View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by CiteSeerX

JMB

Directed Evolution Study of Temperature Adaptation in a Psychrophilic Enzyme

Kentaro Miyazaki[1], Patrick L. Wintrode[1], Rowan A. Grayling[2] Donn N. Rubingh[2] and Frances H. Arnold[1]*

[1]*Division of Chemistry and Chemical Engineering 210-41 California Institute of Technology, Pasadena, CA 91125, USA*

[2]*Miami Valley Laboratories The Procter & Gamble Company, PO Box 538707 Cincinnati, OH 45253, USA*

**Corresponding author*

We have used laboratory evolution methods to enhance the thermostability and activity of the psychrophilic protease subtilisin S41, with the goal of investigating the mechanisms by which this enzyme can adapt to different selection pressures. A combined strategy of random mutagenesis, saturation mutagenesis and *in vitro* recombination (DNA shuffling) was used to generate mutant libraries, which were screened to identify enzymes that acquired greater thermostability without sacrificing low-temperature activity. The half-life of seven-amino acid substitution variant 3-2G7 at 60 °C is ~500 times that of wild-type and far surpasses those of homologous mesophilic subtilisins. The dependence of half-life on calcium concentration indicates that enhanced calcium binding is largely responsible for the increased stability. The temperature optimum of the activity of 3-2G7 is shifted upward by ~10 °C. Unlike natural thermophilic enzymes, however, the activity of 3-2G7 at low temperatures was not compromised. The catalytic efficiency, k_{cat}/K_M, was enhanced ~threefold over a wide temperature range (10 to 60 °C). The activation energy for catalysis, determined by the temperature dependence of k_{cat}/K_M in the range 15 to 35 °C, is nearly identical to wild-type and close to half that of its highly similar mesophilic homolog, subtilisin SSII, indicating that the evolved S41 enzyme retained its psychrophilic character in spite of its dramatically increased thermostability. These results demonstrate that it is possible to increase activity at low temperatures and stability at high temperatures simultaneously. The fact that enzymes displaying both properties are not found in nature most likely reflects the effects of evolution, rather than any intrinsic physical-chemical limitations on proteins.

© 2000 Academic Press

Keywords: adaptation; directed evolution; psychrophile; subtilisin; thermostability

Introduction

Analyses of sequence-structure-function relationships in homologous proteins from organisms adapted to different temperatures have yielded important insights into the molecular basis of protein adaptation. Based on studies of this kind, a number of temperature adaptation mechanisms have been proposed (Russel & Taylor, 1995; Fields & Somero, 1998; Jaenicke & Bøhm, 1998; Zavodszky *et al*., 1998). Comparative studies of evolutionarily related proteins are limited, however, by the many neutral mutations that will generally have arisen during divergent evolution. It is often extremely difficult to identify adaptive mutations against this large background of neutral ones. These studies are further complicated by the fact that it is not always clear which properties are directly subjected to selective pressure. During evolution, a property not under selective pressure may change as a result of random drift, in which

Present address: K. Miyazaki, National Institute of Bioscience and Human Technology, 1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan, E-mail, miyaken@nibh.go.jp

Abbreviations used: CD, circular dichroism; HEPPS, *N*-2-hydroxyethylpiperazine-*N*′-3-propane sulfonic acid; StEP, staggered extension process; s-AAPF-*p*Na, succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide; T_{opt}, optimum temperature.

E-mail address of the corresponding author: frances@cheme.caltech.edu

0022-2836/00/041015–12 $35.00/0

© 2000 Academic Press

case comparative studies will yield little information on adaptive mechanisms. A property not under selective pressure may also change because it is coupled to a selected property (Benner, 1989). These confounding effects severely inhibit our ability to uncover the molecular basis for protein adaptation.

One way to discover sequence (or structural) changes that lead to specific functional adaptation is to use directed evolution (mutation, recombination, and screening *in vitro*) to evolve specific protein functions under highly controlled conditions. This approach generates adaptive mutations, thus avoiding the complications arising from neutral evolution. In addition, directed evolution allows us to choose which properties will be subjected to selective pressure and therefore to test the extent to which different properties are coupled. We can also apply combinations of selective pressures not normally encountered in nature, such as the simultaneous requirement for both stability at high temperatures and high activity at low temperatures (Giver *et al.*, 1998). Biochemical and structural analysis of the evolved proteins will allow us to uncover the molecular mechanisms by which the new properties have been acquired. Researchers have used directed evolution to evolve a wide variety of enzyme properties, in the process discovering mutations that would not have been suggested by rational approaches (Zhang *et al.*, 1997; Giver *et al.*, 1998; Oue *et al.*, 1999; Spiller *et al.*, 1999).

Extensive research has been done on the subtilisin family of serine proteases (EC 3.4.21.14). Over 170 sequences have been reported (Siezen & Leunissen, 1997) and several three-dimensional structures are available at high resolution (e.g. Carlsberg, Bode *et al.*, 1987; BPN′, McPhalen & James, 1988; thermitase, Teplyakov *et al.*, 1990; proteinase K, Betzel *et al.*, 1990; Savinase, Kuhn *et al.*, 1998). Comparative studies of subtilisins have revealed that 22 residues are conserved in the substrate binding pocket, including the perfectly conserved catalytic triad (Asp, His, Ser) and an oxyanion hole (Asn) (Siezen & Leunissen, 1997). As is true for a number of enzymes, a single polypeptide fold has been able to adapt to widely divergent environments. Indeed, subtilisins have been isolated from psychrophilic (Davail *et al.*, 1992; Narinx *et al.*, 1992), mesophilic (Wells *et al.*, 1983; Stahl & Ferrari, 1984; Jacobs *et al.*, 1985; Wati *et al.*, 1997), and extreme thermophilic organisms (Völkl *et al.*, 1994; Voorhorst *et al.*, 1996).

Subtilisin S41 from the Antarctic *Bacillus* TA41 displays the typical properties of cold-adapted enzymes: high catalytic efficiency at low temperatures and relative instability at high temperatures (Davail *et al.*, 1994). The enzyme shows high sequence similarity to other subtilisins from psychrophilic (S39, Narinx *et al.*, 1992), mesophilic (BPN′, Wells *et al.*, 1983; E, Stahl & Ferrari, 1984; Carlsberg, Jacobs *et al.*, 1985; SSII, Wati *et al.*, 1997) and thermophilic (thermitase, Meloun *et al.*, 1985) sources (Figure 1). Although S41 shares the common subtilisin fold, it exhibits some features believed specific to psychrophilic proteins (Davail *et al.*, 1994). Its surface is very rich in hydrophilic residues, particularly Asp. The enzyme contains several extended surface loops not generally found in mesophilic subtilisins. It also lacks several salt bridges and aromatic-aromatic interactions that are commonly found in other subtilisins. However, the closely related subtilisin SSII from the mesophilic *Bacillus sphaericus* (Wati *et al.*, 1997) shares the same high Asp content and extended loops (Figure 1). The thermostability and low temperature activity of SSII are comparable to those of other mesophilic subtilisins. Thus, these features alone cannot account for the behavior of the psychrophilic enzyme.

We have used directed evolution to probe the extent to which we can increase the thermostability of subtilisin S41 without damaging its catalytic activity at low temperature. By limiting the evolutionary process to incorporating point mutations rather than insertions or deletions, we can investigate the adaptability of the S41 framework, including its extended loops. By requiring high activity at low temperatures, we can distinguish between what is physically possible (an enzyme that is both highly thermostable and highly active at low temperature) and what may be biologically relevant (enzymes with one, but not both properties). In addition, we wished to examine whether stabilizing amino acid substitutions discovered by directed evolution parallel amino acid changes found in naturally more stable subtilisins such as SSII. Finally, we wished to know how many such mutations are required to make the conversion to an enzyme as stable as its natural mesophilic homologs.

Results and Discussion

Directed evolution of thermostable subtilisin S41

Directed evolution was used to increase the thermostability of subtilisin S41. A variant with improved stability was only accepted, however, if its activity at 25 °C was equal to or greater than that of wild-type. This requirement was intended to ensure the preservation of high activity in the evolved enzymes, and thus determine the extent to which a cold-active enzyme can also be thermostable. Subtilisins, including S41, possess a high affinity (K_d, $\sim 10^{-6}$ to 10^{-10} M) Ca^{2+} binding site (Bryan *et al.*, 1992; Davail *et al.*, 1994) and are reversibly inactivated when this site is not occupied (Davail *et al.*, 1994). They also possess a low affinity (e.g. BPN′, K_d, ~30 mM, Pantoliano *et al.*, 1988) Ca^{2+} binding site (Siezen *et al.*, 1991). Ca^{2+} binding at this site strongly influences stability, but is not required for activity. Preliminary characterization of S41 showed that its half-life at 60 °C is highly Ca^{2+} dependent. Accordingly, S41 mutant libraries were screened at relatively low Ca^{2+} con-

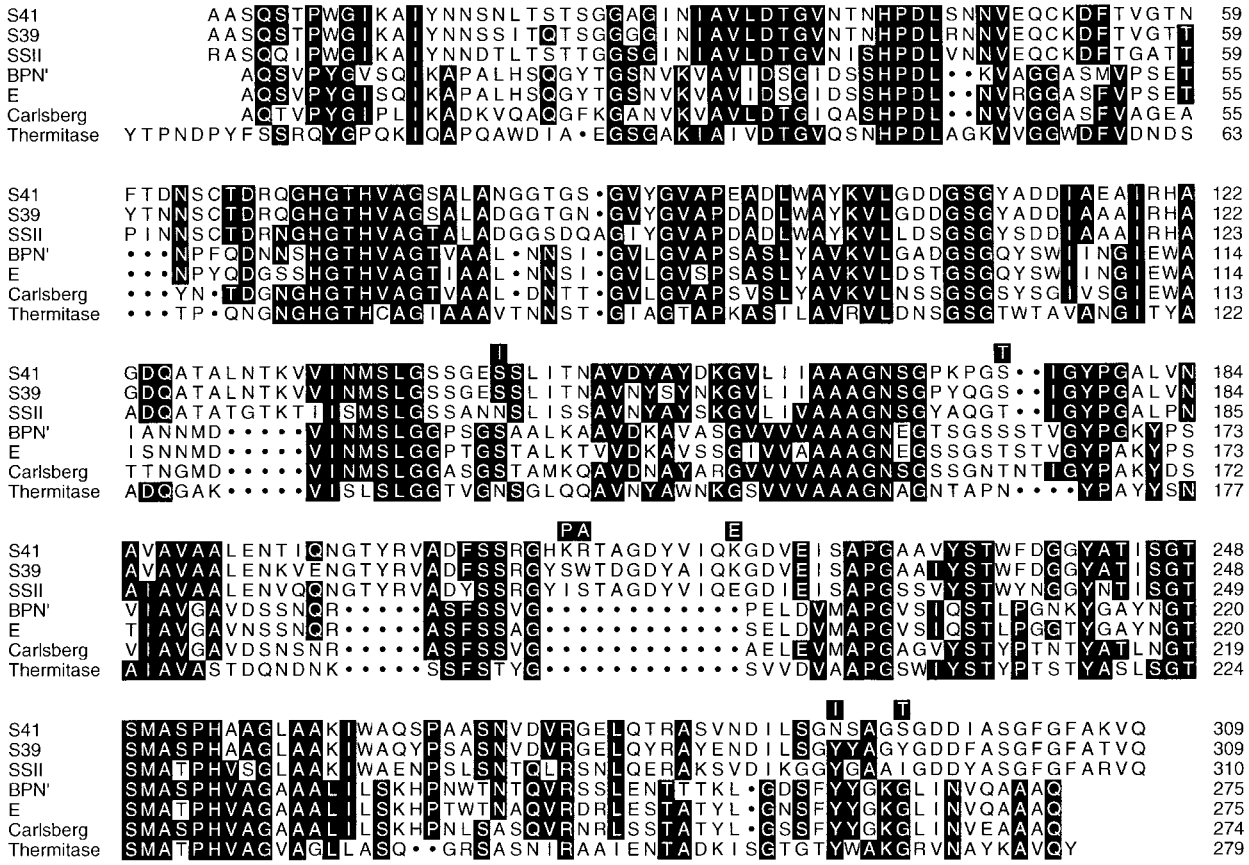

Figure 1. Amino acid sequence alignment of subtilisins S41 (Davail *et al*., 1992), S39 (Narinx *et al*., 1992), SSII (Wati *et al*., 1997), BPN′ (Wells *et al*., 1983), E (Stahl & Ferrari, 1984), Carlsberg (Jacobs *et al*., 1985), and thermitase (Meloun *et al*., 1985). Multiple alignment was carried out using a CLUSTAL W (Thompson *et al*., 1994). Residues found frequently (>50 %) at given positions are shaded. Mutations in 3-2G7 are labeled with filled squares. Gaps in the sequence are shown with (●).

centrations (10 mM), to offer S41 the opportunity for stabilization by increasing its affinity for calcium.

The first generation library was created by error-prone PCR of wild-type S41 under conditions adjusted to generate an average of two to three base-pair substitutions per gene. As reported previously (Miyazaki & Arnold, 1999), nine thermostable S41 variants were identified by comparing the activities of 864 clones before and after incubation at 60 °C. Three contained a substitution at amino acid position 211; position 212 was changed in a fourth. In order to explore what further stabilization was possible beyond that accessible by single base mutations, residues at these two positions were randomized by saturation mutagenesis. Of 1536 clones screened, we identified one (1-14A7) ~ten times more stable than wild-type and whose sequence (Pro211-Ala212) was very different from that of wild-type (Lys211-Arg212) (Miyazaki & Arnold, 1999). Thermostable S41 variants and the lineage during directed evolution are summarized in Figure 2.

Advantageous mutations identified in the first generation were recombined *in vitro* using the StEP (staggered extension process) method (Zhao *et al*., 1998). The probability of finding any particular combination in a random recombination library of six different mutations is extremely low, $1/6^6 = 1/46{,}656$ (Moore *et al*., 1997). To reduce the screening requirements, recombination was performed in two separate experiments, making use of a unique *Msc*I restriction site at nucleotide position 624. This strategy reduces the combinatorial possibilities to only $3^3 = 27$ for each experiment. One recombination experiment involved the mutations at nucleotides 43, 434, and 523; the second recombined genes with mutations at nucleotides 631-636, 661, and 883. A total of 288 clones in each library were assayed before and after incubation at 70 °C, to yield thermostable variants 2-1D2 and 2-1G11. DNA sequencing revealed that 2-1D2 picked up non-synonymous mutations from two parents (Ser145Ile and Ser175Thr), but not the mutation (Asn15Asp) from the third parent. Two newly generated non-synonymous mutations were also found: Ser252Ala and Val272Ile. Variant 2-1G11 contained all the mutations from its three parents. These two separately optimized genes were recombined by restriction digestion with *Msc*I, followed

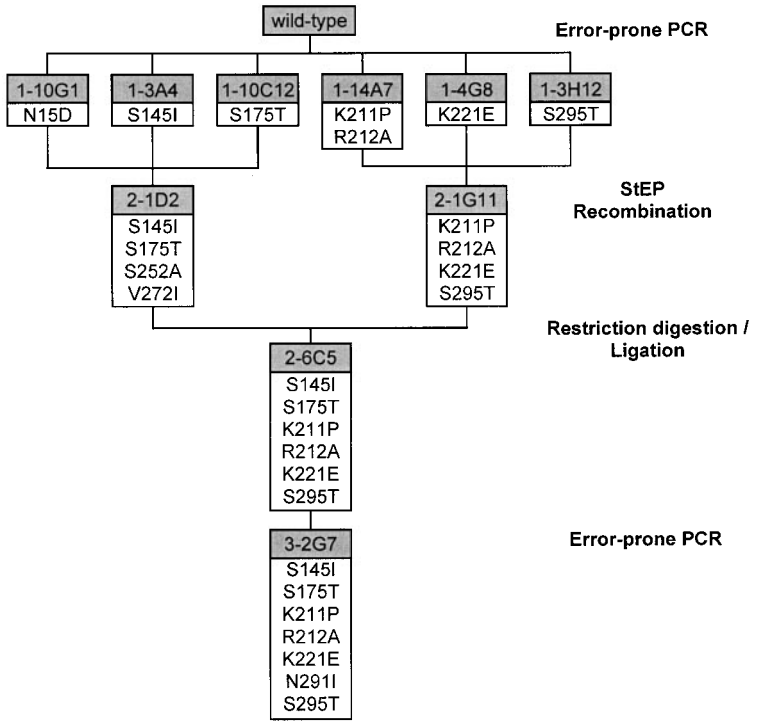

Figure 2. Lineage of subtilisin S41 variants. Variants in the first generation were obtained previously (Miyazaki & Arnold, 1999). The seven mutations obtained from the first generation were recombined by StEP and the best variant, 2-6C5, was obtained in the second generation library. It was parented to create third generation library by error-prone PCR. Only non-synonymous mutations are shown.

by fragment exchange. The *Bsu*36I-*Msc*I gene fragment from 2-1D2 and *Msc*I-*Bam*HI fragment from 2-1G11 were combined to yield 2-6C5. This process eliminated the newly generated (and possibly beneficial) mutations in 2-1D2. The resulting 2-6C5 variant possessed six of the seven amino acid substitutions that were found individually to stabilize wild-type S41, demonstrating that these mutations make cumulative contributions to thermostability.

A third generation library was created by error-prone PCR of 2-6C5. Three thermostable variants were identified upon screening 960 clones. Surprisingly, all three were found to contain different amino acid substitutions at residue 291. Of the three, the Asn291Ile substitution in 3-2G7 gave the highest residual activity upon incubation at 70 °C. Subtilisin S41 variant 3-2G7 contained seven amino acid substitutions, all of which contributed to its increased thermostability.

Thermostability as measured by half-life of autolytic inactivation

The parameter often used to measure subtilisin stability is the rate of irreversible inactivation caused by autolysis at elevated temperature (Voordouw *et al.*, 1976). We compared the half-lives of wild-type S41, 3-2G7, SSII, and BPN′ at 60 °C at different concentrations of $CaCl_2$ (1 or 10 mM) and NaCl (0 or 2 M) (Table 1). Variant 3-2G7 was much more stable than wild-type S41, especially at the low calcium concentration. In the absence of NaCl, 500 and 60-fold increases were obtained in 1 and 10 mM $CaCl_2$, respectively. In 2 M NaCl, 320 and 12-fold increases were obtained. The stability of 3-2G7 far surpassed those of mesophilic subtilisins SSII and BPN′. The stability of 3-2G7 was also much less calcium-dependent in the range 1 to 10 mM. These results suggest an increased affinity for Ca^{2+} in the evolved enzymes. The low affinity of S41 for calcium, relative to its more thermostable counterparts, is known to render the enzyme fragile, particularly at low Ca^{2+} concentrations (Davail *et al.*, 1994).

Calcium dependence of the rate of thermal inactivation

To probe further the role of calcium in stabilization, the half-lives of wild-type S41 and 3-2G7 at 60 °C were determined as functions of $CaCl_2$ concentration (Figure 3). The influence of NaCl was also examined. As expected, increasing the concentration of $CaCl_2$ decreased the rates of inactivation of both enzymes. The dependence of half-life on $CaCl_2$ concentration is sigmoidal. Assuming that the shift in the curve reflects a difference in calcium affinity, we see that 3-2G7 binds Ca^{2+} more tightly than wild-type. In 0 M NaCl, the midpoints occurred at 52 mM for wild-type, *versus* 574 μM for 3-2G7, a nearly 100-fold difference in apparent calcium affinity. The curves for both proteins shifted to higher concentrations in the presence of 2 M NaCl, which likely reflects competition of sodium ions for the binding site, as was observed for BPN′ (Pantoliano *et al.*, 1988). The stability of 3-2G7 in excess calcium (half-life, 602 minutes) is ten times that of wild-type (half-life, 63 minutes),

Table 1. Half-lives of inactivation of wild-type subtilisin S41, variant 3-2G7, and mesophilic subtilisins SSII and BPN′

	1 mM $CaCl_2$		10 mM $CaCl_2$	
Enzyme	0 M NaCl	2 M NaCl	0 M NaCl	2 M NaCl
Wild-type	0.9	3.9	9.3	106
3-2G7	449	1240	566	1303
SSII	34.2	34.5	35.8	147
BPN′	28.0	1201	46.6	897

Enzymes (2 μM) in 50 mM HEPPS-NaOH (pH 8.5) and indicated concentrations of $CaCl_2$ and NaCl were incubated at 60 °C. Residual activity was determined at appropriate intervals in 0.1 M Tris-HCl (pH 8.5), 10 mM $CaCl_2$, 0.2 mM s-AAPF-*p*Na at 30 °C. Values (in minutes) are the averages of three experiments and the standard errors are less than ± 8 %.

corresponding to a difference in free energy of 1.5 kcal/mol.

Similar stabilizing effects of Ca^{2+} and Na^{+} have also been reported for subtilisin BPN′ (Voordouw *et al*., 1976; Pantoliano *et al*., 1988). Calcium binding at the weak site that includes Asp197 (which corresponds to Glu215 in S41) increases thermostability (Pantoliano *et al*., 1988). This site has a K_d for calcium of ~30 mM and, in the absence of high calcium concentrations, will bind sodium or potassium. Based on the sequence homology and the retention of the predicted ligands to Ca^{2+} in S41 (Figure 1), we can predict that this weak Ca^{2+} binding site also exists in S41. The equivalent ligands in S41 are Glu144 (mutation of the corresponding Gly131 to Asp in BPN′ enhances calcium binding, Pantoliano *et al*., 1988), Ala181 (Lys170 in BPN′), Val183 (the corresponding Pro172 to Asp mutation in BPN′ enhances calcium binding, Pantoliano *et al*., 1988), Asp223 (Glu195 in BPN′), Glu225 (Asp197 in BPN′), Arg275 (Arg247 in BPN′), and Gln279 (Glu251 in BPN′).

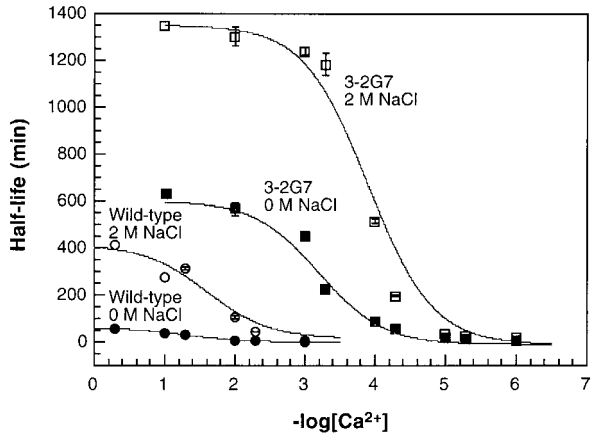

Figure 3. Calcium ion titration of the rate of inactivation at 60 °C. Half-life for thermal inactivation of wild-type S41 and 3-2G7 in the presence of 0 M or 2 M NaCl is plotted as a function of the negative log[Ca^{2+}]. Symbols: (●) wild-type, 0 M NaCl; (○) wild-type, 2 M NaCl; (■) 3-2G7, 0 M NaCl; (□) 3-2G7, 2 M NaCl. The reported half-lives are the averages for two independent experiments, and the error bars represent the range of values. The dotted curve drawn through the data is a theoretical curve for single binding site, $pCa = pK_a + \log([E]/[E\text{-}Ca])$ with an apparent $pK_a = 1.3$ (wild-type, 0 M NaCl); 1.6 (wild-type, 2 M NaCl); 3.2 (3-2G7, 0 M NaCl); 3.9 (3-2G7, 2 M NaCl).

Ion binding constants of psychrophilic proteins are often lower than the corresponding constants of mesophilic homologs (Davail *et al*., 1994; Feller *et al*., 1994). The binding of metal ions, particularly calcium, can provide stabilization superior to most other, weak interactions. Stabilization of proteins by Ca^{2+} binding is well documented and involves unusually high affinity or even extra Ca^{2+} binding sites (Mitani *et al*., 1986; Teplyakov *et al*., 1990). Directed evolution also discovered this stabilization mechanism for S41. Although the contribution of each individual mutation may be small, the net result is an impressive 100-fold increase in affinity. That recombination of the individual mutations increased stability further indicates that these mutations are cumulative in their contributions to the improved binding and stability.

Thermostability as measured by circular dichroism

Circular dichroism (CD) was used to monitor the loss of S41 secondary structure upon heating. To prevent autolysis, the enzymes were pre-treated with the serine protease inhibitor phenylmethylsulfonylfluoride. Figure 4(a) illustrates the CD spectrum of 3-2G7. The profile of wild-type is essentially identical to that of the variant, indicating that the evolution did not result in significant changes in secondary structure content. The spectrum is typical of an α-helical protein, with a large negative ellipticity at 222 nm.

The inactive wild-type S41 and 3-2G7 were heated from 0 to 90 °C (Figure 4(b)). The profiles are essentially identical, except for the shift in transition temperature ~20 °C upward for the evolved enzyme. This shift and the prolonged half-life upon autolysis clearly demonstrate the increased stability of 3-2G7 at high temperature.

Temperature dependence of the specific activity

The specific activities of wild-type and 3-2G7 were determined over a range of temperatures, from 5 to 80 °C (Figure 5). Although the high specific activity at low temperatures observed in psychrophilic enzymes is generally associated with

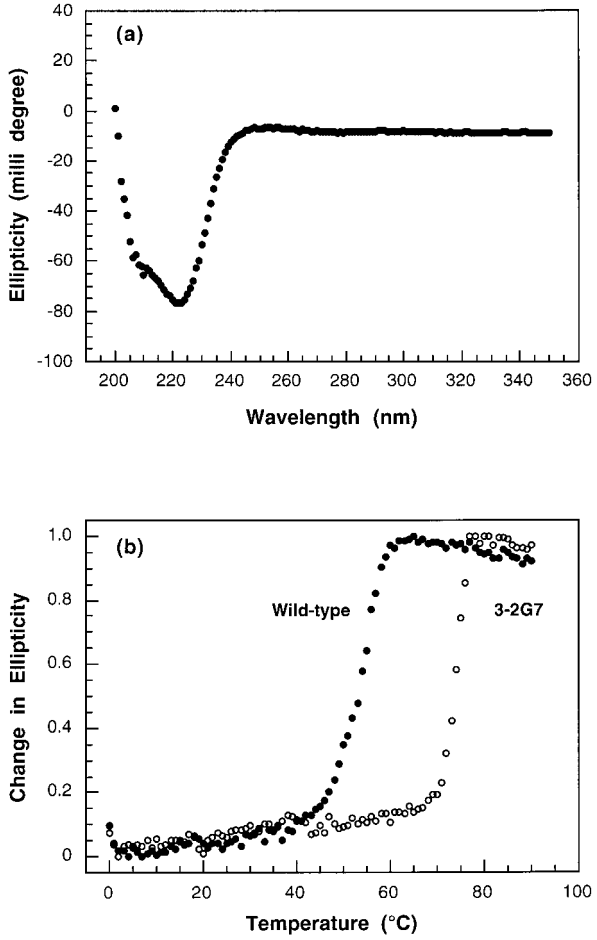

Figure 4. Thermal inactivation profiles of wild-type S41 and 3-2G7, as measured by CD. The enzyme (2 μM) was inactivated with phenylmethylsulfonylfluoride and dialyzed against 50 mM borate-NaOH (pH 8.5), 1 mM $CaCl_2$. (a) CD spectrum of 3-2G7 recorded at 25 °C. (b) The denaturation process monitored by the decrease in ellipticity at 222 nm. Temperature was increased from 0 to 90 °C at a rate ~1.0 deg. C/minute.

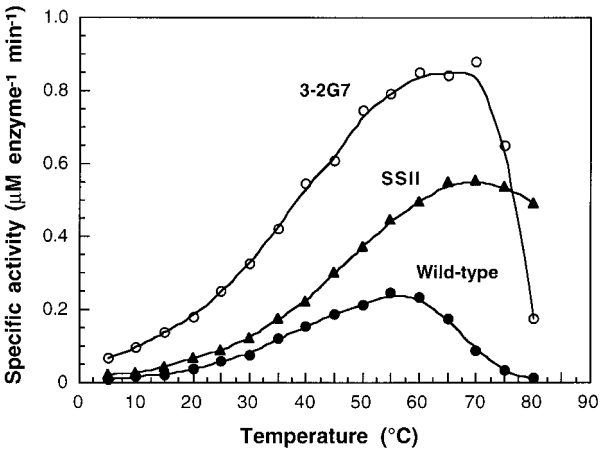

Figure 5. Temperature dependence of specific activities of wild-type subtilisin S41, 3-2G7, and SSII. Symbols: (●), wild-type; (○), 3-2G7; (▲), SSII.

poor thermostability, 3-2G7 combines both cold activity and high stability in a single enzyme. The mutations in fact increased the specific activity ~threefold at all temperatures tested, not just the lower temperature at which the mutants were screened. Stabilization of S41 was accompanied by an increase in the temperature of optimal activity (T_{opt}) of ~10 °C. Similar results were obtained in two recent studies in which directed evolution was used to convert a mesophilic subtilisin E (Zhao & Arnold, 1999) and esterase (Giver *et al.*, 1998) into thermophilic enzymes. In both cases, dramatic increases in both thermostability and activity at all temperatures were observed. The ability of the enzymes to remain folded at higher temperatures translated into upward shifts in T_{opt}. Interestingly, the T_{opt} of 3-2G7 is lower than that of mesophilic SSII (Figure 4), even though its thermostability, as measured by its half-life and its melting temperature, is higher.

Kinetic constants of wild-type and evolved enzymes

Enzyme kinetics were measured at different temperatures and calcium concentrations in order to assess the effect of these parameters on activity. Kinetic constants are summarized in Table 2. The overall catalytic efficiency of 3-2G7, represented by k_{cat}/K_M, far surpassed that of wild-type in the temperature range 10 to 60 °C. Calcium concentration did not influence the kinetics of either enzyme, confirming that Ca^{2+} binding affects stability independently of activity. The increase in k_{cat}/K_M is due primarily to an increase in the intrinsic catalytic rate constant k_{cat}. Thus, although tighter calcium binding increases the resistance of the protein to thermal denaturation, it evidently does not interfere either with the active site geometry or with any protein motions required for catalysis. Variants of BPN′ in which the weak calcium binding site was engineered for higher affinity were also found to retain their catalytic efficiency (Pantoliano *et al.*, 1988).

Preservation of psychrophilic catalysis in 3-2G7

It has been found that psychrophilic enzymes tend to have lower activation energies than their mesophilic counterparts (Low *et al.*, 1973; Simpson & Haard, 1984; Arpigny *et al.*, 1997; Feller *et al.*, 1997). This is not unexpected, since the energy needed to surmount the activation barrier between substrate and product comes from thermal fluctuations (kT), and the magnitude of these fluctuations is reduced at low temperatures. To achieve reaction rates comparable to those of mesophilic

Table 2. Kinetic constants of wild-type subtilisin S41 and evolved variant 3-2G7

Enzyme	Temperature (°C)	[$CaCl_2$] (mM)	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\times 10^5$ M^{-1} s^{-1})
Wild-type	10	1	216	17.3	0.80
		10	282	23.3	0.83
	30	1	192	56.7	2.7
		10	273	63.8	2.0
	60	1	1290	255	1.8
		10	909	264	2.5
3-2G7	10	1	203	51.5	2.5
		10	137	48.2	3.5
	30	1	210	159	6.9
		10	208	179	7.4
	60	1	739	550	6.8
		10	496	553	9.7

Kinetic constants were determined from initial rates at different concentrations of s-AAPF-*p*Na over the range 0.025-1.2 mM, in 50 mM HEPPS-NaOH (pH 8.5), 1 or 10 mM $CaCl_2$. The values are the averages of two experiments and standard errors are less than 10%.

enzymes, enzymes adapted to low temperatures must accomplish a greater reduction of the activation barrier in order to compensate for the reduced thermal energy available. It is appropriate to ask whether our thermostabilized mutant has retained this characteristic. We therefore determined E_a for wild-type S41 and 3-2G7, as well as for the closely related mesophilic homolog SSII. Figure 6 shows the Arrhenius plots for the temperature range 15 to 35 °C. The slopes of the plots are nearly the same for wild-type (7.1 kcal/mol) and 3-2G7 (6.6 kcal/mol). Both have an E_a ~5 kcal/mol lower than that of SSII. These results demonstrate that 3-2G7 preserves the efficient catalysis of the psychrophilic enzyme in spite of its large increase in stability.

Amino acid sequences of evolved subtilisins

Subtilisin S41 differs from SSII at 85 sites (out of 309 overall, 72.5% identity). If we attempt to uncover the molecular basis for the different stabilities of these two proteins, we are faced with the combinatorially insurmountable task of examining the contributions of 85 different mutations and their various combinations. We were able, however, to convert S41 into an enzyme more stable than SSII with only seven amino acid substitutions, demonstrating that the enzyme can adapt rapidly under strong selective pressure. Comparing the sequence of 3-2G7 to those of the homologous subtilisins S39 and SSII, we find that all the stabilizing mutations are located in regions that are variable among the three proteins (Figure 1). Two of the eight mutations, Thr175 and Glu221, are found in the mesophilic homolog SSII.

We can offer two good reasons why the amino acid substitutions found to stabilize a psychrophilic enzyme in directed evolution experiments would not match the mutations that separate it from its more thermostable homologs. One is simply that sequence space is extremely large, and there are probably many roads to thermostabilizing a given enzyme. The second is that the selection pressure applied in the laboratory probably does not resemble the pressures operative during the evolution of the natural enzymes. It is conceivable that the common ancestor to S41 and SSII was also psychrophilic and that SSII evolved under selection for higher thermostability. More likely, however, is that the common ancestor was mesophilic and that S41 experienced pressure to become more cold-active (while thermostability drifted). In this latter case, the differences between the two enzymes would not be expected to shed much light on thermostabilization mechanisms. In addition to the fact that we generally do not know the selection pressures, if any, under which a given enzyme came about, we also do not know how the strength of that selection pressure dictates the solutions. In the

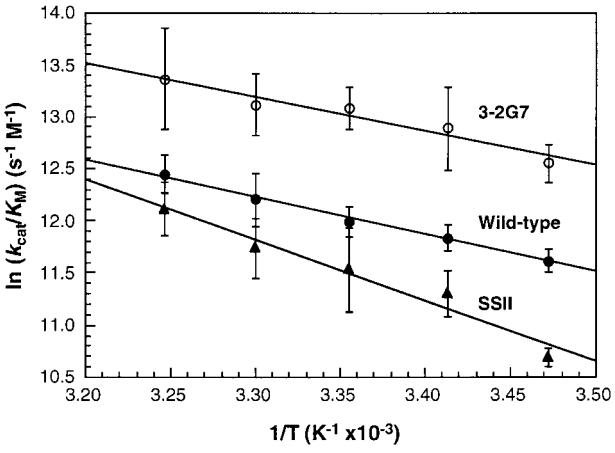

Figure 6. Arrhenius plots for the k_{cat}/K_M of wild-type subtilisin S41, 3-2G7, and SSII. Kinetic constants were determined from initial rates at different concentrations of s-AAPF-*p*Na over the range 0.025-1.2 mM, in 50 mM HEPPS-NaOH (pH 8.5), 10 mM $CaCl_2$. Error bars indicate values from two separate experiments. Symbols: (●), wild-type; (○), 3-2G7; (▲), SSII.

laboratory we may be requiring much larger adaptive changes.

Subtilisin S41 has been noted for its high content of (22) Asp residues. Davail *et al.* (1994) suggested that these Asp residues, which are located mainly in extended surface loops, contribute to the decreased stability of the cold-adapted enzyme by providing a more hydrophilic surface and improved interactions with solvent, which reduces the packing density. However, the mesophilic enzyme SSII also has an unexpectedly large number (19) of Asp residues, most of which are at the same positions as in S41. We observed no substitution of Asp residues during the stabilization of S41. Except for Asn15Asp and Lys221Glu, most mutations were to non-charged residues. It is unclear from the sequence alignment whether these latter mutations affect calcium binding, although the increased negative charge could certainly contribute to enhancing the affinity.

S41 shares high sequence homology with several subtilisins of known structure (BPN′, Wells *et al.*, 1983; Carlsberg, Jacobs *et al.*, 1985; thermitase, Meloun *et al.*, 1985; and Savinase, Betzel *et al.*, 1992). Based on this homology, a three-dimensional model of S41 was constructed which preserves the overall fold shared by other subtilisins (Figure 7). This model was examined for clues to the mechanisms of stabilization in the evolved enzyme. All the mutations, with the sole exception of Asn291Ile, are located on the surface of the enzyme. In addition, with the exception of Ser145-Ile, none occur in elements of regular secondary structure. Specific electrostatic or aromatic interactions observed in BPN′, Carlsberg, and thermitase (Siezen *et al.*, 1991) were not introduced in S41.

Although increased affinity for Ca^{2+} clearly plays an important role in the stabilization of 3-2G7 S41, it is not possible to deduce mechanisms for the observed increase from the structural model. None of the seven mutations occurs in the putative Ca^{2+}-binding sites (Figure 7), which suggests that their effects on Ca^{2+} affinity are through longer-range electrostatic interactions or through subtle structural changes not apparent from the homology model.

It has been suggested that extended loop structures on the molecule surface are responsible for the reduced thermostability of S41 (Davail *et al.*, 1994). Four out of nine thermostable variants in our first random mutant library contained mutations at the beginning of one of the extended loops involving amino acids 210 to 221 (Miyazaki & Arnold, 1999). Moreover, saturation mutagenesis at residues 211 and 212 dramatically improved stability, which indicated that this loop constitutes a weak point of this enzyme, possibly an initiation site for unfolding (Miyazaki & Arnold, 1999). Due to limitations in loop modeling, however, we are unable to identify how the substitutions Lys211Pro and Arg212Ala that occur at the beginning of the extended surface loop (Figure 7) contribute to the large increase in stability. However, it is likely that the introduction of the Pro residue rigidifies the loop by restricting the number of available main-chain conformations. Such stabilization by the introduction of Pro residues into loop regions is a well-documented phenomenon (Matthews, 1993; Watanabe *et al.*, 1994).

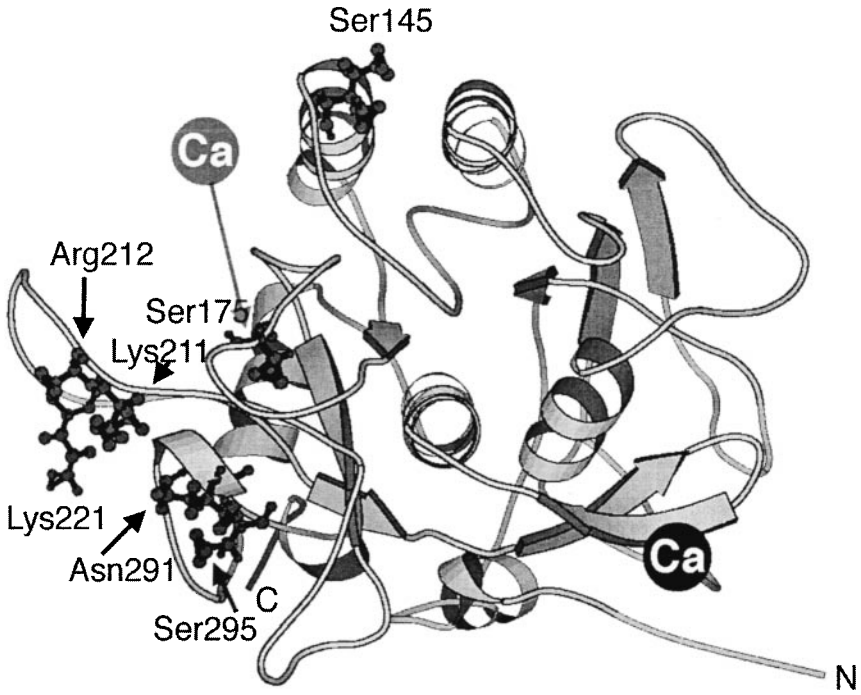

Figure 7. Model of subtilisin S41 constructed based on homology to subtilisins of known structure (see Materials and Methods for details). Seven thermostabilizing mutations identified in variant 3-2G7 are shown. The black Ca indicates the tight Ca^{2+} binding site, and the gray Ca indicates the weak Ca^{2+} binding site, which is located behind the loop containing Ser175. The model was created using MOLSCRIPT (Kraulis, 1991).

Comparison with the site-directed mutagenesis study of homologous psychrophilic subtilisin S39

An effort was made recently to enhance the thermostability of the homologous psychrophilic subtilisin S39 by site-directed mutagenesis (Narinx *et al.*, 1997). The authors created five variants: Arg68Cys, Thr85Asp, Asp106Cys, His121Trp, Ala181Lys/Ser211Glu (reported as Ser210Glu). Among these, only the residue at 211 differs in S41 (Lys211). Considering the striking amino acid sequence identity with S41, 88.3% with neither gaps nor insertions, and shared biochemical properties (Davail *et al.*, 1992; Narinx *et al.*, 1992), the results from S39 should also be applicable to S41. The rationale behind the site-directed mutagenesis experiments was to (a) form a disulfide bond between residues 68 and 106 (Arg68Cys and Asp106Cys), (b) introduce an aromatic interaction in the vicinity of residue 121 (His121Trp), (c) form a salt bridge between residues 181 and 211 (Ala181Lys/Ser211-Glu), and (d) increase calcium binding affinity (Thr85Asp). Variants Arg68Cys and Asp106Cys were not secreted into the supernatant (implying that these mutations interfered with proper folding or function). Variants His121Trp and Ala181Lys/Ser211Glu were less stable than wild-type. Only variant Thr85Asp, which had already been known to increase the stability in S39 (Feller *et al.*, 1996), was effective. None of these mutations was discovered in our experiments. Furthermore, introducing the Thr85Asp substitution into S41 by site-directed mutagenesis did not increase stability, but instead shortened the half-life threefold (at 60 °C in the presence of 10 mM $CaCl_2$) (K.M., unpublished).

Some thoughts on adaptation

It has long been noted that enzymes that are stable at moderate to high temperatures tend to show poor activity at lower temperatures, relative to their cousins that have adapted specifically to colder environments. This has led to speculation that the two properties, stability at elevated temperatures and efficient activity at low temperatures, place mutually exclusive demands on enzymes. Results presented here and in other recent studies (Van den Burg *et al.*, 1998) demonstrate that low temperature activity and thermostability can be improved simultaneously. Enzymes have usually not evolved to a point where such tradeoffs are unavoidable. However, it remains a fact that such enzymes are not generally found in nature (although there are important exceptions, particularly for enzymes that act upon thermolabile substrates). Our work strongly suggests that this fact has its basis in evolution rather than physical chemistry.

Psychrophilic organisms typically live out their lives at low temperatures, just as thermophilic organisms always experience high temperatures. Enzymes in a psychrophilic organism are under no selective pressure to be stable at elevated temperatures. A property not fixed by selection will eventually vanish due to random genetic drift, unless it is coupled to another property that is under selective pressure. It is therefore entirely reasonable and consistent with the results of these protein engineering studies to suppose that psychrophilic enzymes lost thermostability as a result of random drift during the course of divergent evolution from their mesophilic and thermophilic cousins.

A similar argument can be made for the poor low-temperature activity of thermophilic enzymes. The intrinsic rates of chemical reactions are significantly enhanced at high temperatures, and in this sense, thermophilic enzymes are faced with a less formidable catalytic task than their psychrophilic counterparts. One can argue that thermophilic enzymes are subject to less stringent selective pressure on activity than mesophiles or psychrophiles: the degree of optimization sufficient to achieve the biologically necessary reaction rates at high temperatures may be insufficient at low temperatures. Just as one would not expect a psychrophilic enzyme to evolve (or maintain) the additional, unnecessary thermostability required at high temperatures, one would not expect thermophilic enzymes to exhibit the additional catalytic efficiency required at low temperatures.

Of course, adaptation of one property accompanied by genetic drift may not be the only explanation for the stability-activity relationships observed among proteins adapted to different temperatures. It is also possible, for example, that instability provides a fitness advantage. In mesophilic and psychrophilic organisms, excessively stable proteins that resist degradation and accumulate in the cell may be harmful. And, in thermophilic organisms, an enzyme that is highly active at low temperatures might be so active at high temperatures that it consumes substrate or produces product at rates inappropriate for physiological function and therefore provides a fitness disadvantage. We cannot conclude, from the results presented here, which scenario, random drift or selective pressure, is responsible for the paucity of enzymes that are both cold-active and thermostable. However, an intrinsic physical incompatibility between these properties is clearly not a viable explanation. Our work demonstrates that enzymes displaying both properties exist and that few amino acid substitutions are required to obtain them. Thus, if nature wanted them, she could make them.

Proteins do not exist in isolation, but as functional pieces of highly complex organisms surviving in specific environments. Both the physiological requirements of the organism and the particular features of the environment will play a large role in determining the properties of a protein, imposing many constraints over and above those imposed by the laws of physical chemistry. Mistaking biological requirements for physical-chemical requirements is a persistent danger in the

study of natural proteins. Directed evolution is uniquely suited to addressing this problem by removing natural biological and environmental constraints and allowing us to ask what are the physical limitations intrinsic to proteins themselves.

Materials and Methods

Plasmids

Subtilisin S41 was expressed using pCT1, which carries the prosubtilisin sequence of S41 fused to presequence of subtilisin BPN′ (Miyazaki & Arnold, 1999). Subtilisin BPN′ was expressed from plasmid pPG580 (provided by C. W. Saunders, Procter & Gamble), which was constructed as described (Qasim *et al.*, 1997) with the intact BPN′ gene flanked by *Eco*RI and *Bam*HI sites. Subtilisin SSII was expressed from pSPH2R, which was constructed by replacing the mature sequence of S41 in pCT1 with that of subtilisin SSII. A set of primers RA-fusion (5′- CAAACAAACCTGAGGCTCTTTACAACGC-TATGAGAGCTTCTCAACAAATA -3′) and Bsphdown (5′- AAAGGATCCTTATTGAACGCGAGCAAA -3′) permitted amplification of fragment containing the mature sequence of SSII from plasmid M33 (kindly provided by Dr Alan Porter, The National University of Singapore). The reaction mixture contained 10 mM Tris-HCl (pH 8.75), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 1.5 mM $MgSO_4$, 0.1 % (v/v) Triton®X-100, 0.2 mM of each dNTPs, 50 μM of each primers, 10 ng of M33, and 2.5 units *Pfu* polymerase (Stratagene, La Jolla, CA) in a total volume of 100 μl. PCR was carried out on an MJ Research (Watertown, MA) thermal cycler (PTC-200) at 94 °C for 30 seconds, 55 °C for 30 seconds, and 72 °C for 30 seconds, and a total of 30 cycles were performed. The ~1 kb fragment was isolated, purified, and digested with *Bsu*36I and *Bam*HI followed by ligating with pCT1 previously cut with *Bsu*36I and *Bam*HI to exclude the S41 fragment.

Error-prone PCR, saturation mutagenesis, DNA recombination and screening

Error-prone PCR and saturation mutagenesis were carried out as described (Miyazaki & Arnold, 1999). DNA recombination was performed using the StEP method (Zhao *et al.*, 1998). The reaction contained 0.2 μg each template DNA, 10 × PCR buffer, 0.2 mM of each dNTPs, 15 μM of each primer, and 1.25 units *Taq* polymerase 2000 (Stratagene, La Jolla, CA) in 50 μl. The mixture was heated at 95 °C for five minutes and then subjected to thermal cycling (94 °C for 30 seconds; 55 °C for five seconds, 80 cycles). Purified restricted inserts from PCR and recombination reactions were ligated with an expression vector generated by *Bsu*36I-*Bam*HI digestion of pCT1. Libraries were screened by assaying catalytic activity towards a synthetic substrate s-AAPF-*p*Na (Sigma, St. Louis, MO) at 25 °C before and after incubation at elevated temperature as described (Miyazaki & Arnold, 1999).

Protein purification

All the subtilisins used in this study were purified to homogeneity as described (Miyazaki & Arnold, 1999). Concentrations of purified S41 and 3-2G7 were determined using the extinction coefficient $\varepsilon_{280} = 37{,}849\ M^{-1}\ cm^{-1}$ (Davail *et al.*, 1994). Concentration of SSII was determined using a Bio-Rad (Richmond, CA) protein assay kit using wild-type S41 as a standard. Concentration of BPN′ was determined using $E^{0.1\,\%} = 1.20$, which corresponds to an extinction coefficient $\varepsilon_{280} = 32{,}214\ M^{-1}\ cm^{-1}$ (Pantoliano *et al.*, 1989).

Enzyme activity

Proteolytic activity was determined using a small synthetic peptide substrate succinyl-LAla-LAla-LPro-LPhe-*p*-nitroanilide (s-AAPF-*p*Na) by monitoring the formation of released *p*-nitroaniline at 410 nm (DelMar *et al.*, 1979) in a thermostatted Shimadzu (Columbia, MD) BioSpec-1601 spectrophotometer. Concentration of the substrate was determined using an extinction coefficient $\varepsilon_{315} = 14{,}000\ M^{-1}\ cm^{-1}$ (DelMar *et al.*, 1979). Kinetic constants for wild-type S41, 3-2G7, SSII, and BPN′ were determined from a series of initial rates at different concentrations of s-AAPF-*p*Na over the range of 0.025-1.2 mM that bracketed K_M.

Temperature dependence of the specific activity (product formation/μM per minute) was determined in 50 mM *N*-2-hydroxyethylpiperazine-*N*′-3-propane sulfonic acid (HEPPS)-NaOH (pH 8.5), 1 mM $CaCl_2$, 2 mM s-AAPF-*p*Na. The assay solution (100 μl) was pre-incubated at specified temperature for one minute, followed by adding 2 μl of enzyme (2 μM) to initiate the reaction; temperatures were controlled to ±0.3 °C with an MJ Research thermal cycler. The reaction was terminated after one minute by adding 2 μl of 0.1 M phenylmethylsulfonylfluoride and the product formation was analyzed on a Molecular Devices (Sunnyvale, CA) plate reader (THERMOmax) at 405 nm.

Thermal inactivation

Half-lives of irreversible thermal inactivation upon autolysis were determined at 60 °C in 50 mM HEPPS-NaOH (pH 8.5) using 2 μM enzyme. At various time intervals, 5-μl aliquots were removed and diluted into 1 ml of an assay solution (0.1 M Tris-HCl (pH 8.5), 10 mM $CaCl_2$, 0.2 mM s-AAPF-*p*Na) for the measurement of residual activity at 30 °C. Plots of the log of residual activity *versus* time were linear, indicating a first-order decay process under these conditions.

Thermal stability by CD spectroscopy

Thermal denaturation was monitored using the decrease in ellipticity at 222 nm on an Aviv (Lakewood, NJ) CD spectropolarimeter (model 62DS). Enzyme (2 μM) was incubated in 50 mM HEPPS-NaOH (pH 8.5), 1 mM $CaCl_2$, 1 mM phenylmethylsulfonylfluoride, and complete loss of activity was confirmed. The sample was then dialyzed against 50 mM borate-NaOH (pH 8.5), 1 mM $CaCl_2$. Temperature was increased from 0 to 90 °C at a rate of ~1.0 °C/minute.

Homology modeling

A three-dimensional structural model of S41 was constructed based on its homology with subtilisins Carlsberg, Savinase, BPN′ and thermitase. Coordinates (Carlsberg, 1CSE, Bode *et al.*, 1987; BPN′, 2SNI, McPhalen & James, 1988; thermitase; 1TEC, Gros *et al.*, 1989: Savinase, 1SVN, Betzel *et al.*, 1992) were obtained from the RCSB Protein Data Bank (Bernstein *et al.*, 1977).

Sequence alignments and model construction and refinement were carried out using the Homology module of the INSIGHT II molecular modeling software package (Biosym Technologies, San Diego, CA). The Figure was created with MOLSCRIPT (Kraulis, 1991).

Acknowledgments

We thank Yi Tang for assistance with CD experiments. We also thank the Caltech Biopolymer Synthesis and Analysis Resource Center for DNA synthesis and the Caltech DNA Sequencing Core Facility for DNA sequencing. K.M. acknowledges the National Institute of Bioscience and Human Technology (Japan) for its generous financial support. This research was supported by Procter & Gamble.

References

Arpigny, J. L., Lamotte, L. & Gerday, C. (1997). Molecular adaptation to cold of an Antarctic bacteria lipase. *J. Mol. Catal.* **B3**, 29-35.

Benner, S. A. (1989). Enzyme kinetics and molecular evolution. *Chem. Rev.* **89**, 789-806.

Bernstein, F., Koetzle, T., Williams, G., Meyer, E. J., Brice, M., Rodgers, J., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: a computer based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-542.

Betzel, C., Teplyakov, A. V., Harutyunyan, E. H., Saenger, W. & Wilson, K. S. (1990). Thermitase and proteinase K: a comparison of the refined three-dimensional structures of the native enzymes. *Protein Eng*. **3**, 161-172.

Betzel, C., Klupsch, S., Papendorf, G., Hastrup, S., Branner, S. & Wilson, K. S. (1992). Crystal structure of the alkaline proteinase savanase from *Bacillus lentus* at 1.4 Å resolution. *J. Mol. Biol.* **223**, 427-445.

Bode, W., Papamokos, E. & Musil, D. (1987). The high-resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and eglin c, an elastase inhibitor from the leech *Hirudo medicinalis*. Structural analysis, subtilisin structure and interface geometry. *Eur. J. Biochem.* **166**, 673-692.

Bryan, P., Alexander, P., Strausberg, S., Schwarz, F., Lan, W., Gilliland, G. & Gallagher, D. T. (1992). Energetics of folding subtilisin BPN′. *Biochemistry,* **31**, 4937-4945.

DelMar, E. G., Largman, C., Brodrick, J. W. & Geokas, M. C. (1979). A sensitive new substrate for chymotrypsin. *Anal. Biochem.* **99**, 316-320.

Davail, S., Feller, G., Narinx, E. & Gerday, C. (1992). Sequence of the subtilisin-encoding gene from an Antarctic psychrotroph *Bacillus* TA41. *Gene,* **119**, 143-144.

Davail, S., Feller, G., Narinx, E. & Gerday, C. (1994). Cold-adaptation of proteins - purification, characterization, and sequence of the heat-labile subtilisin from the Antarctic psychrophile *Bacillus* TA41. *J. Biol. Chem.* **269**, 17448-17453.

Feller, G., Payan, F., Theys, F., Qian, M., Haser, R. & Gerday, C. (1994). Stability and structural analysis of α-amylase from the antarctic psychrophile *Alteromonas haloplanctis* A23. *Eur. J. Biochem.* **222**, 441-447.

Feller, G., Narinx, E., Aspigny, J. L., Aittaleb, M., Baise, E., Genicot, S. & Gerday, C. (1996). Enzymes from psychrophilic organisms. *FEMS Microbiol. Rev.* **18**, 189-202.

Feller, G., Zekhnini, Z., Lamotte-Brasseur, J. & Gerday, C. (1997). Enzymes from cold-adapted microorganisms. The class C β-lactamase from the antarctic psychrophile *Psychrobacter immobilis* A5. *Eur. J. Biochem.* **244**, 186-191.

Fields, P. A. & Somero, G. N. (1998). Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. *Proc. Natl Acad. Sci. USA,* **95**, 11476-11481.

Giver, L., Gershenson, A., Freskgard, P.-O. & Arnold, F. H. (1998). Directed evolution of a thermostable esterase. *Proc. Natl Acad. Sci. USA,* **95**, 12809-12813.

Gros, P., Betzel, C., Dauter, Z., Wilson, K. S. & Hol, W. G. J. (1989). Molecular dynamics refinement of a thermitase-eglin C complex at 1.98 Å resolution and comparison of two crystal forms that differ in calcium content. *J. Mol. Biol.* **210**, 347-367.

Jacobs, M., Eliasson, M., Uhlen, M. & Flock, J. I. (1985). Cloning, sequencing and expression of subtilisin Carlsberg from *Bacillus licheniformis*. *Nucl. Acids Res.* **13**, 8913-8926.

Jaenicke, R. & Bøhm, G. (1996). The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* **8**, 738-748.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Kuhn, P., Knapp, M., Soltis, S. M., Granshaw, G., Thoene, M. & Bott, R. (1998). The 0.78 Å structure of a serine protease: *Bacillus lentus* subtilisin. *Biochemistry,* **37**, 13446-13452.

Low, P. S., Bada, J. L. & Somero, G. N. (1973). Temperature adaptation of enzymes: roles of the free energy, the enthalpy, and the entropy of activation. *Proc. Natl Acad. Sci. USA,* **70**, 430-432.

Matthews, B. W. (1993). Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* **62**, 139-160.

McPhalen, C. A. & James, M. N. G. (1988). Structural comparison of two serine proteinase-protein inhibitor complexes: eglin-c-subtilisin Carlsberg and CI-2-subtilisin Novo. *Biochemistry,* **27**, 6582-6598.

Meloun, B., Baudys, M., Kostka, V., Hausdorf, G., Froemmel, C. & Hoehne, W. E. (1985). Complete primary structure of thermitase from *Thermoactinomyces vulgaris* and its structural features related to the subtilisin-type proteinases. *FEBS Letters,* **183**, 195-200.

Mitani, M., Harushima, Y., Kuwajima, K., Ikeguchi, M. & Sugai, S. (1986). Innocuous character of [ethylenebis(oxyethylenenitrilo)]tetraacetic acid and EDTA as metal-ion buffers in studying Ca^{2+} binding by α-lactalbumin. *J. Biol. Chem.* **261**, 8824-8829.

Miyazaki, K. & Arnold, F. H. (1999). Exploring non-natural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. *J. Mol. Evol.* **49**, 716-720.

Moore, J. C., Jin, H. M., Kuchner, O. & Arnold, F. H. (1997). Strategies for the *in vitro* evolution of protein function: enzyme evolution by random recombination of improved sequences. *J. Mol. Biol.* **272**, 336-347.

Narinx, E., Davail, S., Feller, G. & Gerday, C. (1992). Nucleotide and derivative amino acid sequence of the subtilisin from the Antarctic psychrotroph *Bacillus* TA39. *Biochim. Biophys. Acta,* **1131**, 111-113.

Narinx, E., Baise, E. & Gerday, C. (1997). Subtilisin from psychrophilic antarctic bacteria: characterization and site-directed mutagenesis of residues possibly involved in the adaptation to cold. *Protein Eng.* **10**, 1271-1279.

Oue, S., Okamoto, A., Yano, T. & Kagamiyama, H. (1999). Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues. *J. Biol. Chem.* **274**, 2344-2349.

Pantoliano, M. W., Whitlow, M., Wood, J. F., Rollence, M. L., Finzel, B. C., Gilliland, G. L., Poulos, T. L. & Bryan, P. N. (1988). The engineering of binding affinity at metal ion binding sites for the stabilization of proteins: subtilisin as a test case. *Biochemistry,* **27**, 8311-8317.

Pantoliano, M. W., Whitlow, M., Wood, J. F., Dodd, S. W., Hardman, K. D., Rollence, M. L. & Bryan, P. N. (1989). Large increases in general stability for subtilisin BPN′ through incremental changes in the free energy of unfolding. *Biochemistry,* **28**, 7205-7213.

Qasim, M. A., Ganz, P. J., Saunders, C. W., Bateman, K. S., James, M. N. & Laskowski, M., Jr (1997). Interscaffolding additivity. Association of P1 variants of eglin c and of turkey ovomucoid third domain with serine proteinases. *Biochemistry,* **36**, 1598-1607.

Russel, R. J. M. & Taylor, G. L. (1995). Engineering thermostability: lessons from thermophilic proteins. *Curr. Opin. Biotech.* **6**, 370-374.

Siezen, R. J. & Leunissen, J. A. M. (1997). Subtilases: the superfamily of subtilisin-like serine protease. *Protein Sci.* **6**, 501-523.

Siezen, R. J., de Vos, W. M., Leunissen, J. A. M. & Dijkstra, B. W. (1991). Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases. *Protein Eng.* **4**, 719-737.

Simpson, B. K. & Haard, N. F. (1984). Trypsin from Greenland cod, *Gadus ogac*. Isolation and comparative properties. *Comp. Biochem. Physiol.* **B79**, 613-622.

Spiller, B., Gershenson, A., Arnold, F. H. & Stevens, R. C. (1999). A Structural view of evolutionary divergence. *Proc. Natl Acad. Sci. USA,* **96**, 12305-12310.

Stahl, M. L. & Ferrari, E. (1984). Replacement of the *Bacillus subtilis* subtilisin structural gene with an *in vitro*-derived deletion mutation. *J. Bacteriol.* **158**, 411-418.

Teplyakov, A. V., Kuranova, I. P., Harutyunyan, E. H., Vainshtein, B. K., Frommel, C., Hohne, W. E. & Wilson, K. S. (1990). Crystal structure of thermitase at 1.4 Å resolution. *J. Mol. Biol.* **214**, 261-279.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673-4680.

Van den Burg, B., Vriend, G., Veltman, O. R., Venema, G. & Eijsink, V. G. (1998). Engineering an enzyme to resist boiling. *Proc. Natl Acad. Sci. USA,* **95**, 2056-2060.

Vølkl, P., Markiewicz, P., Stetter, K. O. & Miller, J. H. (1994). The sequence of a subtilisin-type protease (aerolysin) from the hyperthermophilic archaeum *Pyrobaculum aerophilum* reveals sites important to thermostability. *Protein Sci.* **3**, 1329-1340.

Voordouw, G., Milo, C. & Roche, R. S. (1976). Role of bound calcium ions in thermostable, proteolytic enzymes. Separation of intrinsic and calcium ion contributions to the kinetic thermal stability. *Biochemistry,* **15**, 3716-3724.

Voorhorst, W. G., Eggen, R. I., Geerling, A. C., Platteeuw, C., Siezen, R. J. & Vos, W. M. (1996). Isolation and characterization of the hyperthermostable serine protease, pyrolysin, and its gene from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **271**, 20426-20431.

Watanabe, K., Masuda, T., Ohashi, H., Mihara, H. & Suzuki, Y. (1994). Multiple proline substitutions cumulately thermostabilize *Bacillus cerues* ATCC7064 oligo-1,6-glucosidae. Irerefragable proof supporting the proline rule. *Eur. J. Biochem.* **226**, 277-283.

Wati, M. R., Thanabalu, T. & Porter, A. G. (1997). Gene from tropical *Bacillus sphaericus* encoding a protease closely related to subtilisins from Antarctic bacilli. *Biochim. Biophys. Acta,* **1352**, 56-62.

Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A. & Chen, E. Y. (1983). Cloning, sequencing, and secretion of *Bacillus amyloliquefaciens* subtilisin in *Bacillus subtilis*. *Nucl. Acids Res.* **11**, 7911-7925.

Zavodszky, P., Kardos, J., Svingor, A. & Petsko, G. A. (1998). Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl Acad. Sci. USA,* **95**, 7406-7411.

Zhang, J. H., Dawes, G. & Stemmer, W. P. (1997). Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc. Natl Acad. Sci. USA,* **94**, 4504-4509.

Zhao, H. & Arnold, F. H. (1999). Directed evolution converts subtilisin E into a functional equivalent of thermitase. *Protein Eng.* **12**, 47-53.

Zhao, H., Giver, L., Shao, Z., Affholter, J. A. & Arnold, F. H. (1998). Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nature Biotechnol.* **14**, 258-261.

Edited by J. A. Wells

(Received 25 October 1999; received in revised form 4 February 2000; accepted 18 February 2000)