Cloning and thermostability of *TaqI* endonuclease isoschizomers from *Thermus* species SM32 and *Thermus filiformis* Tok6A1

Weiguo CAO*, Jing LU*, Simon G. WELCH†, Ralph A. D. WILLIAMS† and Francis BARANY*¹

*Department of Microbiology, Hearst Microbiology Research Center, Cornell University Medical College and Strang Cancer Prevention Center, New York, NY 10021, U.S.A., and †Department of Biochemistry, Faculty of Basic Medical Sciences, Queen Mary and Westfield College, University of London, Mile End Road, London E1 4NS, U.K.

Two *TaqI* endonuclease (hereafter referred to as *TaqI*) isoschizomer genes, *tsp32I^R* from *Thermus* species SM32 of Azores and *tfiTok6A1I* from *T. filiformis* Tok6A1 of New Zealand, were cloned in *Escherichia coli*. The overexpressed enzymes were partly purified and their thermostability was determined. In the medium-salt buffer, *Tsp32I^R*, *TfiTok6A1I* and one previously cloned *TaqI* isoschizomer (*TthHB8I*) were more thermostable than *TaqI*. *Tsp32I^R* remained partly active up to 90 °C in the low-salt buffer. Six amino acid residues that are identical in the three high thermostability isoschizomers (*Tsp32I^R*, *TfiTok6A1I* and

TthHB8I) but differ in *TaqI* might provide added rigidity for thermostabilization. These include four proline residues located in or near loop regions, and one alanine and one arginine located at helix regions in the predicted *TaqI* endonuclease secondary structure. The possible role of these residues in thermostabilization was evaluated by mutagenizing the *TaqI* enzyme. Mutants generated at these six positions were less thermostable than wild-type *TaqI*. The results suggest that the surrounding sequence or structural context might be as important as the mutation itself.

INTRODUCTION

TaqI endonuclease (hereafter referred to as *TaqI*) activity was observed in *Thermus* strains isolated in the United States, Japan, New Zealand, Portugal and Azores [1]. On the basis of the nucleotide sequences, we classified the eight *TaqI* isoschizomer genes into four geographic groups: USA group as represented by *TaqI*, Japan by *TthHB8I*, New Zealand by *TfiTok6A1I*, and Portugal and Azores by *Tsp32I^R*. Overall, the amino acid sequences show considerable divergence, ranging from 54 % to 75 % identity among the eight isoschizomers. Yet within the same homology group, the sequences are virtually identical with each other, with either no differences at all or only a few point mutations [1].

Although the protein sequences are generally conserved and amino acid residues important for enzyme activity are highly conserved, their biochemical properties can vary significantly. For example, *TaqI* and *TthHB8I* share 75 % sequence identity, but *TthHB8I* is more thermostable than *TaqI* because the latter readily loses its activity at 70 °C. A *Thermus* strain isolated from Azores contains two fractions of *TaqI* isoschizomer that differ in optimal concentrations of metal cofactor, thermostabilities and DEAE-Sephacel elution profiles [2]; they were subsequently named *Tsp32I* and *Tsp32II*. Hence although these homologous *TaqI* isoschizomers use the same catalytic mechanism, the 30–45 % amino acid difference distal from the active site could offer sufficient subtle changes to render one isoschizomer more thermostable than another.

The genes encoding *TaqI* and *TthHB8I* have been cloned and overexpressed in *Escherichia coli* [3–5]. Here we describe the cloning and expression of *TfiTok6A1I* and *Tsp32I^R* from the two new geographic groups (New Zealand and Portugal). The thermostability of the cloned *TfiTok6A1I* and *Tsp32I^R* were compared with the cloned *TaqI* and *TthHB8I*, as well as with the native *Tsp32I* and *Tsp32II*. The thermostabilities are in the order *Tsp32I^R*, *Tsp32II* > *TfiTok6A1I*, *TthHB8I*, *Tsp32I* > *TaqI* in the medium-salt buffer. *Tsp32I^R* and *Tsp32II* are partly active up to 90 °C in the low-salt buffer. Amino acid sequence alignment reveals that six amino acid replacements (four to proline, one to alanine, one to arginine) are identical among the isoschizomers with elevated thermostability (*Tsp32I^R*, *TfiTok6A1I*, and *TthHB8I*). The possible role of these residues was evaluated by site-directed mutagenesis. The resultant 12 *TaqI* mutants maintained wild-type or close to wild-type enzyme activity, but their thermostability varied considerably depending on the position and type of mutation.

EXPERIMENTAL

Reagents, media and strains

All routine chemical reagents were purchased from Sigma Chemicals (St. Louis, MO, U.S.A.) or Fisher Scientific (Fair Lawn, NJ, U.S.A.). Restriction enzymes, T4 DNA ligase, T4 polynucleotide kinase and Klenow fragment of *E. coli* DNA polymerase were obtained from New England Biolabs (Beverly,

Abbreviations used: DTT, dithiothreitol; *TaqI*, *TaqI* endonuclease; *taqI*, *TaqI* endonuclease gene; *taqIM*, *TaqI* methylase gene; *tfiTok6A1I*, *TaqI* isoschizomer gene in *Thermus filiformis* Tok6A1; *Tsp32I^R*, *TaqI* isoschizomer in *Thermus* species SM32; *tsp32I^R*, gene encoding *Tsp32I^R*; *Tsp32I*, native *Tsp32I* protein purified from *Thermus* species SM32; *Tsp32II*, native *Tsp32II* protein purified from *Thermus* species SM32 (we name the *TaqI* isoschizomer gene that we identified in *Thermus* species SM32 *tsp32I^R* and its gene product *Tsp32I^R* to distinguish them from natively purified *Tsp32I* and *Tsp32II*).

¹ To whom correspondence should be addressed (e-mail barany@mail.med.cornell.edu).

The nucleotide sequence data reported in this paper for *TfiTok6A1I* and *Tsp32I^R* appear in the GenBank Database under the accession numbers U86869 and U86870 respectively.

MA, U.S.A.). Oligonucleotide synthesis reagents, DNA sequencing kits and PCR kits were obtained from the Applied Biosystems Division of Perkin-Elmer Corporation (Foster City, CA, U.S.A.). TE buffer contained 10 mM Tris/HCl, pH 8.0, and 1 mM EDTA. Agarose gel loading buffer (10×) consisted of 50% (v/v) Ficoll, 0.1 M EDTA, 1% (w/v) SDS, 0.25% Bromophenol Blue and 0.25% Xylene Cyanol; 2× *TaqI* storage buffer consisted of 50 ml of glycerol, 1 ml of 1 M Tris/HCl, pH 8.0, 200 µl of 0.5 M EDTA, pH 8.0, 1 ml of 100 mM dithiothreitol (DTT) and 20 mg of BSA. TC medium (pH 9.0), Fortified Broth and Mops medium were prepared as described [3,6,7]. The plasmid pFB76 containing a single *TaqI* recognition site was used for *TaqI* endonuclease activity assay [8]. *E. coli* host strain AK53 (*mrrB*[−], MM294) was from our laboratory collection.

Cloning, expression and partial purification

A polylinker sequence, 5'-GGA TCC AGC TGG TAC CAA GCT TCT GCA G-3', containing *Bam*HI, *Pvu*II, *Kpn*I, *Hind*III and *Pst*I sites was inserted after the stop codon of the *taqI* gene in pBLT88 [3] by overlap extension PCR method [9]. The resulting plasmid was designated pEx*TaqR*.1. The isoschizomer genes were amplified from genomic DNA by a PCR procedure described previously [1]. The DNA sequence of the *phoA* signal peptide was amplified from pEx*TaqR*.1 (see Figure 1) by two-step PCR: denaturation at 94 °C for 30 s, then annealing and extension at 60 °C for 1.5 min for a total of 25 cycles. The PCR mixture (100 µl) consisted of 50 ng of pEx*TaqR*.1 as template, 0.6 µM Iso*Taq*-clon3R and Iso*Taq*-clon1 as primers, 1× PCR buffer II, each dNTP at 50 µM, 1.5 mM MgCl₂ and 2.5 units of Ampli*Taq* DNA polymerase (Figure 1). The *phoA* signal peptide was fused into the N-terminus of the isoschizomer genes by overlap extension PCR. The PCR reaction was performed in the two-step format as described above except that the template was replaced by 1 µl of each PCR solution from the isoschizomer gene and the *phoA* signal peptide amplification mixtures, and Iso*Taq*-clon1 and Iso*Taq*-clon2R were used as primers. The *taqI* gene in pEx*TaqR*.1 was replaced by isoschizomer genes by conventional restriction digestion and ligation procedures [10]. Plasmids containing the new isoschizomer inserts were transformed into *E. coli* AK53 competent cells. The complete open reading frames of *tsp32I^R* and *tfiTok6A1I* were sequenced to ensure that the plasmid constructs were free of PCR error. Expression of these genes was induced by growing cells in low-phosphate Mops medium. Cell pastes from 15 ml of culture were suspended in sonication buffer consisting of 10 mM Tris/HCl, pH 8.0, 1 mM EDTA, 10 mM DTT, 150 µM PMSF and 200 mM NaCl. After two cycles of sonication, crude protein extracts were separated from cell debris by centrifugation, then heat treated at 65 °C for 15 min. Denatured *E. coli* proteins were removed by centrifugation. Partly pure *TaqI* isoschizomer proteins were mixed with an equal volume of 2× *TaqI* storage buffer and stored at −20 °C.

Purification of native *Tsp32I* and *Tsp32II*

Tsp32I and *Tsp32II* were partly purified from 16 g of *Thermus* species SM32 cells [2]. One unit of restriction endonuclease activity is defined as the amount of enzyme required to completely digest 1 µg of λ DNA at 65 °C in 'one-Phor-All' universal buffer from Pharmacia (Piscataway, NJ, U.S.A.).

Enzyme activity assay

To quantify the amount of expressed *TaqI* or its isoschizomers, partly purified protein samples were separated by SDS/PAGE

and revealed by standardized Coomassie stain along with pure *TaqI* protein previously purified by chromatographic methods [3]. The intensities of the protein bands were captured by scanning the dried protein gels and converted into molar concentration by using pure *TaqI* protein as standard. Partly purified enzymes (2 nM) were incubated with 20 nM pFB76 in the medium-salt buffer [10 mM Tris/HCl buffer (pH 8.0 at 60 °C)/0.1 mg/ml BSA/50 mM NaCl/10 mM DTT/10 mM MgCl₂] in a final volume of 10 µl for 5 min at 60 °C to achieve limited digestion. When 2 mM MnCl₂ was used as the metal cofactor, the reaction mixtures were incubated for 40 min at 60 °C. The reaction was quenched by adding 2 µl of 10× agarose gel loading buffer. Linearized plasmids were separated from supercoiled plasmids on a 1% (w/v) agarose gel.

Thermostability assay

To evaluate thermostability, partly purified enzymes were incubated in medium-salt (50 mM NaCl) or low-salt (0 mM NaCl) buffer in the absence of the metal cofactor and plasmid substrate at designated temperatures for 5 min. The reaction mixtures were supplemented with 20 nM pFB76 and 10 mM MgCl₂ to measure the remaining *TaqI* activity under the conditions described above.

Site-directed mutagenesis

Six positions in *TaqI* were chosen for mutational study: Thr¹², Ala⁵⁶, Leu⁶⁴, Lys¹⁶⁶, Glu²²¹, Arg²⁵¹ (see Figure 5). Two types of mutation were designed, one of which reverted *TaqI* to *Tsp32IR*. An overlap extension PCR procedure was used for site-directed mutagenesis [1,9]. Partly degenerated internal primers were designed to generate two types of mutation at a particular site. PCR products, digested with *Bgl*II and *Mlu*I, or with *Mlu*I and *Kpn*I, were ligated to cloning vector pEx*TaqR*.1 treated with the same pair of restriction endonucleases. The ligated vectors were transformed into *E. coli* strain AK53 (*mrrB*[−], MM294). Plasmids containing inserts were reisolated and sequenced on an ABI sequencer by using dye-deoxy terminator chemistry to identify the mutated sequence and to ensure that the constructs were free of PCR or ligation error. Mutant proteins were expressed and partly purified as described above. The protein concentration was estimated by SDS/PAGE.

RESULTS AND DISCUSSION

Growth of *Thermus* at different temperatures

Thermus species were isolated from hot spring environments with temperatures ranging from 53 to 85 °C. Under enriched culture conditions, *Thermus* isolates grow at temperatures ranging from 40 to 79 °C with an optimum at 70 °C [11]. *T. aquaticus* YT1, *T. thermophilus* HB8, *T. filiformis* Tok6A1 and *Thermus* species SM32, which contain *TaqI* endonuclease or its isoschizomers, were originally isolated in the United States, Japan, New Zealand and Azores, respectively. Growth in TC medium at pH 9.0 was monitored at four different temperatures (Table 1). All four strains accumulated substantial cell mass after overnight cultivation at 70–80 °C. *T. filiformis* Tok6A1 was the only strain that could grow at a higher temperature up to 85 °C. *T. aquaticus* YT1 generated yellow pigments (carotenoid) as previously documented [11] at 70–75 °C, but not at 80 °C, indicating either that the pathway was switched off or the gene product(s) became inactive. As shown later, *TaqI* endonuclease isolated from *T. aquaticus* YT1 was inactive at 80 °C in the medium salt *in vitro*. As a non-essential gene, its change in expression and activity might not correlate closely with cellular viability. The two new endonuclease genes, *tsp32I^R* and *tfiTok6A1I*, representing two

Table 1 Growth of *Thermus* bacteria at different temperatures

Thermus was grown in TC medium at indicated temperatures with shaking at 125 rev./min. Symbols: +, significant growth was observed in overnight culture; —, growth was not observed after 72 h of incubation. For the results in the last column, the nucleotide sequences of *TaqI* and its isoschizomers were determined and their corresponding genes expressed in *E. coli*.

Strain	Growth at 70 °C	Growth at 75 °C	Growth at 80 °C	Growth at 85 °C	Pigmentation	Expression of endonuclease in <i>E. coli</i>
<i>T. aquaticus</i> YT1	+	+	+	—	Yellow	[3]
<i>T. thermophilus</i> HB8	+	+	+	—	Yellowish	[5]
<i>T. filiformis</i> Tok6A1	+	+	+	+	Yellowish	This study
<i>Thermus</i> species SM32	+	+	+	—	Yellowish	This study

Figure 1 Cloning of *TaqI* isoschizomer gene from *Thermus* species SM32

Seven amino acid residues (part of *phoA* signal peptide, short striped bar in the *taqI* gene) were fused into the N-termini of isoschizomer genes by overlap extension PCR. The resulting fragments were digested with *Bam*HI and then blunt-ended with the Klenow fragment of *E. coli* DNA polymerase. The 5' overhang was generated by *Bgl*II digestion. The vector pExTaqR.1 was digested with *Pvu*II and *Bgl*II to delete the *taqI* gene, then ligated with the PCR fragments. A similar strategy was used for cloning *tfiTok6A1I*. See the text for details. Abbreviations: *amp*^R, ampicillin resistance gene; MCS, multiple cloning sites.

distinct homology groups [1], were cloned and expressed in *E. coli* by this work.

Cloning of *tsp32I^R* and *tfiTok6A1I*

The complete open reading frames of the two new isoschizomer genes, *tsp32I^R* and *tfiTok6A1I*, have been determined [1]. Their unusually high divergence from the previously studied *TaqI* and *TthHB8I* calls for further characterization. Additionally, one of the natively purified *TaqI* isoschizomers from *Thermus* species SM32, *Tsp32II*, has been shown to be more thermostable than any of the known *TaqI* isoschizomers [2]. The cloning scheme shown in Figure 1 illustrates the recombinant DNA procedures used for cloning the *tsp32I^R* gene. The *tsp32I^R* gene was readily

amplified by PCR with primers IsoTaq-clon2R and IsoTaq-clon4 (Figure 1). To facilitate protein expression [3], a seven-residue *phoA* signal peptide sequence (MKQSGGN) was amplified by PCR with primers IsoTaq-clon3R and IsoTaq-clon1, and fused into the N-terminus of the *tsp32I^R* gene by overlap extension PCR [9]. To facilitate directional cloning, a plasmid containing both *taqI* and *TaqI* methylase (*taqIM*) genes (pFBLT88; for plasmid map see [3]) was reconstructed so that a polylinker sequence was inserted immediately after the stop codon of the *taqI* gene. The replacement of the *taqI* gene in pExTaqR.1 with the *tsp32I^R* gene gave rise to plasmid pTSP32I.1. *TaqI* methylase encoded by the *taqIM* gene upstream of the *tsp32I^R* gene helps to prevent *Tsp32I^R* from nicking or cleaving the host genome. The *tfiTok6A1I* gene from *T. filiformis* Tok6A1 was cloned into

Figure 2 SDS/PAGE analysis of *TaqI* isoschizomers

Partly purified protein samples were subjected to SDS/PAGE [12.5% (w/v) gel]. Protein bands were detected by the routine Coomassie staining method. *TaqI* and the three isoschizomer proteins are indicated by arrows. Sizes of molecular mass standards are indicated (in kDa) at the left. Lane A, molecular mass standards; lane B, pure *TaqI* protein; lane C, negative control (partly purified *E. coli* proteins without *TaqI* or its isoschizomer proteins); lane D, partly purified *TaqI* protein as indicated; lane E, partly purified *TthHB8I* protein; lane F, partly purified *TfiTok6A1I* protein; lane G, partly purified *Tsp32I^R* protein. Both *TaqI* and *TfiTok6A1I* proteins migrated more slowly than their actual molecular masses suggest.

E. coli by a similar approach. The transcription of *tsp32I^R* and *tfiTok6A1I* is regulated by the *phoA* promoter, which can be switched on under phosphate-deficient culture conditions.

Expression and partial purification

Under low-phosphate growth conditions (Mops medium), *Tsp32I^R* and *TfiTok6A1I* accounted for approx. 5% of total soluble proteins. The majority of host *E. coli* proteins were separated from *Tsp32I^R* or *TfiTok6A1I* by thermal denaturation at 60 °C for 15 min followed by centrifugation. Although all four proteins have a calculated molecular mass of approx. 32 kDa, *TfiTok6A1I* migrated more slowly than expected on SDS/PAGE (Figure 2), in a manner identical to that of *TaqI* [3]. *Tsp32I^R* and *TthHB8I*, however, migrated normally at approx. 32 kDa. This phenomenon does not seem to be correlated with the degree of amino acid homology, as *TaqI* shares the least identity with *TfiTok6A1I* of the three isoschizomers [1]. The thermostabilities of the four isoschizomers discussed later are also not correlated with the proteins' mobility.

Enzyme activity

The partly purified enzymes were assayed for *TaqI* enzyme activity by using a plasmid containing a single *TaqI* site (Figure 3). A single-strand break nicks the plasmid, whereas a double-strand break should linearize the plasmid. The assay conditions (enzyme-to-substrate ratio 1:10) were designed to achieve limited plasmid cleavage so that differences in enzyme activity could be observed. The minor protein impurity from *E. coli* host did not interfere with *TaqI* activity assays, suggesting that the partly purified enzymes were essentially free of host nucleases (results not shown). In a medium-salt buffer (50 mM NaCl) with either Mg^{2+} or Mn^{2+} as the metal cofactor, *TaqI* was enzymically the most active of the four proteins, followed by *TthHB8I* and

Figure 3 Enzyme activity of *TaqI* and its isoschizomers

Partly purified enzymes (2 nM) were incubated with 20 nM pFB76 in the medium-salt buffer: (A) assayed at 60 °C for 5 min with 10 mM $MgCl_2$; (B) assayed at 60 °C for 40 min with 2 mM $MnCl_2$. Lane 1, undigested pFB76 plasmid as control; lane 2, pFB76 digested with *TaqI*; lane 3, pFB76 digested with *TthHB8I*; lane 4, pFB76 digested with *TfiTok6A1I*; lane 5, pFB76 digested with *Tsp32I^R*.

TfiTok6A1I; *Tsp32I^R* was the least active (Figure 3). This disparity might be due to discrete pH or temperature optima, salt requirements or differences in catalytic efficacy. For example, *Tsp32II* purified from *Thermus* species SM32 has a slightly higher pH optimum than *TaqI* and a narrow optimal range of Mg^{2+} cofactor concentration [2]. The replacement of Mg^{2+} with Mn^{2+} lowered the cleavage activity of all four enzymes to approx. one-seventh as demonstrated by the requirement of extended incubation (40 min) with Mn^{2+} . This result is in keeping with the previously observed k_{cat} difference of *TaqI* enzyme in the two metal cofactor buffer conditions [12], indicating that the enzymic mechanism for DNA cleavage is similar for the four enzymes.

Thermostability

Thermostability of *TaqI* and its three isoschizomers was initially determined at 70–90 °C with 5 °C increments in a medium-salt buffer (50 mM NaCl). *TaqI* lost most of its activity after heat inactivation at 70 °C for 5 min (Figure 4A). *TthHB8I* and *TfiTok6A1I* retained some activity after incubation at 80 °C; *Tsp32I^R* was partially active after incubation at 85 °C. *Tsp32I* purified from *Thermus* sp. SM32 was as thermostable as *TthHB8I* and *TfiTok6A1I*, whereas natively purified *Tsp32II* demonstrated similar thermostability to that of cloned *Tsp32I^R*. We further compared the thermostabilities of *Tsp32I^R* and *Tsp32II* at 88 and 90 °C in medium-salt and low-salt buffer (0 mM NaCl). Both proteins were less thermostable in the medium-salt buffer (Figure 4B). At 88 °C partial enzyme activities were retained after incubation in the low-salt buffer and to a smaller extent in the medium-salt buffer. At 90 °C they were partially active only after incubation in the low salt. Thus the recombinant *Tsp32I^R* and native *Tsp32II* were at least 15 °C more thermostable than *TaqI* and 5–10 °C more thermostable than *TthHB8I* and *TfiTok6A1I*. It is therefore likely that *Tsp32II* is the native gene product of *tsp32I^R*, but ultimate proof will need further verification, for

Figure 4 Comparison of thermostabilities

(A) Comparison of thermostabilities of *TaqI* isoschizomers in medium-salt buffer. The partly purified cloned and native proteins ranging from 0.04 to 0.25 unit were incubated at the indicated temperature for 5 min, then the remaining *TaqI* activity was assayed as described in the Experimental section with 10 mM $MgCl_2$. Lanes C, control reactions without exposing the enzymes to thermal inactivation. (B) Comparison of thermostabilities of the cloned *Tsp32I^R* and the native *Tsp32II* in medium-salt and low-salt buffers at 88 and 90 °C.

example N-terminal sequencing of highly purified *Tsp32I* and *Tsp32II*.

Although *TaqI*, *TthHB8I*, *TfiTok6A1I* and *Tsp32I^R* share more than 50% amino acid sequence identity, they differ considerably in their thermostabilities. The catalytic activity of the enzymes seems to be inversely proportional to their thermostability (Figures 3 and 4), i.e. the more thermostable the enzyme is, the less catalytically active it becomes. This observation supports the hypothesis that when a thermostable protein gains rigidity in its three-dimensional structure it decreases its catalytic efficiency because the conformational flexibility required for catalysis is decreased. Crystallographic studies of *EcoRV* and *BamHI* endonucleases demonstrate that these enzymes undergo considerable conformational changes on substrate binding [13,14]. Similar changes are expected for *TaqI* endonuclease and have been implicated by kinetic analysis [12] and by Fe^{2+} -mediated cleavage analysis of *TaqI* in the presence and the absence of DNA substrate (W. Cao and F. Barany, unpublished work). Hence the increased rigidity of *TaqI* protein might result in decreased DNA cleavage activity.

Some early studies on the mechanisms of thermostability compared the differences in amino acid composition of proteins from thermophilic and mesophilic sources [15,16]. A number of thermostability parameters have been proposed on the basis of these analyses, which suggest that some secondary and tertiary structural elements, such as hydrophobic index, α -helix and β -sheet contents, could alter the thermostability of proteins. Argos et al. [17] summarized the general direction of amino acid shuffling between thermophilic and mesophilic proteins, which showed the trend of substitutions as Gly to Ala, Ser to Ala, Ser to Thr, Lys to Arg, and Asp to Glu from mesophiles to thermophiles. Although these rules could provide a general trend

Table 2 Thermostability of *TaqI* mutants

For relative activities, the partly purified enzymes were assayed for DNA cleavage activity in medium-salt buffer at 60 °C with no thermal denaturation. For relative thermostabilities, the partly purified enzymes were incubated at 80 °C for 5 min in low-salt buffer, then assayed for DNA cleavage activity in medium-salt buffer at 60 °C; see the Experimental section for details. The enzyme activity and thermostability of the wild-type enzyme were taken as 100%.

Mutant	Relative activity (%)	Relative thermostability (%)
<i>TaqI</i> wild-type	100	100
<i>TaqI</i> -T12A	100	80
<i>TaqI</i> -T12S	100	80
<i>TaqI</i> -A56P	100	80
<i>TaqI</i> -A56S	100	75
<i>TaqI</i> -L64P	100	90
<i>TaqI</i> -L64Q	100	80
<i>TaqI</i> -K166R	100	80
<i>TaqI</i> -K166Q	100	10
<i>TaqI</i> -E221P	100	65
<i>TaqI</i> -E221Q	80	70
<i>TaqI</i> -R251P	100	70
<i>TaqI</i> -R251H	100	10

towards higher thermostability from mesophiles, we did not find them applicable to the intra-genus thermostability difference of *TaqI* isoschizomers. Because the thermostability differences are in a relatively narrow temperature range, the macroscopic parameters derived from comparisons of mesophilic and thermophilic proteins might not reveal the subtlety of amino acid changes relevant to thermostability. Jaenicke and co-workers

Figure 5 Sequence alignment of *TaqI* and its isoschizomer proteins

Amino acid residues differing from *TaqI* are boxed. Amino acid residues identical in the four isoschizomers have been deleted for clarity. Dashes have been introduced to achieve maximal alignment. Amino acid residues identical between the three isoschizomers but different from *TaqI* are indicated by ▼. The secondary structure assignment of *TaqI* endonuclease was predicted by the PHD method [32]. Abbreviations: α, α-helix; β, β-strand; L, loop. GenBank accession numbers: U86869 for *TfiTok6A11*; U86870 for *Tsp32I^R*.

[18–20] have also questioned these rules on the basis of sequence statistics of D-glyceraldehyde-3-phosphate dehydrogenase.

To understand the structural basis for increased thermostability of *Tsp32I^R*, *TfiTok6A11* and *TthHB8I*, we sought to identify the amino acid residues that were identical in the three isoschizomers, but different from *TaqI*. Because these three isoschizomers were more thermostable than *TaqI* and evolved from a common ancestor, amino acid residues that favoured a high thermostability could have been preserved through evolution. As shown in Figure 5, six such sites were identified: four of them were proline residues (Ala⁵⁶ → Pro, Leu⁶⁶ → Pro, Glu²²¹ → Pro and Arg²⁵¹ → Pro), one was arginine (Lys¹⁶⁶ → Arg) and one was alanine (Thr¹² → Ala).

Site-directed mutagenesis of *TaqI*

Given the potential role of these residues in stabilizing the more thermostable isoschizomers, we used the site-directed mutagenesis approach to study further the possible involvement of alanine, proline and arginine residues in thermostabilization. We made two types of change for each residue: one reverted to *TaqI* to become identical to the three more thermostable isoschizomers at a particular position, and the other change was a new amino acid residue not found in any of the four proteins. Our results are summarized in Table 2 and discussed below.

Mutations at Thr¹² and Lys¹⁶⁶

Both mutants at Thr¹² were less stable than the wild-type enzyme (Table 2). These results suggest that Thr¹² might exert a subtle stabilizing effect that could not be simply explained by helix propensities [21]. Studies on barnase stability by Fersht's group has led to the conclusion that it is invalid to assign to each amino acid a unique *s*-value (helix-forming propensities) that is generally applicable to all positions in all helices in all proteins [22].

Sequence comparison suggests a high frequency of Lys to Arg replacement between mesophilic and thermophilic molecules [17]. More thorough sequence comparison concludes that decreased flexibility and increased hydrophobicity in α-helical regions are the main stabilizing principles [23]. The Lys¹⁶⁶ residue is located in the middle of a 14-residue region exhibiting a strong tendency to form an α helix according to molecular modelling (Figure 5). To test the possible role of Lys and Arg in protein stabilization, Lys¹⁶⁶ in *TaqI* was mutated to Arg and Gln. Lys¹⁶⁶ → Arg was less thermostable than the wild-type *TaqI* but more thermostable than Lys¹⁶⁶ → Gln (Table 2). It seems that Arg might be the 'best choice' for *Tsp32I^R*, but a 'good choice' not necessarily the 'best choice' for *TaqI* because of the difference in sequence context.

Mutations at Ala⁵⁶, Leu⁶⁴, Glu²²¹ and Arg²⁵¹

Proline is known as a protein stabilizer by virtue of increased

conformational rigidity, thereby providing a favourable entropic effect against protein denaturation [24]. On the basis of linearity of proline content and thermostability observed in *Bacillus* oligo-1,6-glucosidases, a 'proline theory' was initially proposed by Suzuki [25]; later, sequence comparison and site-directed mutagenesis studies confirmed the critical roles of proline in rigidifying the β -turns or coils in the loop regions [26–28].

We made proline substitutions at the four positions (Ala⁵⁶, Leu⁶⁴, Glu²²¹ and Arg²⁵¹) in *TaqI*. Thermostability comparisons showed that all of them were less stable than the wild-type enzyme at 80 °C (Table 2). All the mutants were as active as wild-type enzyme except for Glu²²¹ → Gln, which was presumably located at the C-cap of a long helix region (Figure 5). Although Pro could strengthen the loop region to deter thermal denaturation, the positive thermostabilization effect of Pro substitutions might be undermined by highly variable surrounding sequences. A closer look of the flanking sequences of Ala⁵⁶ is revealing (Figure 5). In *TthHB8I*, PNG forms an ideal β -turn when Pro takes the second position $i+1$ (see [29] for definition). Pro-Asn in *Tsp32I^R* and Pro-Gly in *TfiTok6AII* are also suitable for β -turn structure [29]. In *TaqI*, however, Aap-Leu-Glu might form a less thermoresistant β -turn or loop. Data from multiple proline substitutions in oligo-1,6-glucosidase based on the sequence alignment of *Bacillus cereus* and *B. thermoglucosidasius* also demonstrate that the 'proline effect' might be influenced by flanking sequences [27]. Proline replacements Asn¹⁰⁹ → Pro and Lys¹²¹ → Pro in a highly conserved region resulted in a positive ΔT_m of 1.4 °C, whereas a proline replacement of Glu²⁷⁰ in a highly variable region decreased the T_m by 0.3 °C [27]. In *E. coli* RNase H, a single Gln¹¹³ → Pro change within the R7 loop decreased the T_m by approx. 2 °C, but replacement of the loop region (R7) increased the T_m by 2.4 °C at pH 5.5 [30]. A recent analysis of β -turn propensities also emphasizes the need for coherence between a particular residue and its structural motif for engineering stability in an immunoglobulin variable domain [31]. To enhance the stabilization effect of proline in *TaqI*, additional changes surrounding the proline substitutions might be necessary.

Concluding remarks

We have cloned and expressed two new *TaqI* isoschizomers in *E. coli*. *Tsp32I^R* is at least 15 °C more thermostable than *TaqI*. The variation in thermostability caused by 30–45 % sequence difference could not be simply explained by empirical rules derived from comparisons of amino acid composition of proteins from thermophilic and mesophilic sources. As our first attempt at understanding the underlying principle of *TaqI* and its isoschizomers' thermostability, we have mutagenized six positions in *TaqI* endonuclease that are identical in the three more thermostable isoschizomers, *Tsp32I^R*, *TfiTok6AII* and *TthHB8I*. In every case the corresponding amino acid from the hyperstable protein does not confer additional stability on *TaqI*. These results suggest that when placing a thermo-favourable residue into a particular site, its sequence or structural context has to be

compatible. More structural and biochemical studies on *TaqI* (for example, changing the flanking sequences of the six mutants) and its isoschizomers (for example, mutating the corresponding six positions in *Tsp32I^R*) will deepen our understanding of the underlying mechanism of *Tsp32I^R* thermostability.

W.C. was a Norman and Rosita Winston Foundation fellow in Biomedical Research. This work was supported by grants from the National Institutes of Health (GM-41337-07) and Applied Biosystems Division of Perkin-Elmer Corporation.

REFERENCES

- Cao, W., Lu, J. and Barany, F. (1997) *Gene* **197**, 205–214
- Welch, S. G. and Williams, R. A. D. (1995) *Biochem. J.* **312**, 505–510
- Barany, F. (1988) *Gene* **63**, 167–177
- Barany, F., Slatko, B., Danzitz, M., Cowburn, D., Schildkraut, I. and Wilson, G. G. (1992) *Gene* **112**, 91–95
- Barany, F., Danzitz, M., Zebala, J. and Mayer, A. (1992) *Gene* **112**, 3–12
- Neidhardt, F. C., Bloch, P. L. and Smith, D. F. (1974) *J. Bacteriol.* **119**, 736–747
- Ramaley, R. F. and Hixson, J. (1970) *J. Bacteriol.* **103**, 526–528
- Barany, F. (1987) *Gene* **56**, 13–27
- Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. and Pease, L. R. (1989) *Gene* **77**, 51–59
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Brock, T. D. and Freeze, H. (1969) *J. Bacteriol.* **98**, 289–297
- Cao, W., Mayer, A. N. and Barany, F. (1995) *Biochemistry* **34**, 2276–2283
- Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K. and Wilson, K. S. (1993) *EMBO J.* **12**, 1781–1795
- Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I. and Aggarwal, A. K. (1995) *Science* **269**, 656–663
- Singleton, Jr., R., Middaugh, C. R. and MacElroy, R. D. (1977) *Int. J. Pept. Protein Res.* **10**, 39–50
- Merkler, D. J., Farrington, G. K. and Wedler, F. C. (1981) *Int. J. Pept. Protein Res.* **18**, 430–442
- Argos, P., Rossman, M. G., Grau, U. M., Zuber, H., Frank, G. and Tratschin, J. D. (1979) *Biochemistry* **18**, 5698–5703
- Bohm, G. and Jaenicke, R. (1994) *Int. J. Pept. Protein Res.* **43**, 97–106
- Korndorfer, I., Steipe, B., Huber, R., Tomsch, A. and Jaenicke, R. (1995) *J. Mol. Biol.* **246**, 511–521
- Jaenicke, R., Schurig, H., Beaucamp, N. and Ostendorp, R. (1996) *Adv. Protein Chem.* **48**, 181–269
- Chakrabarty, A. and Baldwin, R. L. (1995) *Adv. Protein Chem.* **46**, 141–176
- Serrano, L., Neira, J. L., Sancho, J. and Fersht, A. R. (1992) *Nature (London)* **356**, 453–455
- Menendez-Arias, L. and Argos, P. (1989) *J. Mol. Biol.* **206**, 397–406
- Matthews, B. W., Nicholson, H. and Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6663–6667
- Suzuki, Y., Oishi, K., Nakano, H. and Nagayama, T. (1987) *Appl. Microbiol. Biotechnol.* **26**, 546–551
- Watanabe, K., Chishiro, K., Kitamura, K. and Suzuki, Y. (1991) *J. Biol. Chem.* **266**, 24287–24294
- Watanabe, K., Masuda, T., Ohashi, H., Mihara, H. and Suzuki, Y. (1994) *Eur. J. Biochem.* **226**, 277–283
- Watanabe, K., Kitamura, K. and Suzuki, Y. (1996) *Appl. Environ. Microbiol.* **62**, 2066–2073
- Creighton, T. E. (1993) *Proteins: Structures and Molecular Properties*, W. H. Freeman, New York
- Kimura, S., Nakamura, H., Hashimoto, T., Oobatake, M. and Kanaya, S. (1992) *J. Biol. Chem.* **267**, 21535–21542
- Ohage, E. C., Graml, W., Walter, M. M., Steinbacher, S. and Steipe, B. (1997) *Protein Sci.* **6**, 233–241
- Rost, B. and Sander, C. (1993) *J. Mol. Biol.* **232**, 584–599