Eur. J. Biochem. 226, 277–283 (1994)

Multiple proline substitutions cumulatively thermostabilize *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase

Irrefragable proof supporting the proline rule

Kunihiko WATANABE, Tomoko MASUDA, Hiroyuki OHASHI, Hisaaki MIHARA and Yuzuru SUZUKI

Department of Agricultural Chemistry, Kyoto Prefectural University, Japan

(Received July 18/September 2, 1994) – EJB 94 1075/2

Nine residues of *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase were replaced stepwise with proline residues. Of the nine residues, Lys121, Glu208 and Glu290 were at second sites of β turns; Asn109, Glu175 and Thr261 were at N-caps of α helices; Glu216, Glu270 and Glu378 were in coils within loops. The replacements were carried out in the order, Lys121→Pro, Glu175→Pro, Glu290→Pro, Glu208→Pro, Glu270→Pro, Glu378→Pro, Thr261→Pro, Glu216→Pro and Asn109 →Pro. The resultant nine active mutant enzymes contained 1–9 more proline residues than *B. cereus* oligo-1,6-glucosidase. The thermostability of these mutants was additively enhanced with the increase in the number of proline residues introduced. The increase in the thermostability was most remarkable when proline residues were introduced at second sites of β turns or at N-caps of α helices. The above results afforded irrefragable proof for the proline rule as an effective principle for increasing protein thermostability [Suzuki, Y., Oishi, K., Nakano, H. & Nagayama, T. (1987) *Appl. Microbiol. Biotechnol. 26*, 546–551].

Suzuki et al. [1–4] had found a strong correlation between the increase in the number of proline residues and the rise in the thermostability of bacillary oligo-1,6-glucosidases (dextrin 6-α-glucanohydrolase). On the basis of this finding, Suzuki et al. [4, 5] proposed an idea (the proline rule, formerly named the proline theory) which stated that the increases in the frequency of proline occurrence at β turns (their second sites, see Note) and in the total number of hydrophobic residues can enhance protein thermostability. This was first given strong support by Matthews et al. [6]. These authors increased the thermostability of bacteriophage T4 lysozyme by replacing an alanine residue with a proline residue at one of the β turns, so as to decrease the backbone entropy of unfolding. The enhanced thermostability by single-site proline substitutions has also been reported for hen egg white lysozyme [7], human lysozyme [8], *Escherichia coli* ribonuclease HI [9, 10] and *Bacillus stearothermophilus* neutral protease [11].

Watanabe et al. [12–14] deduced primary structures of oligo-1,6-glucosidases for *B. cereus* ATCC7064 (mesophile), *Bacillus coagulans* ATCC7050 (facultative thermophile) and *Bacillus thermoglucosidasius* KP1006 (DSM2542, obligate thermophile) from the corresponding gene structures. Very recently, Kizaki et al. [15] determined the tertiary structure of *B. cereus* oligo-1,6-glucosidase by X-ray crystallography. This enzyme consists of three domains. The N-terminal domain forms a $(\alpha/\beta)_8$ barrel structure. The subdomain forms a loop-rich structure with an α helix and a β sheet. Eight antiparallel β strands, folded in two Greek-key motifs, create the C-terminal domain. The subdomain and the loops binding C-termini of β strands to N-termini of α helices of the $(\alpha/\beta)_8$ barrel structure form a deep cleft at the active site (Fig. 1). Primary structures of three oligo-1,6-glucosidases were analysed in the light of the tertiary structure of the *B. cereus* enzyme [12–15]. This analysis gave four definite findings. (a) Each enzyme has at least 22 critical sites, at which proline residues can occur randomly with increasing thermostability. (b) The majority of the critical sites are in solvent-exposed β turns (their second sites) or coils within loops forming the active-site cleft and its wall. (c) A few of the critical sites are at N-caps of α helices of the $(\alpha/\beta)_8$ barrel structure. (d) Proline residues mainly replace charged or polar residues (Glu, Lys) at the critical sites. These substitutions result from two-base exchanges in codons by transversion.

Correspondence to Y. Suzuki, Department of Agricultural Chemistry, Kyoto Prefectural University, Shimogamo, Sakyo-ku, Kyoto-shi, Japan 606

Fax: +81 75 781 6043.

Abbreviations. t_m, temperature at which enzyme was 50% inactivated at pH 7.0 in 10 min; Δt_m, the difference between t_m values of *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase and its mutant.

Enzyme. Oligo-1,6-glucosidase or dextrin 6-α-glucanhydrolase (EC 3.2.1.10).

Note. The β turn involves four consecutive residues *i*, *i*+1, *i*+2 and *i*+3 in a protein where the polypeptide chain folds back on itself by nearly 180° [Wilmot, C. M. & Thornton, J. M. (1988) *J. Mol. Biol. 203*, 221–232] The position *i*+1 is referred to the second site of the β turn. There is a hydrogen bond between the CO group of residue *i* and the NH group of residue *i*+3. The type-I turn has $(\phi,\psi)_2 = (-60°, -30°)$ and $(\phi,\psi)_3 = (-90°, 0°)$, while the type-II turn has $(\phi,\psi)_2 = (-60°, 120°)$ and $(\phi,\psi)_3 = (80°, 0°)$, for residues *i*+1 and *i*+2 of the β bend, respectively. Of seven critical residues in β turns of *B. cereus* oligo-1,6-glucosidase, Lys121(→Pro in Mut-1), Glu290 (→Pro in Mut-3), Pro443, Lys457 and Glu487 are at the second sites of type-I turns, while Glu208 (→Pro in Mut-4) and Pro257 are at those of type-II turns [Kizaki, H., Hata, Y., Watanabe, K., Katsube, Y. & Suzuki, Y. (1993) *J. Biochem. (Tokyo) 113*, 646–649].

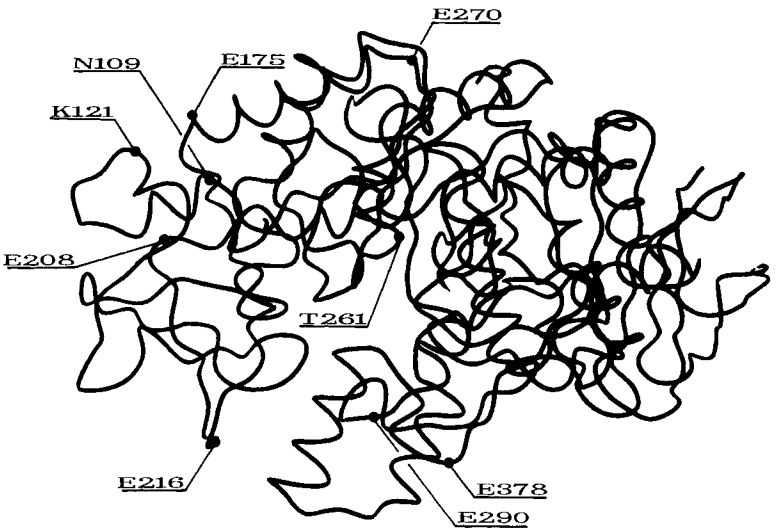

Fig. 1. Positions of proline substitutions in *B. cereus* oligo-1,6-glucosidase. The enzyme model is from [15]. The residues substituted for proline residues are indicated by their corresponding single letters and residue numbers.

On the basis of findings a–d, we selected nine of the 22 critical sites in *B. cereus* oligo-1,6-glucosidase (Figs 1 and 2). Of the nine sites, three were at second sites of β turns, three were at N-caps of α helices, and three were in coils within loops. In the present study, we found that the thermostability of *B. cereus* oligo-1,6-glucosidase can be cumulatively increased by adding proline residues to these nine critical sites. The most effective sites are second sites of β turns or N-caps of α helices. These results have afforded irrefragable proof for the proline rule as an effective principle for increasing protein thermostability [4, 5].

MATERIALS AND METHODS

Materials

Restriction enzymes, DNA-modifying enzymes, the M13 sequencing kit and the site-directed mutagenesis kit were purchased from Toyobo Co. Ltd or from Takara Shuzo Co. Ltd. [α-^{32}P]dCTP (400 Ci/mmol) for sequencing of DNA was purchased from Amersham Co. All other chemicals were of analytical grade.

Plasmids and hosts

The hybrid plasmid pBCE4-2 carried the *B. cereus* ATCC7064 oligo-1,6-glucosidase gene [16]. This plasmid was used for site-directed mutagenesis of the gene. Plasmids resulting from cumulative mutagenesis were used for the expression of mutant oligo-1,6-glucosidases in *E. coli* MV1184 [16–18]. *E. coli* BW313 was used for the preparation of single-stranded uracil-containing DNA templates for site-directed mutagenesis [19].

Mutagenesis

Site-directed mutagensis was carried out according to the method of Kunkel [19]. The mutagenesis of the *B. cereus* ATCC7064 oligo-1,6-glucosidase gene was performed stepwise using pBCE4-2 to direct the following mutations, Lys121→Pro, Glu175→Pro, Glu290→Pro, Glu208→Pro, Glu270→Pro, Glu378→Pro, Thr261→Pro, Glu216→Pro and Asn109→Pro (Fig. 1). Primers used for the mutations were 5′-AAGATAATCCATATAGA-3′ for Lys121→Pro, 5′-GGGATAATCCAAAGGTA-3′ for Glu175→Pro, 5′-TTCAGGACCAGGTGGA-3′ for Glu290→Pro, 5′-CTAAAGAACCGGGATTAC-3′ for Glu208→Pro, 5′-ATACTGGGCCAGAACGAA-3′ for Glu270→Pro, 5′-TCCGATTTCCGTCGATTGA-3′ for Glu378→Pro, 5′-GTGTAACGCCAGAAGAAG-3′ for Thr261→Pro, 5′-TGTTGAAACACCAGAAGAGGGG-3′ for Glu216→Pro, 5′-GATGAACATCCTTGGTTTAT-3′ for Asn109→Pro (mismatches are underlined). Resultant mutant clones of pBCE4-2 were selected and verified by dideoxy DNA sequencing [20]. Mutant oligo-1,6-glucosidases expressed in *E. coli* MV1184 cells were purified to homogeneity as described earlier [16]. Primers used were synthesized with an Applied Biosystems 380B DNA synthesizer.

Heat inactivation

B. cereus ATCC7064 oligo-1,6-glucosidase [16] and its mutants were assayed spectrophotometrically at 35 °C and at pH 6.8 with *p*-nitrophenyl-α-D-glucopyranoside as the substrate [12]. 1 U each enzyme was defined as the amount of the enzyme hydrolyzing 1 μmol substrate in 1 min [12]. Each enzyme (17 μg/ml) was heated for 10 min at different temperatures (0 °C, 30–65 °C) in 50 μl 50 mM potassium phosphate, 5 mM EDTA, pH 7.0 (buffer A). The mixture was assayed immediately after a fourfold dilution with chilled buffer A. The enzyme (17 μg/ml) was incubated at 45 °C or at 48 °C for different intervals in buffer A. The remaining activity was determined as described above.

RESULTS

Proline-introduced oligo-1,6-glucosidases

Nine residues of *B. cereus* ATCC7064 oligo-1,6-glucosidase were replaced stepwise with proline residues by site-directed mutagenesis. Of the nine residues, Lys121, Glu208 and Glu290 were at second sites of β turns; Asn109, Glu175 and Thr261 were at N-caps of α helices and Glu216, Glu270 and Glu378 were in coils within loops (Figs 1 and 2). The majority of these were exposed to solvent (Fig. 1). The resul-

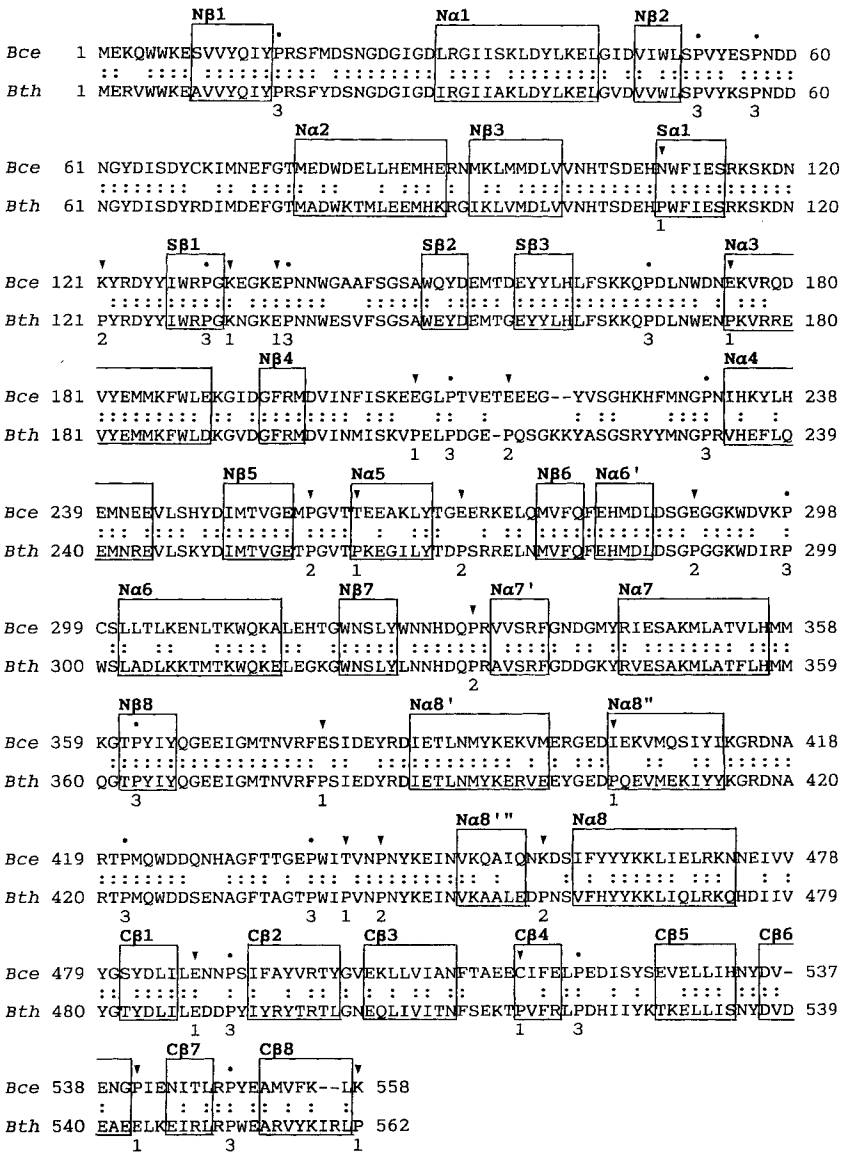

Fig. 2. 15 sites of conservative proline residues and 21 critical sites of proline occurrence among oligo-1,6-glucosidases from *B. cereus*, *B. coagulans* and *B. thermoglucosidasius*. The amino acid sequences of *B. cereus* oligo-1,6-glucosidase (upper line, *Bce*) and of *B. thermoglucosidasius* oligo-1,6-glucosidase (lower line, *Bth*) are from the previous reports [12, 14]. The sequence of *B. coagulans* oligo-1,6-glucosidase is not presented in this figure [13]. Colons (:) show the identical residues between the *B. cereus* enzyme and the *B. thermoglucosidasius* enzyme. The 15 sites of conservative proline residues among three bacillary oligo-1,6-glucosidases are indicated with ● on the sequences of the *B. cereus* enzyme and the *B. thermoglucosidasius* enzyme. The 21 critical sites at which proline residues occur randomly among the three enzymes are presented with ▼ on the two sequences. One critical site equivalent to Pro293 of *B. coagulans* oligo-1,6-glucosidase [13] (see Discussion) is not shown in this figure, since the site is in the regions of residues 292–294 of the same enzyme, and these residues are lacked on the two sequences [12–14]. The numbers typed under the 15 conservative proline residues and the 21 critical sites denote the number of proline occurrences among three bacillary oligo-1,6-glucosidases. α helices (Nα) and β strands (Nβ) in the N-terminal domain, an α helix (Sα) and β strands (Sβ) in the subdomain, and β strands (Cβ) in the C-terminal domain are boxed. Their assignments are based on the result of X-ray crystallographic analysis of *B. cereus* oligo-1,6-glucosidase [15]. The positions of secondary structures in *B. cereus* oligo-1,6-glucosidase are as follows. Nβ1 is residues 9–15; Nα1 is residues 30–43; Nβ2 is residues 47–50; Nα2 is residues 78–90; Nβ3 is residues 93–100; Sα1 is residues 109–114; Sβ1 is residues 127–131; Sβ2 is residues 149–152; Sβ3 is residues 157–161; Nα3 is residues 175–190; Nβ4 is residues 195–198; Nα4 is residues 233–243; Nβ5 is residues 250–255; Nα5 is residues 261–267; Nβ6 is residues 277–280; Nα6′ is residues 282–286; Nα6 is residues 301–314; Nβ7 is residues 320–324; Nα7′ is residues 333–337; Nα7 is residues 344–356; Nβ8 is residues 361–365; Nα8′ is residues 386–397; Nα8″ is residues 403–412; Nα8‴ is residues 450–455; Nα8 is residues 460–473; Cβ1 is residues 481–485; Cβ2 is residues 492–499; Cβ3 is residues 502–509; Cβ4 is residues 515–518; Cβ5 is residues 527–533; Cβ6 is residues 536–540; Cβ7 is residues 544–547; Cβ8 is residues 552–557.

tant nine active mutant enzymes were expressed in *E. coli* MV1184 cells. These mutants were named Mut-1, Mut-2, Mut-3, Mut-4, Mut-5, Mut-6, Mut-7, Mut-8 and Mut-9 in increasing order of proline residues introduced. Mutations in these enzymes were as follows: Lys121→Pro in Mut-1, Lys121→Pro/Glu175→Pro in Mut-2, Lys121→Pro/Glu175→Pro/Glu290→Pro in Mut-3, Lys121→Pro/Glu175→Pro/Glu290→Pro/Glu208→Pro in Mut-4, Lys121→Pro/Glu175→Pro/Glu290→Pro/Glu208→Pro/Glu270→Pro in Mut-5, Lys121→Pro/Glu175→Pro/Glu290→Pro/Glu208→Pro/Glu-

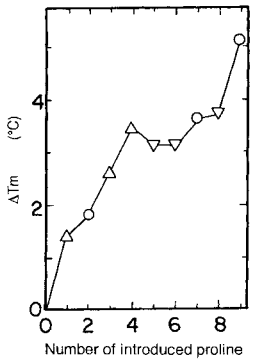

Fig. 3. A relationship between the thermostability and the number of proline residues added in the mutants of *B. cereus* oligo-1,6-glucosidase. The tempeartures t_m at which these enzymes were 50% inactivated, were determined within the errors of ±0.2°C, as described in the text. The difference Δt_m between t_m values of the *B. cereus* enzyme and its mutants was used to evaluate the thermostability of the mutants. The t_m of the wild-type enzyme was 44.5°C. The number of proline residues introduced into each mutant corresponds to the mutant number, i.e. nine proline residues added to Mut-9. Symbols △, O and ▽ indicate second sites of β turns, N-caps of α helices and coils within loops, respectively, at which the last proline residues were added on cumulative proline substitutions in mutants (see text).

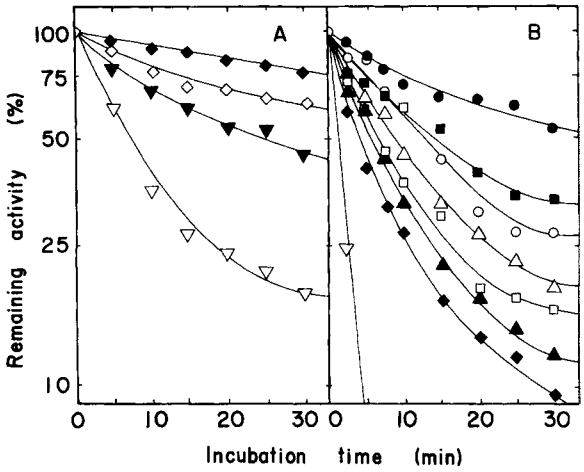

Fig. 4. Kinetics of the inactivation of *B. cereus* oligo-1,6-glucosidase and its mutants Mut-1 to Mut-9 at 45°C (A) and at 48°C (B). ▽ wild-type enzyme; ▼, Mut-1; ◇ Mut-2; ◆, Mut-3; △, Mut-4; ▲ Mut-5; □, Mut-6; ■, Mut-7; ○, Mut-8; ●, Mut-9.

270→Pro/Glu378→Pro in Mut-6, Lys121→Pro/Glu175→Pro/Glu290→Pro/Glu208→Pro/Glu270→Pro/Glu378→Pro/Thr261→Pro in Mut-7, Lys121→Pro/Glu175→Pro/Glu290→Pro/Glu208→Pro/Glu270→Pro/Glu378→Pro/Thr261→Pro/Glu216→Pro in Mut-8, and Lys121→Pro/Glu175→Pro/Glu290→Pro/Glu208→Pro/Glu270→Pro/Glu378→Pro/Thr261→Pro/Glu216→Pro/Asn109→Pro in Mut-9. Specific activities of the purified mutant enzymes were determined as 293 U/mg for Mut-1, 242 U/mg for Mut-2, 217 U/mg for Mut-3, 238 U/mg for Mut-4, 220 U/mg for Mut-5, 205 U/mg for Mut-6, 203 U/mg for Mut-7, 201 U/mg for Mut-8, and 237 U/mg for Mut-9. These values were about comparable with that (282 U/mg) of *B. cereus* oligo-1,6-glucosidase [16].

Secondary structures versus thermostability

B. cereus ATCC7064 oligo-1,6-glucosidase and its mutants were exposed to temperatures of 30−65°C at pH 7.0. The temperature t_m, at which the enzyme was 50% inactivated in 10 min, was determined using the plots of residual activities *versus* treated temperatures. This test was repeated separately 3−5 times with each enzyme. Mean values of t_m achieved within errors of ±0.2°C were 44.5°C for *B. cereus* oligo-1,6-glucosidase, 45.9°C for Mut-1, 46.3°C for Mut-2, 47.1°C for Mut-3, 47.9°C for Mut-4, 47.6°C for Mut-5, 47.6°C for Mut-6, 48.1°C for Mut-7, 48.2°C for Mut-8, and 49.6°C for Mut-9. The difference Δt_m between these t_m values of *B. cereus* oligo-1,6-glucosidase and of its mutant was used to evaluate the thermostability of the mutant. The thermostability of mutants was additively enhanced with the increase in the number of proline residues introduced, although the thermostability decreased only slightly on the transition Mut-4→Mut-5 (Fig. 3). This trend was far clearer with data from the kinetics of the inactivation of mutants (Fig. 4). Their half-life times were; at 45°C, 6.0 min for *B. cereus* oligo-1,6-glucosidase, 25 min for Mut-1, 120 min for Mut-2 and >120 min for Mut-3 (Fig. 4A); at 48°C, 1.5 min for *B. cereus* oligo-1,6-glucosidase, 3.8 min for Mut-3, 8.8 min for Mut-4, 5.8 min for Mut-5, 6.7 min for Mut-6, 14 min for Mut-7, 13 min for Mut-8 and 40 min for Mut-9 (Fig. 4B). The increase ($\Delta\Delta t_m$) in the thermostability of mutants was most marked when proline residues were added at second sites of β turns (Lys121→Pro in Mut-1, Glu290→Pro in Mut-3, Glu208→Pro in Mut-4) or at N-caps of α helices (Glu175→Pro in Mut-2, Thr261→Pro in Mut-7, Asn109→Pro in Mut-9; Fig. 3). The addition of proline residues into coils was less effective or neutral for the thermostability (Glu270→Pro in Mut-5, Glu378→Pro in Mut-6, Glu216→Pro in Mut-8). The increase ($\Delta\Delta t_m$) in Δt_m due to one additional proline residue ranged over 0.8−1.4°C for second sites of β turns and from 0.4−1.4°C for N-caps of α helices.

DISCUSSION

Oligo-1,6-glucosidases from *B. cereus* ATCC7064, *B. coagulans* ATCC7050 and *B. thermoglucosidasius* KP1006 contain 19, 24 and 32 proline residues, respectively [12−14]. 15 of these occur at the common sites of three enzymes, which refer to positions 16, 52, 57, 130, 137, 168, 211, 231, 298, 362, 421, 437, 490, 520 and 549 of the *B. cereus* enzyme (Fig. 2). Each enzyme has at least 22 critical sites, at which proline can occur randomly with increasing thermostability (see above) [12−14]. These sites correspond to Asn109, Lys121, Lys132, Glu136, Glu175, Glu208, Glu216, Pro257, Thr261, Glu270, Glu290, Pro331, Glu378, Ile403, Thr440, Pro443, Lys457, Glu487, Cys515, Pro541 and Lys558 of the *B. cereus* enzyme, and Pro293 of the *B. coagulans* enzyme (Fig. 2).

X-ray crystallography of *B. cereus* oligo-1,6-glucosidase has clarified its 21 critical sites and its 15 conservative proline residues [15]. (a) 19 of the critical sites are exposed to solvent, but only two (Pro257 and Thr261) are buried within the molecule. Of the critical sites, seven (Lys121, Glu208, Pro257, Glu290, Pro443, Lys457 and Glu487) are at second sites of β turns, nine (Lys132, Glu136, Glu216, Glu270, Pro331, Glu378, Thr440, Pro541 and Lys558) are in random coils, four (Asn109, Glu175, Thr261 and Ile403) are at N-caps of α helices, and one (Cys515) is at the N-terminus of a β strand (Fig. 2). (b) In contrast, 11 of the conservative proline residues are buried within the molecule, and only

four (Pro130, Pro231, Pro298 and Pro520) are on its surface. However, Pro231, Pro520 and Pro549 are locked at second sites of β turns, and Pro130 and Pro362 are within β strands (Fig. 2). All other 10 proline residues are in loops (Fig. 2). This finding (b) suggests that the conservative proline residues contribute towards maintaining the basic conformational integrity of the enzyme. This is consistent with results from *E. coli* tryptophan synthetase [21], *Aspergillus oryzae* ribonuclease T1 [22], bovine ribonuclease A [23] and *Dictyostelium* nucleoside diphosphate kinase [24], which demonstrated that the protein stability is dramatically decreased by replacing conservative or buried proline residues by other amino acids. Indeed, Alber et al. [25] found that mutations substantially destabilizing bacteriophage T4 lysozyme are only observed in its rigid parts. (c) Ramachandran plots of the conformational angles (ϕ and ψ) [26] show that all residues except for four (Lys132, Glu216, Cys515 and Lys558) at the 21 critical sites of *B. cereus* oligo-1,6-glucosidase are clustered around two distinct regions of α ($\phi = -47°$−approximately $-82°$, $\psi = -10°$−approximately $-54°$) and β ($\phi = -48°$−approximately $-72°$, $\psi = 128°$approximately 167°), as described using the nomenclature of Efimov [15, 27]. Their mean values of ϕ and ψ are $-60°$ and $-29°$ for the α region, and $-61°$ and 150° for the β region. These regions are near two theoretically predicted energy minima for a proline conformation in the polypeptide chain [28−30]. The ϕ and ψ angles are $-100°$ and $-53°$ for Lys132; $-123°$ and 37° for Glu216; $-149°$ and 170° for Cys515; and 78° and $-77°$ for Lys558. The latter two (ϕ,ψ) angles are not surprising in terms of Cys515 adopting a β conformation at the N-end of a β strand [30] and of Lys558 being at the C-terminus of the enzyme (Fig. 2). (d) 17 of the residues, with four exceptions (Met256, Gly289, Leu486 and Gly540), preceding the 21 critical sites take extended β conformations with $\phi = -56°$−approximately $-158°$ and $\psi = 117°$−approximately 176° [15]. Met256 and Leu486 occur in the regions of δ ($\phi = -135°$, $\psi = 87°$) and α_L ($\phi = 44°$, $\psi = 45°$), respectively [27]. The ϕ and ψ angles are $-105°$ and $-155°$ for Gly289, and 172° and 144° for Gly540. These 20 residues, with the exception of Leu486, preceding proline sites have ϕ and ψ angles within the allowed regions given by Schimmel and Flory [28] and by other investigators [29, 30].

Site-directed mutagenesis has proven that the thermostability of *B. cereus* oligo-1,6-glucosidase can be cumulatively increased by introducing multiple proline residues to its critical sites (Figs 3 and 4). Nine residues replaced by proline residues at the critical sites are Asn109, Lys121, Glu175, Glu208, Glu216, Thr261, Glu270, Glu290 and Glu378. All of these are in the N domain and the subdomain of the enzyme (Fig. 1). Lys121, Glu208 and Glu290 are at second sites of β turns on the loops between α helix Sα1 and β-strand Sβ1, between β-strand Nβ4 and α helix Nα4, and between α helices Nα6′ and Nα6, respectively (Fig. 2) [15]. Asn109, Glu175 and Thr261 are at N-caps of α helices Sα1, Nα3 and Nα5, respectively (Fig. 2). Glu216, Glu270 and Glu378 are in random coils on the flexible loops between β-strand Nβ4 and α helix Nα4, between α helix Nα5 and β-strand Nβ6, and between β-strand Nβ8 and α helix Nα8′, respectively (Fig. 2). These critical residues, with the exception of Glu270, are located on the C-terminal side of the $(\alpha/\beta)_8$ barrel structure, i.e. on the side of the active-site cleft (Fig. 1). Glu270 is on the N-terminal side of the barrel (Fig. 1). The proline substitutions for the nine critical sites may cause no significant change on the enzyme. This is compatible with the following findings. Proline has a propensity for occurring at the second sites of β turns, at the N-caps of α helices and in coils in proteins [30−35]. All of the mutants with 1−9 proline residues added are as active as *B. cetreus* oligo-1,6-glucosidase. Nine critical residues (Asn109, Lys121, Glu175, Glu208, Glu216, Thr261, Glu270, Glu290 and Glu378) of the *B. cereus* enzyme are replaced by proline residues in its thermostable counterpart of *B. thermoglucosidasius* (Fig. 2) [12, 13]. Both enzymes are highly similar in their amino acid sequences around the nine critical sites (Fig. 2). The nine critical residues and their preceding residues of the *B. cereus* enzyme have ϕ and ψ angles that are within the allowed range for proline replacements, although Glu216 is in the range of γ [27] (seven critical residues with $\phi = -49°$−approximately $-82°$, $\psi = -10°$−approximately $-54°$, Glu208 with $\phi = -48°$ and $\psi = 128°$; eight preceding residues except Gly289 with $\phi = -72°$−approximately $-158°$, $\psi = 120°$−approximately 167° [15]; see c and d in the former paragraph). The above critical residues except Thr261 are exposed to solvent (Fig. 1). Proline residues substituted for these residues would less perturb their neighboring conformations than those for the buried residues (a and b) [25, 36]. Indeed, enhanced thermostability has been achieved by adding proline residues to the surface regions of various proteins [6−11]. Simple estimates of the contribution by the mutations Lys→Pro, Glu→Pro, Asn→Pro and Thr→Pro to the backbone entropy of unfolding are -17−approximately -8 J mol^{-1} K^{-1} [6, 37]. The corresponding increases in the free energy of unfolding are 5.4−approximately 2.9 kJ/mol with *B. cereus* oligo-1,6-glucosidase at 45°C and pH 7.0 [6, 37]. However, the observed increase in thermostability varies according to the sites mutated (Figs 3 and 4). The *B. cereus* enzyme is thermostabilized most remarkably when proline residues are substituted for the critical residues at second sites (Lys121, Glu208 and Glu290) of β turns or at N-caps (Asn109, Glu175 and Thr261) of α helices. This may be caused because of a fall in the entropy of the denatured state of the enzyme without a significant change in the native backbone conformation (see above) [6]. This agrees well with results from T4 lysozyme, human lysozyme and *B. stearothermophilus* neutral protease [6, 8, 11]. However, the substitution of proline residues for the critical residues (Glu216, Glu270 and Glu378) in coils of *B. cereus* oligo-1,6-glucosidase are less effective for the thermal stabilization (Figs 3 and 4). These proline residues may increase the conformational energy in the native state, which can offset the entropic stability gain, as demonstrated with hen egg white lysozymre [7]. In contrast, *E. coli* ribonuclease HI and *B. stearothermophilus* neutral protease are stabilized by replacing residues on loops by proline residues [9, 11].

The present study supports the proline rule [4, 5] (Figs 3 and 4). The rule reveals a simple and steady strategy used in evolution, whereby a protein can be stepwise adapted to thermal environments without largely affecting its functional and structural integrity. This strategy has three characteristics [38−44]. Proline residues critical for thermal stabilization favor second sites of β turns and N-caps of α helices (Figs 3 and 4). Most of these critical sites occur randomly over the protein surface (Figs 1 and 2). Proline residues added to critical sites independently contribute to thermal stabilization, i.e. such effects are additive, but not cooperative (Figs 3 and 4). These features indicate that the thermal adaptation of a protein is accomplished by the cumulation of many small stabilization effects on its peripheral neutral regions, resulting in the molecule being more packed.

It remains to be solved whether or not the proline rule applies in general to a wide range of enzymes with different structural motifs [4, 14, 45, 46]. Apart from such an evolutional problem, however, it is likely that the principle of the proline rule is useful in protein engineering, since it suggests that the potential sites and residues (such as Glu or Lys) of substitution in the protein improve thermostability [5].

REFERENCES

1. Suzuki, Y., Aoki, R. & Hayashi, H. (1982) Assignment of a *p*-nitrophenyl-α-D-glucopyranoside-hydrolyzing α-glucosidase of *Bacillus cereus* ATCC7064 to an exo-oligo-1,6-glucosidase, *Biochim. Biophys. Acta 704*, 476–483.
2. Suzuki, Y. & Tomura, Y. (1986) Purification and characterization of *Bacillus coagulans* oligo-1,6-glucosidase, *Eur. J. Biochem. 158*, 77–83.
3. Suzuki, Y., Fujii, H., Uemura, H. & Suzuki, M. (1987) Purification and characterization of extremely thermostable exo-oligo-1,6-glucosidase from a caldoactive *Bacillus* sp. KP1228, *Starch/Stärke 39*, 17–23.
4. Suzuki, Y., Oishi, K., Nakano, H. & Nagayama, T. (1987) A strong correlation between the increase in number of proline residues and the rise in thermostability of five *Bacillus* oligo-1,6-glucosidase, *Appl. Microbiol. Biotechnol. 26*, 546–551.
5. Suzuki, Y. (1989) A general principle of increasing protein thermostability, *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci. 65*, 146–148.
6. Matthews, B. W., Nicholson, H. & Becktel, W. J. (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding, *Proc. Natl Acad. Sci. USA 84*, 6663–6667.
7. Ueda, T., Tamura, T., Maeda, Y., Hashimoto, Y., Miki, T., Yamada, H. & Imoto, T. (1993) Stabilization of lysozyme by the introduction of Gly-Pro sequence, *Protein Eng. 6*, 183–187.
8. Herning, T., Yutani, K., Inaka, K., Kuroki, R., Matsushima, M. & Kikuchi, M. (1992) Role of proline residues in human lysozyme stability: a scanning calorimetric study combined with X-ray structure analysis of proline mutants, *Biochemistry 31*, 7077–7085.
9. Kimura, S., Nakamura, H., Hashimoto, T., Oobatake, M. & Kanaya, S. (1992) Stabilization of *Escherichia coli* ribonuclease HI by strategic replacement of amino acid residues with those from the thermophilic counterpart, *J. Biol. Chem. 267*, 21535–21542.
10. Ishikawa, K., Kimura, S., Kanaya, S., Morikawa, K. & Nakamura, H. (1993) Structure study of mutants of *Escherichia coli* HI with enhanced thermostability, *Protein Eng. 6*, 85–91.
11. Hardy, F., Vriend, G., Veltman, O. R., van der Vinne, B., Venema, G. & Eijsink, V. G. H. (1993) Stabilization of *Bacillus stearothermophilus* neutral protease by introduction of proline residues, *FEBS Lett. 317*, 898–92.
12. Watanabe, K., Kitamura, K., Iha, H. & Suzuki, Y. (1990) Primary structure of the oligo-1,6-glucosidase of *Bacillus cereus* ATCC7064 deduced from the nucleotide sequence of the cloned gene, *Eur. J. Biochem. 192*, 609–620.
13. Watanabe, K., Kitamura, K., Iha, H., Chishiro, K. & Suzuki, Y. (1990) Proline residues of oligo-1,6-glucosidases and their thermostability – The structure of an oligo-1,6-glucosidase from the facultative thermophile *Bacillus coagulans* ATCC7050, *Nippon Nogeikagaku Kaishi 64*, 575.
14. Watanabe, K., Chishiro, K., Kitamura, K. & Suzuki, Y. (1991) Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermoglucosidasius* KP1006, *J. Biol. Chem. 266*, 24287–24294.
15. Kizaki, H., Hata, Y., Watanabe, K., Katsube, Y. & Suzuki, Y. (1993) Polypeptide folding of *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase revealed by 3.0 Å resolution X-ray analysis, *J. Biochem. (Tokyo) 113*, 646–649.
16. Watanabe, K., Kitamura, K., Hata, Y., Katsube, Y. & Suzuki, Y. (1991) Overproduction, purification and crystallization of *Bacillus cereuds* oligo-1,6-glucosidase, *FEBS Lett. 290*, 221–223.
17. Yanisch-Peron, C., Vieira, J. & Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC vectors, *Gene (Amst.) 33*, 103–119.
18. Vieira, J. & Messing, J. (1987) Production of single-stranded plasmid DNA, *Methods Enzymol. 153*, 3–11.
19. Kunkel, T. A., Roberts, J. & Zakour, R. (1987) Mutagenesis and protein engineering, *Methods Enzymol. 154*, 367–382.
20. Sanger, F. (1981) Determination of nucleotide sequences in DNA, *Science 214*, 1205–1210.
21. Yutani, K., Hayashi, S., Sugisaki, Y. & Ogasawara, K. (1991) Role of conserved proline residues in stabilizing tryptophane synthase α subunit: analysis by mutants with alanine or glycine, *Proteins Struct. Funct. Genet. 9*, 90–98.
22. Mayr, L. M., Landt, O., Hahn, U. & Schmid, F. X. (1993) Stability and folding kinetics of ribonuclease T1 are strongly altered by the replacement of *cis*-proline 39 with alanine, *J. Mol. Biol. 231*, 897–912.
23. Schultz, D. A. & Baldwin, R. L. (1992) Cis proline mutants of ribonuclease A. I. Thermal stability, *Protein Sci. 1*, 910–916.
24. Lascu, I., Deville-Bonne, D., Glaser, P. & Vèron, M. (1993) Equilibrium dissociation and unfolding of nucleoside diphosphate kinase from *Dictyostelium discoideum*, *J. Biol. Chem. 268*, 20268–20275.
25. Alber, T. (1989) Mutational effects on protein stability, *Annu. Rev. Biochem. 58*, 765–798.
26. Ramakrishnan, C. & Ramachandran, G. N. (1965) Stereochemical criteria for polypeptide and protein chain conformations, *Biophys. J. 5*, 909–933.
27. Efimov, A. V. (1986) Standard conformations of a polypeptide chain in irregular regions of proteins, *Mol. Biol. (Moscow) 20*, 250–260.
28. Schimmel, P. R. & Flory, P. J. (1968) Conformational energies and configurational statics of copolypeptides containing L-proline, *J. Mol. Biol. 34*, 105–120.
29. Hurley, J. H., Mason, D. A. & Matthews, B. W. (1992) Flexible-geometry conformational energy maps for the amino acid residue preceding a proline, *Biopolymers 32*, 1443–1446.
30. MacArther, M. W. & Thornton, J. M. (1991) Influence of proline conformation, *J. Mol. Biol. 218*, 397–412.
31. Chou, P. Y. & Fasman, G. D. (1974) Prediction of protein conformation, *Biochemistry 13*, 222–245.
32. Richardson, J. S. & Richardson, D. C. (1988) Amino acid preferences for specific locations at the ends of α-helices, *Science 240*, 1648–1652.
33. Argos, P. & Palau, J. (1982) Amino acid distribution in protein secondary structures, *Int. J. Pept. Protein Res. 19*, 380–393.
34. Yun, R. H., Anderson, A. & Hermans, J. (1991) Proline in α-helix: stability and conformation studied by dynamics simulation, *Proteins Struct. Funct. Genet. 10*, 219–228.
35. Levitt, M. (1978) Conformational preferences of amino acids in globular proteins, *Biochemistry 17*, 4227–4285.
36. Matthews, B. (1993) Structural and genetic analysis of protein folding and stability, *Curr. Opin. Struct. Biol. 1*, 17–21.
37. Nemethy, G., Leach, S. J. & Scheraga, H. A. (1966) The influence of amino acid side chains on the free energy of helix-coil transitions, *J. Phys. Chem. 70*, 998–1006.
38. Jaenicke, R. & Zàvodszky, P. (1990) Proteins under extreme conditions, *FEBS Lett. 268*, 344–349.
39. Wells, J. A. (1990) Additivity of mutational effects in proteins, *Biochemistry 29*, 8509–8517.
40. Privalov, P. & Tsalkova, T. N. (1979) Micro- and macro-stabilities of globular proteins, *Nature 280*, 692–696.
41. Jaenicke, R. (1991) Protein stability and molecular adaptation to extreme conditions, *Eur. J. Biochem. 202*, 715–728.
42. Argos, P., Rossman, M. G., Grau, U. M., Zuber, H., Frank, G. & Tratschin, J. D. (1979) Thermal stability and protein structure, *Biochemistry 18*, 5698–5703.

43. Creighton, T. E. (1984) in *Protein. Structures and molecular properties*, pp. 252–259, Freeman and Company, NY, USA.
44. Siezen, R. J., de Vos, W. M., Leunissen, J. A. M. & Dijkstra, B. W. (1991) Homology modeling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteases, *Protein Eng. 4*, 719–737.
45. Suzuki, Y., Hatagaki, K. & Oda, H. (1991) A hyperthermostable pullulanase produced by an extreme thermophile, *Bacillus flavocaldarius* KP1228, and evidence for the proline theory of increasing protein thermostability, *Appl. Microbiol. Biotechnol. 34*, 707–714.
46. Suzuki, Y. (1992) A general pronciple of increasing protein thermostability – The proline theory, in *Industrial biotechnology* (Malik, V. S. & Sridhar, P., eds) pp. 163–178, Oxford & IBH Publishing Co., New Delhi, India.
47. Wilmot, C. M. &. Thornton, J. M. (1988) Analysis and prediction of the different types of β turn in poroteins, *J. Mol. Biol. 203*, 221–232.