55 °C is approximately four times that at 35 °C indicating that, as expected, the rate doubled every 10 °C. However, it is clear that within the first few minutes at 55 °C the enzyme began to lose activity and was completely inactivated within 10 minutes. There was an obvious precipitate in the 55 °C reaction that did not appear at 35 °C, suggesting that the

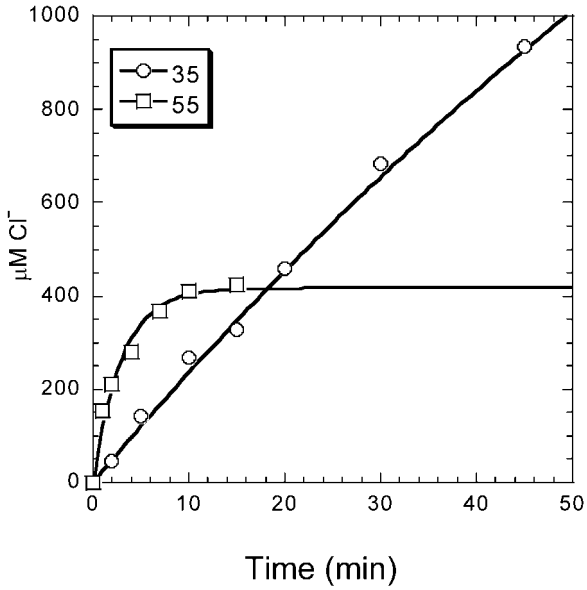


Figure 1. Enzyme progress curves for parental dehalogenase at 35 °C (○) and 55 °C (□). Activity was measured as described in the experimental section using 10 mM TCP as the substrate.

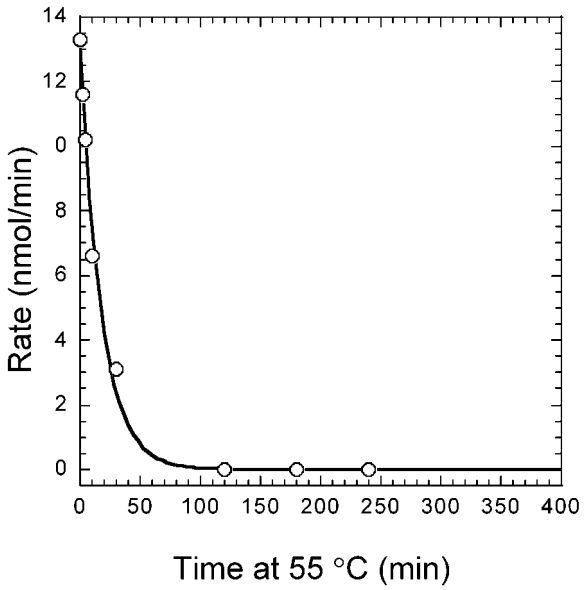


Figure 2. Thermal inactivation of the parental dehalogenase. Purified dehalogenase was incubated at 55 °C for the indicated times, removed and assayed for the ability to hydrolyze TCP at 35 °C. The data were fit to a single exponential resulting in a half-life of approximately 11 min.

protein tended to aggregate after denaturing. Thermal tolerance was measured by incubating a protein solution at 55 °C for a specified period of time, cooling it on ice, and then assaying it for activity at 35 °C. This experiment determines how well a protein will refold after thermal denaturation. Figure 2 shows that the parental dehalogenase presumably is irreversibly de-

natured at 55 °C with an apparent half-life of approximately 11 minutes.

Thermal Stability of Single Site Variants

We considered two different high-throughput screening protocols to assess thermal stability of mutant enzyme variants. In one procedure, the reaction is carried out, and activity assayed at high temperature, whereas in the second method enzyme variants are first heated to an elevated temperature for a specified period of time, allowed to cool to a lower temperature and then assayed for residual activity. The first method would allow discovery of enzymes that retain both the native structure and activity at the high temperature (i.e., a higher T_m). The second method would not directly interrogate enzyme activity at high temperatures, but rather would select for enzymes that either retain structure at elevated temperatures or that undergo reversible thermal denaturation. In reality a combination of both will occur. Technically it is difficult to analyze large numbers of samples continuously at high temperature, therefore we employed the second method for primary screening of mutagenized dehalogenases. An additional constraint placed on the screen was that no clones with room temperature activity less than the parental dehalogenase were selected even if they appeared to be stable to the heating conditions.

GSSM™ was used to mutate the parental dehalogenase gene resulting in 315 pools of mutagenized genes, with each pool randomized at one of the amino acid positions. Following cell sorting and growth, high throughput screens were conducted as described in the experimental section. Those clones retaining more activity than the wild-type controls after heating were recovered and the gene was sequenced to determine the substitution conferring increased stability. Eight single-site mutants were discovered that retained higher activity following heating to 55 °C. Only a single position (G182) resulted in multiple mutations with similar thermal stabilities. All eight mutant enzymes were isolated and half-lives at 55 °C determined. As shown in Table 1, these eight mutations increased the half-life of the wild-type dehalogenase between 30- and 250-fold. From this screen, we found that the substitutions I220L and P302A were the most effective with regard to the improvement in thermal tolerance.

Table 1. Summary of stability of parental and mutant dehalogenases. Half-lives were derived from fits to thermal inactivation curves. The values are averages of three measurements with the associated standard deviations. ND: not determined.

Clone	Half-life (min) at 55 °C	Half-life (min) at 80 °C
Wild-type	11	ND
D89G	580 ± 150	ND
F91S	700 ± 100	ND
T159L	300 ± 50	ND
G182V, Q, A	750 ± 100	ND
I220L	2700 ± 200	ND
N238T	700 ± 100	ND
W251Y	450 ± 100	ND
P302A	1000 ± 300	ND
Dhla5	29,000	13
Dhla8	ND	138

Multi-Site Mutant Enzymes Dhla5 and Dhla8

We next explored the possibility that combinations of single-site mutations may manifest themselves in an additive, if not synergistic, manner to produce enzymes with even greater thermal tolerance. The first test of this additivity effect was conducted using a single gene incorporating five of the beneficial mutations (D89G, F91S, T159L, G182Q, and I220L) to produce Dhla5. In addition, we also combined all eight beneficial mutations to generate the dehalogenase variant Dhla8. Initial comparison of the stability at 55 °C of Dhla5 and parental dehalogenase is shown in Figure 3A. As can be seen, these five amino acid substitutions cumulatively produced an enzyme with dramatically improved stability at 55 °C. In order to assess the half-life of Dhla5 through activity loss, measurements had to be collected by allowing the enzyme to remain at 55 °C over an extended period of time. Through this procedure, it was determined that Dhla5 had an apparent half-life of 29,000 minutes at 55 °C (Figure 3B), corresponding to a 3000-fold improvement over the parental enzyme.

We next examined the thermal tolerance of the eight-mutant enzyme, Dhla8 and found that the stability was impossible to measure at 55 °C. Accordingly, the temperature was raised to 80 °C where Dhla8 was found to be about ten times more stable than Dhla5, corresponding to about a 30,000-fold improvement over the wild-type protein (Figure 4).

Differential scanning calorimetry (DSC) was used to establish the melting temperature (T_m) of each enzyme and to investigate the possibility of reversible thermal denaturation at the high temperature of our experiments. DSC traces associated with the parental dehalogenase, Dhla5 and Dhla8 are shown in Figure 5. The forward scans were performed from 40 °C to 90 °C, followed by cooling back to 40 °C and then rescanning to 90 °C. As can be observed, the parental dehalogenase displays a T_m of 65 °C and the denaturation appears irreversible. Interestingly, Dhla5 and Dhla8 both exhibit an identical 8 °C increase in T_m (73 °C). Relative to Dhla5, the three additional mutations incorporated within Dhla8 apparently have no influence on the melting temperature, however, Dhla5 was shown to only partially renature after

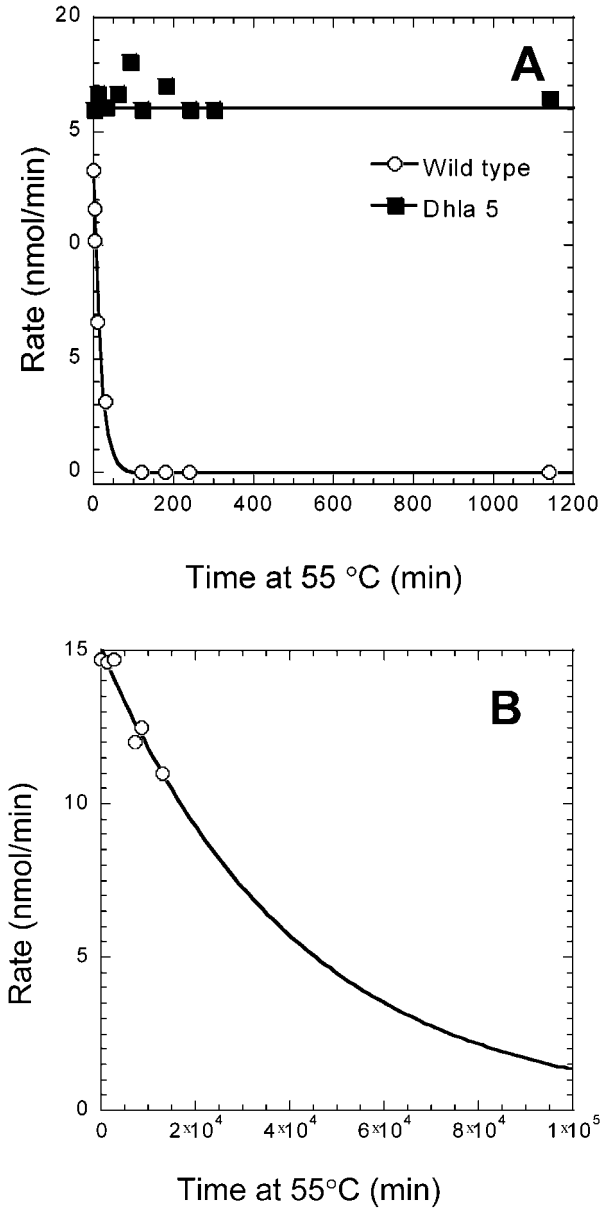


Figure 3. (**A**) Thermal inactivation of the parental dehalogenase as compared to Dhla5. The conditions were the same as Figure 2. (**B**) Long incubation times of Dhla5 at 55 °C. The data are fit to a single exponential with a half-life of 29,000 minutes.

cooling, while denaturation was completely reversible in the case of Dhla8.

High Temperature Activity

The high-throughput screening method used in the studies outlined above selected for enzymes that could tolerate high temperatures and retain activity at room temperature. However, these studies provided no indication of the activity of these enzymes

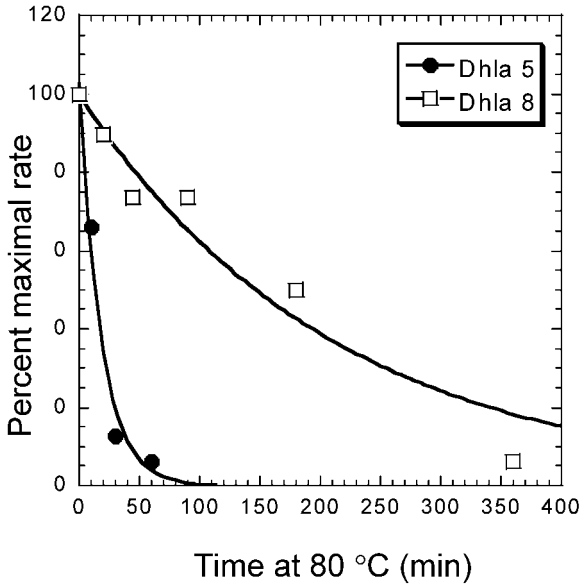


Figure 4. Thermal inactivation of Dhla5 and Dhla8 at 80 °C. The conditions were the same as Figure 1B. Fits to the data resulted in half-lives of 13 minutes and 138 minutes for Dhla5 and Dhla8, respectively.

at higher temperatures. However, the DSC results shown above demonstrate that both Dhla5 and Dhla8 have a higher T_m than the parental protein, therefore both enzymes maintain the native structure at higher temperatures. Since an objective of this study is to develop dehalogenases that could perform reactions at elevated temperatures and thus with higher rates and productivity, we have surveyed the enzymatic activity of Dhla5 at two different temperatures and compared the results with those obtained with the parental enzyme. Figure 6 compares the ability of Dhla5, Dhla8, and parental dehalogenase to catalyze the conversion of TCP to DCH at 35 °C and 55 °C. As previously observed, parental dehalogenase inactivates quickly at the higher temperature, however both Dhla5 and Dhla8 continue to turn over throughout the time of the reaction without any apparent inactivation at 55 °C. Moreover, Figure 6 demonstrates that the mutations in Dhla5 and Dhla8 do not negatively affect the specific activity of the enzymes at either temperature. As a matter of fact Dhla8 has a k_{cat} of approximately 1.5- to 2-fold higher than the wild-type protein (data not shown). Giver et al.[26] observed a similar pattern during the development of a thermally stable esterase. They too observed that a thermal stable version of the esterase was equally as active as the parent at room temperature. Interestingly, throughout this work, we have been able to demonstrate that the phenotypic properties of activity and thermal stability may be independently addressed in the evolution of enzymes. Hence, in the present study we have shown that the thermal stability of dehalo-

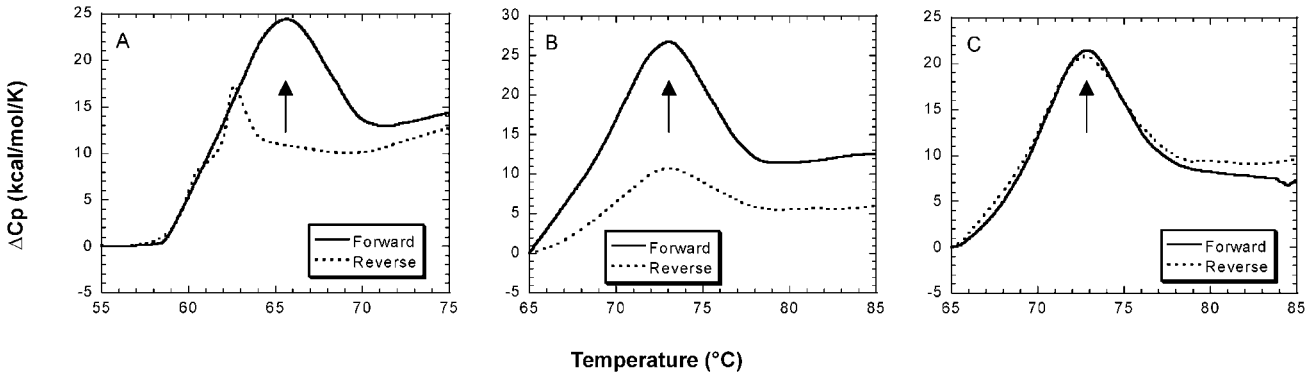


Figure 5. Differential scanning calorimetry of parental (**A**), Dhla5 (**B**), and Dhla8 (**C**) dehalogenases. The conditions are described in the experimental section. Arrows in each figure indicate the melting temperature, T_m.

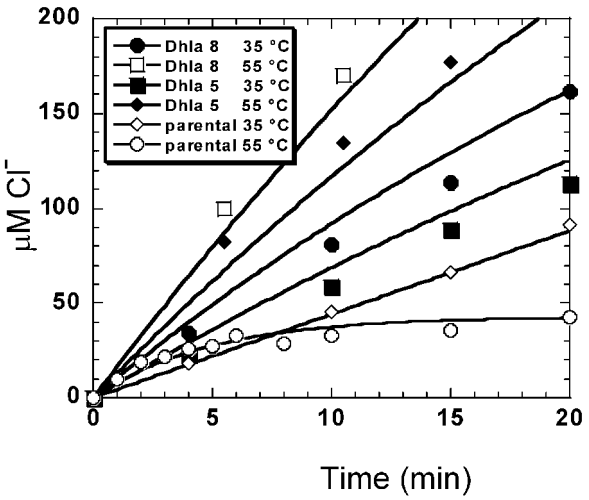


Figure 6. Enzyme progress curves for parental (◊, ○), Dhla5 (■, ♦), and Dhla8 (•, □) dehalogenase at 35 °C and 55 °C, respectively. Activity was measured as described in the experimental section using 10 mM TCP as the substrate.

genases and the reversible nature of folding may be greatly improved without affecting the activity of the enzyme.

Reactor Performance

As part of our efforts to establish the commercial viability of our new thermostable dehalogenase, we sought to demonstrate that the transformation of TCP to DCH could be performed in a bioreactor that is maintained at an elevated temperature, thereby improving catalyst productivity. One would expect an approximate four-fold increase in rate by increasing the reactor temperature by 20 °C. A large percentage of commercial biocatalytic processes involve the use of immobilized enzymes. Accordingly, we have immobilized the parental dehalogenase as well as Dhla5 and Dhla8 onto a cross-linked alumina support for bioreactor studies. We believed that data collected with soluble enzyme would not translate directly to results achieved with an immobilized enzyme system since several factors in addition to thermostability contribute to reactor performance. For example, if any of the variants are more or less sensitive to glutaraldehyde cross-linking some effect will be observed in productivity. We did not attempt to measure these other parameters. Table 2 summarizes the reactor performance of the three dehalogenases. Immobilization of the parental enzyme appears to have a stabilizing effect relative to the soluble enzyme and, at 55 °C, a half-life of approximately 35 h was observed. Dhla5 and Dhla8 are considerably more stable than the parental protein (150 and 630 h, respectively), however not to the same degree as soluble protein. Both enzymes have higher initial and operational productivity. Overall, reactor productivity of Dhla8 is improved 25-fold over parental dehalogenase. Eighty percent of this improvement is due to the increase in $T_{1/2}$ and the other 20% comes from the improvement in initial productivity. The improvement in initial productivity may be partially a result of the higher turn over number of Dhla8 relative to wild-type. Therefore the improvements we observed with the soluble form of the enzyme correlate with a reactor system.

Table 2. Comparison of continuous reactor performance for thermal stable dehalogenase enzymes at 55 °C. Initial productivity was determined at 50% conversion of TCP to DCH. Operational productivity is the product of operation lifetime (4 half lives) and initial productivity.

Enzyme	$T_{1/2}$ (h)	Initial productivity (lb product/lb enzyme/h)	Operational productivity (lb product/lb enzyme)
WT	35	1.54	72
Dhla5	150	1.97	394
Dhla8	630	2.13	1788

Discussion

There are innumerable reports of the use of mutagenesis and screening to develop improved variants of wild-type proteins, for example see[18,19,25–28]. However, to our knowledge, this is the first report of a gene that has systematically been mutagenized so that every position within the protein would contain all twenty amino acids. GSSM offers a tremendous advantage over other less comprehensive mutagenic techniques such as error prone PCR, since all possible amino acids can be accessed through this method. Furthermore, the mutagenic pools can be stably maintained for long periods of time and screened for any number of parameters, for example, stability, turnover, inhibition, etc. It is thus possible to map every position within a particular protein for its sensitivity to mutation and determine which positions have a positive, negative, or neutral affect on a particular property. For instance, GSSM may quickly be able to find active site residues since most changes at those positions would result in a severely inhibited protein. Also, those positions that are involved in important structural interactions may also appear to be very inflexible to mutation.

In this report we completely scanned the primary structure of the haloalkane dehalogenase from *Rhodococcus rhodochrous* to discover which mutations improve the thermostability of the enzyme. This was done so that we could capture the rate enhancement coincident with a reaction temperature increase without thermally inactivating the enzyme. The screening effort resulted in the discovery of eight single-site substitutions that individually conferred some degree of thermostability to the dehalogenase. The mutations were then sequentially added to a single gene to generate two multi-site variants, Dhla5 and Dhla8 with five and eight substitutions, respectively (see Figure 7 for the positions within the three-dimensional structure of the protein). Thermostability of each mutant was measured by heating the protein to an elevated temperature (55 °C) for specified periods of time, cool the protein, and then measure residual enzymatic activity at 35 °C. In actuality this analysis is a measure of either the ability of the protein to remain native at high temperature (increase in T_m) or the ability of the protein to refold after thermal denaturation.

A general scheme to describe the events that occur during heating is shown in Equation 3 where k_1 and k_{-1} represent the equilibrium between the folded and unfolded state of a protein, and k_2 represents the composite rate of a number of irreversible events prohibiting a protein from refolding after thermal denaturation. These events may be aggregation, misfolding, or chemistry at a side chain. Melting temperature (T_m) affects the equilibrium between native and denatured forms of the protein, i.e., at this temperature 50% of the protein is in the native, folded state and 50% is in an unfolded, or denatured state. On the other hand, half-life at a temperature approaching T_m is a measure of k_2. In the following discussion we consider T_m and half-life as two separable properties and discuss how each of the substitutions may contribute to the observed effects.

$$\text{Native} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{Denatured} \xrightarrow{k_2} \text{Irreversible events} \qquad (3)$$

Melting Temperature

We did not measure the T_m values of each of the single-site mutants, however we know that the combination of five of these mutations into a single protein elevated T_m by 8 °C. These five mutations (D89G,

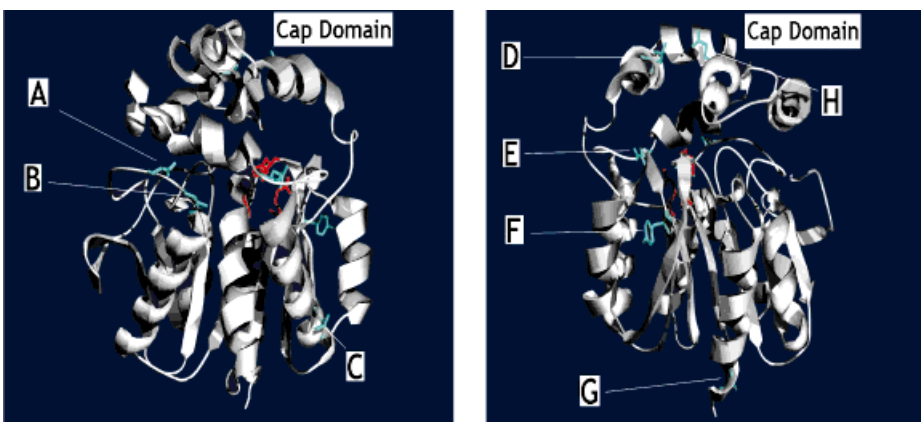


Figure 7. A ribbon structure[6] of the *Rhodococcus* dehalogenases is displayed using the Swiss pdb viewer.[45] The catalytic triad residues are shown in red and the thermostable mutations (A-H) are shown in blue. A = D89G, B = F91S, C = N238T, D = T159L, E = I220L, F= W251Y, G = P302A, H = G182Q. For reference the cap domain is also labeled.

Table 3. Location of each of the eight substitutions based on the three-dimensional crystal structure of the *Rhodococcus* dehalogenase[6] and a proposed structural effect of the mutation.

Mutation	Location	Structural Speculation
D89G	Located on the loop between β-sheet 4 and α-helix 2.	
F91S	Located at the N' position of α-helix 2.	Removal of a "hydrophobic staple" between N', Phe91 and N4, Val96.[44] Potential additional hydrogen bond with the carboxylate of E219.
T159L	Located in the C4 position of α-helix 4.	Removal of a β-branch in α-helix 4.
G182Q	Located on α-helix 6.	
I220L	Located at C' of α-helix 8.	
N238T	Located in the C4 position of a-helix 9.	Potential additional H-bond between hydroxy of T238 and main chain carbonyl group. Introduction of β-branch.
W251Y	Located on β-sheet 7.	Potential additional H-bond between hydroxy of Y251 and carboxylate of D277.
P302A	Located on α-helix 11.	Potential additional H-bond with main chain carbonyl group.

F91S, T159L, G182Q, I220L, and N238T) must, therefore, contribute to the increase in T_m. There are several theories as to the structural determinants conferring thermal stability on proteins. These include increased hydrophobic packing,[29,30,31] decreased length of surface loops[32], hydrogen bonding patterns,[33] and the role of ionic interactions.[20] There have also been numerous experiments demonstrating the stabilizing effect induced by helix capping in proteins[34,35] and peptides.[36–39] However, the significance of these observations to protein thermostability is not fully understood. Table 3 lists the mutations that were discovered and offers some plausible explanation as to how these changes may improve T_m. For example, T159 forms a β-branch in α-helix 4 and β-branches have been shown to destabilize helices,[40,41] therefore the substitution, T159L may well stabilize this helix by removing the β-branch and improve the overall stability of the protein. The substitution F91S has the potential to form an additional hydrogen bond between the serine hydroxy group and an acceptor. In doing so this substitution could increase the stability of the native form of the enzyme. Similarly, the mutation N238T occurs in α-helix 9 and may introduce an additional hydrogen bond with a main chain carbonyl group. The combination of these interactions may act synergistically and enhance T_m far greater than the additive effects of the individual substitutions, however without T_m measurements for the single sites it is not possible to ascertain this.

Half-Life at High Temperature

In the following discussion we consider only the increase in thermal tolerance, i.e., the ability of the protein to refold after being thermally denatured. The actual molecular processes by which a protein denatures and aggregates are very complex and is beyond the scope of this discussion, however, in general as k_2 approaches zero, the half-life should get longer. A completely reversible process will have k_2 equivalent to zero. Furthermore, irreversible denaturation and aggregation may be caused by some chemical event at particular side chains (e.g., oxidation, deamination, etc.) or by non-productive hydrophobic intra- and intermolecular interactions of the unfolded form of the protein. The mutation with the most dramatic affect on half-life is I220L with an approximate 250-fold improvement. In the crystal structure of the *Rhodococcus* enzyme (Figure 7) I220 can be found at the C-terminal position of α-helix 8 and is predicted to be solvent inaccessible. The side chains of isoleucine and leucine have almost equivalent hydrophobicity values,[42] therefore it is difficult to explain why this substitution has increased the half-life, however it may alter hydrophobic packing or the way the side chain interacts in the unfolded state and subsequently decrease k_2.

One of the most interesting results presented here is the difference between Dhla5 and Dhla8. Both Dhla5 and Dhla8 have identical melting temperatures yet Dhla8 is approximately ten times more tolerant to 80 °C than Dhla5, therefore these three additional mutations must contribute solely to the increase in half-life. It has been predicted that the effects of individual stabilizing mutations are roughly additive[43] yet Dhla5 has a substantially longer half-life at 55 °C than the simple addition of all five single sites and the same is true of Dhla8. However, since half-life is somewhat independent of T_m, additivity may not hold. Again, using a kinetic argument, k_2 must be much slower in Dhla8 relative to Dhla5. This rate, however, is not zero, as the DSC would suggest, since eventually Dhla8 does irreversibly inactivate at 80 °C. One of the three mutations in Dhla8 not present in Dhla5 is P302A, and *cis-trans* isomerization of proline peptide bonds is intrinsically very slow. Potentially the substitution of alanine at this site has increased the efficiency by which Dhla8 can refold after denaturation resulting in a much longer half-life at high temperatures. There are 27 other proline residues within the dehalogenase and it's interesting that only a substitution at P302 had the stabilizing effect. The other two substitutions, W251Y and N238T,

replace relatively non-polar side chains with more polar side chains potentially, again, decreasing the chances of non-productive hydrophobic interactions occurring when the protein is in the unfolded state and causing aggregation. This may well increase the possibility that the protein will more efficiently refold after thermal denaturation.

Conclusions

Using GSSM and high throughput screening we have developed very thermostable derivatives of a haloalkane dehalogenase and discovered substitutions that would have been impossible, *a priori*, to predict. These substitutions have improved both the melting temperature and the reversibility of thermal denaturation of the enzyme. Only after analyzing the biochemical properties and the three-dimensional structure of the protein can we begin to rationalize these substitutions. This information is important in designing future evolution studies.

Experimental Section

Mutagenesis

The *Rhodococcus* Dhla gene was cloned into the expression vector pTrcHis (Invitrogen) with an N-terminal His tag and a C-terminal EXFLAG tag as previously described[6]. GSSM™ was performed as described[25] using oligonucleotides to direct the randomization at each position within the protein. In brief, 32-fold degenerate oligonucleotides (N,N,G/T) were used for each codon in the gene so that all possible amino acids would be encoded by the resulting codons. Separate degenerate oligonucleotides were used to mutagenize each codon in the gene. Dhla5 and Dhla8 were constructed by the sequential addition of each of the single site changes by site directed mutagenesis.

Enzyme Assays

Substrates were 1,2,3-trichloropropane (TCP), 1-bromopropane (BP), or 1,2-dibromoethane (DBE) at saturating concentrations. None of the variants were tested on substrates other than these. All substrates were purchased from Sigma. Buffer was saturated with TCP, BP, or DBE by stirring vigorously overnight. After saturation the pH of the buffer was readjusted. Dehalogenase activity was assayed by either a pH-indicator system as previously described[5] or by a Cl^- indicator system.[24] pH shift assays were carried out in 1 mM Tris/SO_4^{2-}, pH 7.8 using 100 μM bromthymol blue as the indicator. Enzymatic activity was monitored by observation of a decrease in absorbance at 600 nm. Chloride ion specific assays were carried out in 50 mM Tris/SO_4^{2-}, pH 7.8. At defined time points the reaction was quenched by the addition of $Fe(NO_3)_3$ in $HClO_4$ and the color (450 nm) developed with $Hg(SCN)_2$ in ethanol. Standard curves showed a linear response between 0 and 0.3 mM Cl^-.

High Throughput Screen

We have adapted the pH-indicator assay to be used in microtiter plates in which individual clones have been arrayed and grown. Briefly, *E. coli* transformants from each mutagenic reaction were pooled and single cells were arrayed by fluorescence activated cell sorting (FACS) into 384 well plates containing media (LB with 25 μg/ml ampicillin) and inducer (0.1 mM isopropyl thiogalactoside). One plate was used for each position within the protein and each plate contained 16 wells of parental transformants. Growth was allowed to proceed for two days at 30 °C, after which the plates were centrifuged and media removed. Cell pellets were resuspended in 1 mM Tris/SO_4^{2-}, pH 7.8. From each master plate two daughter plates were made of which one was heated to 55 °C for 30 minutes and the other remained at 25 °C. Heated plates were allowed to cool and then substrate (8 mM TCP in 1 mM Tris/SO_4^{2-}, pH 7.8) and indicator (100 μM bromthymol blue) were added to each well to initiate the reaction. Enzyme turnover was monitored periodically by measuring absorbance at 600 nm in a microplate reader. Rates of the heat-treated and non-heat-treated plates were compared in order to measure thermal stability.

Enzyme Purification

Parental dehalogenase and selected variants were purified by Ni-chelate chromatography (Pharmacia). Induced overnight cultures were centrifuged and suspended in 50 mM NaP_i, 100 mM NaCl, pH 7.4. Cells were disrupted by sonication and the soluble protein fraction was isolated by centrifugation. The cell-free extract was applied to a Ni-chelate sepharose fast flow column and then washed with 20 mL of wash buffer (50 mM NaPi, 300 mM NaCl, pH 7.4). Bound protein was eluted with 10 mL of elution buffer (50 mM NaPi, 300 mM NaCl, 300 mM imidazole, pH 7.4). Enzyme purity was assessed by SDS-PAGE. Fractions with >90% purity were collected and dialyzed at 4 °C against 5 mM Tris-SO_4^{2-}, pH 7.8.

Differential Scanning Calorimetry

DSC was performed on a Perkin Elmer Pyris1 calorimeter. Temperature calibration was performed using indium. Samples (40 μL) were sealed in stainless steel pans, equilibrated at 40 °C, and run from 40 °C to 90 °C at a scan rate of 10 °C/min. Following the initial run, samples were cooled back to 40 °C and a second scan was initiated to investigate reversibility. Protein concentrations were 10 – 13 mg/mL. The reference compartment contained an empty sealed stainless steel pan. The raw data (mW) were corrected for instrument baseline by subtracting the air versus air scan. The data were converted to heat capacity by dividing by the scan rate (0.167 °C/s) and the number of moles of protein in each sample yielding data with units of kcal/mol/K.

Preparation of Supported Dehalogenase Biocatalyst

PEI alumina support was first rinsed with distilled water to remove finely divided materials, degassed, and then treated for 2 hours at room temperature with 25% (w/v) freshly

thawed aqueous glutaraldehyde. Excess glutaraldehyde was removed from the support by exhaustive washing with distilled water. The support then was rinsed in 10 mM potassium phosphate/10 μM EDTA buffer (pH 7.0) prior to enzyme loading. In preparation for immobilization, purified dehalogenases were exchanged into 10 mM potassium phosphate/10 μM EDTA buffer (pH 7.0) and brought to a concentration of ~10 mg/mL using an Amicon YM-10 membrane. Protein was offered in excess to the activated alumina at a ratio of 300 mg protein/g support. Loading was conducted for 16 h with gentle agitation at room temperature. Unbound protein was decanted and the support was loaded into the reactor. Enzyme loading was estimated by measuring the difference in A_{280} of the enzyme solution before and after the loading procedure.

Immobilized Enzyme Reactor

A stainless steel 30-cm "shell and tube" reactor was fashioned from 3/8 in and 3/4 in 316 SS and Suage-Lock fittings. Reactor inlet and outlet tubing consisted of 1/16 in SS HPLC tubing fitted with 3-way switching valves for analytical sampling. A circulating water bath was used to maintain constant temperature. Reactor feed consisted of degassed 100 mM potassium phosphate buffer, pH 7.0 saturated with TCP (~8 mM), and maintained at ~5 °C above the operating temperature of the reactor. Feed was delivered continuously with an Alltech 301 HPLC pump. The supported enzyme suspension was packed into the reactor by gravity. Excess reactor volume was made up using 2-mm glass beads. The reactor feed was delivered in the upflow direction to minimize channeling of the bed.

Analysis of TCP and DCH

Gas chromatographic analysis of reactor streams for TCP and DCH was conducted using an HP 5890 gas chromatograph equipped with an HP 7673 autosampler. A 30-meter megabore (0.53 mm), 1.0 micron film, DB-Wax column (J & W Scientific) was used in splitless mode (1/50) with a linear flow injector. Helium was used as carrier gas with FID detection. Quantitation of TCP and DCH was performed by the internal standard method, using two separate internal standards. 1,1,1,2-Tetrachloroethane and 3-chloro-1-propanol were used as internal standards for TCP and for DCH, respectively. For analysis, aqueous samples from the reactor were transferred to capped tubes, saturated with sodium sulfate, and extracted with two volumes of chloroform containing the two internal standards at 10 mM. The chloroform layer was dried with sodium sulfate prior to GC analysis.

Determination of Enzyme Reactor Productivity

In this continuous enzyme reactor system, productivity varies significantly with overall conversion of TCP to DCH. In order to compare enzymes, the productivity for each was determined at constant conversion. Initial reactor productivity was first established by adjusting the feed rate to provide 50% conversion of TCP to DCH. Productivity was continuously monitored with time by decreasing the feed rate in order to reestablish 50% conversion as the biocatalyst activity decays. Data were collected until approximately 10% original activity was reached, and the half-life was determined by fitting with an exponential function.

Acknowledgements

We wish to thank Dr. Ramy Farid, Rutgers University, for performing and interpreting the DSC experiments, Jonathan Eads (Diversa Corp.) for the analysis of the three dimensional structure of the Rhodococcus enzyme, and Dr. Mark Burk (Diversa Corp.) for critically reviewing the manuscript.

References

[1] J. Affhoter, P. Swanson, H. Kan, R. Richard (Dow Chemical Company), *US Patent* WP 9836 080, **1998**.
[2] J. Damborsky, M. G. Nyandoroh, M. Nemec, I. Holoubek, A. T. Bull, D. J. Hardman, *Biotechnol. Appl. Biochem.* **1997,** *26* (Pt 1), 19–25.
[3] P. E. Swanson, *Curr. Opin. Biotechnol.* **1999,** *10*, 365–369.
[4] D. B. Janssen, F. Pries, J. R. van der Ploeg, *Annu. Rev. Microbiol.* **1994,** *48*, 163–191.
[5] J. F. Schindler, P. A. Naranjo, D. A. Honaberger, C. H. Chang, J. R. Brainard, L. A. Vanderberg, C. J. Unkefer, *Biochemistry* **1999,** *38*, 5772–5778.
[6] J. Newman, T. S. Peat, R. Richard, L. Kan, P. E. Swanson, J. A. Affholter, I. H. Holmes, J. F. Schindler, C. J. Unkefer, T. C. Terwilliger, *Biochemistry* **1999,** *38*, 16105–16114.
[7] J. Marek, J. Vevodova, I. K. Smatanova, Y. Nagata, L. A. Svensson, J. Newman, M. Takagi, J. Damborsky, *Biochemistry* **2000,** *39*, 14082–14086.
[8] S. M. Franken, H. J. Rozeboom, K. H. Kalk, B. W. Dijkstra, *Embo J.* **1991,** *10*, 1297–1302.
[9] F. Pries, A. J. van den Wijngaard, R. Bos, M. Pentenga, D. B. Janssen, *J. Biol. Chem.* **1994,** *269*, 17490–17494.
[10] F. Pries, J. Kingma, G. H. Krooshof, C. M. Jeronimus-Stratingh, A. P. Bruins, D. B. Janssen, *J. Biol. Chem.* **1995,** *270*, 10405–10411.
[11] J. P. Schanstra, I. S. Ridder, G. J. Heimeriks, R. Rink, G. J. Poelarends, K. H. Kalk, B. W. Dijkstra, D. B. Janssen, *Biochemistry* **1996,** *35*, 13186–13195.
[12] J. P. Schanstra, J. Kingma, D. B. Janssen, *J. Biol. Chem.* **1996,** *271*, 14747–14753.
[13] G. H. Krooshof, E. M. Kwant, J. Damborsky, J. Koca, D. B. Janssen, *Biochemistry* **1997,** *36*, 9571–9580.
[14] G. H. Krooshof, I. S. Ridder, A. Tepper, G. J. Vos, H. J. Rozeboom, K. H. Kalk, B. W. Dijkstra, D. B. Janssen, *Biochemistry* **1998,** *37*, 15013–15023.
[15] G. H. Krooshof, R. Floris, A. W. Tepper, D. B. Janssen, *Protein Sci.* **1999,** *8*, 355–360.
[16] T. Bosma, E. Kruizinga, E. J. de Bruin, G. J. Poelarends, D. B. Janssen, *Appl. Environ. Microbiol.* **1999,** *65*, 4575–4581.
[17] J. G. Zeikus, C. Vieille, A. Savchenko, *Extremophiles* **1998,** *2*, 179–183.
[18] B. Van den Burg, G. Vriend, O. R. Veltman, G. Venema, V. G. Eijsink, *Proc. Natl. Acad. Sci. USA* **1998,** *95*, 2056–2060.

[19] F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **1998,** *95*, 2035–2036.
[20] C. Vetriani, D. L. Maeder, N. Tolliday, K. S. Yip, T. J. Stillman, K. L. Britton, D. W. Rice, H. H. Klump, F. T. Robb, *Proc. Natl. Acad. Sci. USA* **1998,** *95*, 12300–12305.
[21] M. Lehmann, L. Pasamontes, S. F. Lassen, M. Wyss, *Biochim. Biophys. Acta* **2000,** *1543*, 408–415.
[22] C. Vieille, G. J. Zeikus, *Microbiol. Mol. Biol. Rev.* **2001,** *65*, 1–43.
[23] J. M. Short (Diversa Corporation), *US Patent* 6,171,820, **2001**.
[24] T. M. Florence, Y. J. Farrar, *Anal. Chem. Acta* **1971,** *54*, 373–377.
[25] O. Kuchner, F. H. Arnold, *Trends Biotechnol.* **1997,** *15*, 523–530.
[26] L. Giver, A. Gershenson, P. O. Freskgard, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **1998,** *95*, 12809–12813.
[27] S. Oue, A. Okamoto, T. Yano, H. Kagamiyama, *J. Biol. Chem.* **1999,** *274*, 2344–2349.
[28] K. Chen, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **1993,** *90*, 5618–5622.
[29] D. E. Anderson, J. H. Hurley, H. Nicholson, W. A. Baase, B. W. Matthews, *Protein Sci.* **1993,** *2*, 1285–1290.
[30] J. H. Hurley, W. A. Baase, B. W. Matthews, *J. Mol. Biol.* **1992,** *224*, 1143–1159.
[31] K. L. Britton, P. J. Baker, K. M. Borges, P. C. Engel, A. Pasquo, D. W. Rice, F. T. Robb, R. Scandurra, T. J. Stillman, K. S. Yip, *Eur. J. Biochem.* **1995,** *229*, 688–695.
[32] R. J. Russell, D. W. Hough, M. J. Danson, G. L. Taylor, *Structure* **1994,** *2*, 1157–1167.
[33] J. J. Tanner, R. M. Hecht, K. L. Krause, *Biochemistry* **1996,** *35*, 2597–2609.
[34] R. Thapar, E. M. Nicholson, P. Rajagopal, E. B. Waygood, J. M. Scholtz, R. E. Klevit, *Biochemistry* **1996,** *35*, 11268–11277.
[35] L. Serrano, A. R. Fersht, *Nature* **1989,** *342*, 296–299.
[36] M. D. Bruch, M. M. Dhingra, L. M. Gierasch, *Proteins* **1991,** *10*, 130–139.
[37] P. C. Lyu, M. I. Liff, L. A. Marky, N. R. Kallenbach, *Science* **1990,** *250*, 669–673.
[38] A. Chakrabartty, A. J. Doig, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1993,** *90*, 11332–11336.
[39] B. Forood, E. J. Feliciano, K. P. Nambiar, *Proc. Natl. Acad. Sci. USA* **1993,** *90*, 838–842.
[40] L. Piela, G. Nemethy, H. A. Scheraga, *Biopolymers* **1987,** *26*, 1273–1286.
[41] S. Padmanabhan, S. Marqusee, T. Ridgeway, T. M. Laue, R. L. Baldwin, *Nature* **1990,** *344*, 268–270.
[42] M. A. Roseman, *J. Mol. Biol.* **1988,** *200*, 513–522.
[43] J. A. Wells, *Biochemistry* **1990,** *29*, 8509–8517.
[44] V. Munoz, F. J. Blanco, L. Serrano, *Nat. Struct. Biol.* **1995,** *2*, 380–385.
[45] N. Guex, M. C. Peitsch, *Electrophoresis* **1997,** *18*, 2714–2723.