Available online at www.sciencedirect.com
SCIENCE DIRECT
Enzyme and Microbial Technology 37 (2005) 402–409
ENZYME AND MICROBIAL TECHNOLOGY
www.elsevier.com/locate/emt

Biochemical characterization of a novel thermostable glucose-1-phosphate thymidylyltransferase from *Thermuscaldophilus*: Probing the molecular basis for its unusual thermostability

Niranjan Parajuli [a], Devi B. Basnet [a], Young Soo Chung [b], Hei Chan Lee [a], Kwangkyoung Liou [a], Jae Kyung Sohng [a,*]

[a] *Institute of Biomolecule Reconstruction, Sun Moon University, #100 Kalsan-ri, Tangjeon-myun, Asan-si, Chungnam 336-708, Republic of Korea*
[b] *GeneChem Corp. 461-6, Jeonmin-Dong, Yusong-Gu, Daejon 305-390, Republic of Korea*

Received 7 June 2004; accepted 2 February 2005

Abstract

We have found that the thermophilic activity of glucose-1-phosphate thymidylyltransferase (stRmlA) is conditional upon the presence certain substances. In particular, it showed its thermal stability only in the presence of 50 mM dTTP or dTMP, thermal stability being the highest at 70 °C. The purified enzyme was stable up to 90 °C within a broad pH range from 2.0 to 13.0, and its maximum activity was measured at a pH of 11.5 at 70 °C. Unlike other mesophilic counterparts, it showed catalytic activity in the presence of various metal ions in the following order of reactivity: $Mg^{2+} > Zn^{2+} > Cu^{2+} > Co^{2+} > Fe^{2+} > Ca^{2+} > Fe^{3+} > Ni^{2+}$. Its catalytic activity was not inhibited even by the denaturants 50 mM guanidine hydrochloride and 50 mM urea.

To explore the molecular basis for its unusual thermostability, homology structural modeling, codon usage comparisons, and amino acid composition analyses were performed. The CCC, CUC and UCC codons are significantly higher, and distributed uniformly in *strmlA* as contrasted with other GPTTs. Our analyses revealed that the GC rich coding sequences may not be the reason for thermostability of stRmlA. The β-sheets are also found more likely in stRmlA, which contributes to thermal stability by increasing the number of hydrogen bonds. Further results suggest that a large number of apolar functional groups exposed to solvent accessible surface area, a significant number of residues sensitive to oxidation or deamination, and a higher hydrophobicity, any or all of the which could be reasons for its unique thermophilic behavior. Finally, we postulate that dTTP or dTMP enters the active site and binds tightly to inhibit the defolding of the protein, which, in turn, increases its thermal stability. Such findings will be useful for further investigations on thermophilic behavior of enzymes.

Keywords: Codon usage; dTTP; Glucose-1-phosphate thymidylyltransferase; *Thermus caldophilus* GK24; Thermostability; β-Sheets

1. Introduction

Glucose-1-phosphate thymidylyltransferase (GPTT), found in several bacterial species, catalyses the reaction that combines dTTP with glucose-1-phosphate (G-1-P) to yield pyrophosphate and dTDP-glucose [1]. Recently, three-dimensional structures of four-mesophilic glucose-1-phosphate thymidylyltransferases have been reported. These include RmlA from *Pseudomonas aeruginosa* [1], RmlA from *Salmonella enterica*LT2 [2], RffH and RmlA from *Escherichia coli* [3,4]. Many antibiotics, such as macrolides, anthracyclines, glycopeptides and coumarin glycosides contain deoxysugar as structural components, and are usually essential for the biological activity of the particular antibiotic. dTDP-glucose is a leading precursor of many deoxysugars, including L-rhamnose, which is associated with the formation of bacterial cell walls. Thermophilic enzymes have considerable potential for many industrial applications because of their high specificity and unusual stability towards

Abbreviations: GPTT, glucose-1-phosphate thymidylyltransferase; G-1-P, glucose-1-phosphate; RSCU, relative synonymous codon usage

* Corresponding author. Tel.: +82 41 530 2246; fax: +82 41 544 2919.
E-mail address: sohng@sunmoon.ac.kr (J.K. Sohng).

0141-0229/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.enzmictec.2005.02.024

many biotransformations. Therefore, an understanding of thermophilic glucose-1-phosphate thymidylyltransferase can provide new insights to deoxysugar biosynthesis.

Thermophilic enzymes are identified in several organisms and many theories have been postulated to differentiate them from their mesophilic counterparts on a molecular basis, but none of them is satisfactory. Thus, it is reasonable to continue further investigation in this area. One approach to a better theoretical understanding involves the comparison of structures and sequences of homologous proteins from thermophilic and mesophilic organisms [5,6]. Previous analyses have suggested that several factors such as hydrophobicity, residues sensitive to oxidation or deamination, and electrostatic interactions may contribute to enhance thermostability [7]. A better way to distinguish thermophiles and mesophiles has been suggested [8], and consists of the following three levels: nucleotide content, codon usage and amino acid composition. It has been found that highly expressed genes often use a more restricted set of preferred synonymous codons than genes with low expression [9], and they were strongly preferential for codons with C or G at the third nucleotide position. It is interesting to note that thermophiles contain more hydrophobic polar amino acids (glutamic acid and tyrosine, etc.) and hydrophobic neutral amino acids (valine and isoleucine, etc.) [10–13].

Many research groups have suggested that the formation of salt bridges, and hydrogen bonding interactions between side chains are other important factors governing thermostability [11,14]. Both events are considerably higher in the majority of the thermophilic proteins. Similarly, thermolabile amino acids like methionine and glutamine are reported in the thermophiles [15]. Secondary structure analyses have revealed that β-strand content might be responsible for the thermal stability of proteins [13]. Thus, one can believe that the structural analyses of thermophilic glucose-1-phosphate thymidylyltransferase can provide significant informations.

We report the biochemical characterization and determinants of thermophilic activity of stRmlA from *Thermus caldophilus* GK24, which is recognized as the first enzyme involved in the dTDP-L-rhamnose biosynthetic gene cluster. In particular, we found that stRmlA is thermostable up to 90 °C only in the presence of dTTP and dTMP. It is the first report of an enzyme, whose thermal stability is dependent on dTTP or dTMP. The comparative study of thermophilic stRmlA with other mesophilic and thermophilic GPTTs, therefore, reveals useful guidelines to pursue further study on this area.

2. Materials and methods

2.1. Bacterial strains and data acquisition

strmlA (highly expressed thermophilic GPTT) was obtained from *T. caldophilus*, whereas *rffH* (highly expressed mesophilic GPTT) was isolated from *Escherichia coli* K12 [16,3]. Six other highly expressed mesophilic GPTTs were also studied on the molecular basis (Table 3). Sixteen putative mesophilic GPTTs were searched in a recent released NCBI database. These include *dnrL* (*Streptomyces peucetius*), *rmlA* (*Vibrio cholerae*), *mtmD* (*Streptomyces agillaceus*), *strD* (*Streptomyces griseus*), *blmD* (*Streptomyces bluensis*), *rffH* (*Photorhabdus luminescens*), *cloV (Streptomyces roseuchromogenes*), *rfbA* (*Streptococcus agalactiae*), *rmlA* (*Mycobacterium bovis*), *rfbA* (*Closteridium perfringens*), *rfbA* (*Neisseria meningitides*), *nanG*1 (*Streptomyces nanchangensis*), *novV* (*Streptomyces caeruleus*), *sim*23 (*Streptomyces antibioticus*), *knaD* (*Streptomyces kanamycetius*) and *prmD* (*Streptomyces rimosus* subsp). Similarly, seven putative thermophilic GPTTs were searched by genome sequence of thermophilic organisms: GPTT (YP_005731.1) from *Thermus thermophilus* HB27, GPTT (NP_228671.1) from *Thermotoga maritime* MSB8, GPTT (Ta1074) from *Thermoplasma acidophilum*, GPTT (NP_376333.1) from *Sulfolobus tokodii* str.7, G1PT (PH1925) from *Pyrococcus horikoshii*, GPTT (NP_126714.1) from *Pyrococcus abyssi*, GPTT (APE1181) from *Aeropyrom pernix*. Thermophilic GPTT *(rmlA)* from *Aneurinibacillus thermoaerophilus* was also considered.

2.2. DNA manipulations and purification of enzymes

strmlA was amplified from *T. caldophilus* (EMBL accession number: AJ605742), and mesophilic GPTT (*rffH*) was derived from *Escherichia coli* K12 (E.C.2.7.7.24). Both products of *strmlA* and *rffH* were cloned into pET32a to generate pSNP1 10 (*NdeI* and *HindIII*) and pSNP230 (*BamHI* and *HindIII*). Finally, *E. coli* BL21 (DE3) was transformed with pSNP110 and pSNP230. *E. coli* BL21 (DE3)/pSNP110 was overexpressed by induction of cells growing at 37 °C to an optical density of 0.6 at a frequency of 600 nm with 0.4 mM IPTG and was again incubated at 20 °C for 21 h. Similarly, overexpression of *E. coli* BL21 (DE3)/pSNP230 was achieved by induction of cells growing at 37 °C to an optical density of 0.6 at a frequency 600 nm with 0.4 mM IPTG and was again incubated at 16 °C for 37 h. The overexpressed cells harboring recombinants *rffH* and *strmlA* were collected by centrifugation separately and washed with a 50 mM potassium phosphate buffer (pH 7.5). The suspended pellets were sonicated for 4.5 min.

The purification of thermophilic stRmlA was performed by heat treatment. Cell free extract of *E. coli* BL21 (DE3)/pSNP110 was incubated at 85 °C for 30 min with 50 mM dTTP or dTMP. After centrifugation, purified prostRmlA was collected from the supernatant.

2.3. Protein determination

The protein concentrations were determined by the method of Bradford (Bio-Rad, Richmond, CA, USA), using brovine serum albumin as the standard.

2.4. Thermal stability of stRmlA

The thermal stability of the stRmlA was determined by measuring relative activity after incubating the enzyme in 50 mM Tris–HCI buffer (pH 7.5) containing 20 mM $MgCl_2$, 50 mM G-1-P and one unit of inorganic pyrophosphatase at a concentration of 15.7 mg/ml at different temperatures for 20 min. The time course of the heat inactivation of the enzyme was examined at 60, 70, 80 and 90 °C in the presence of dTTP, and separately for dTMP. At various time intervals, samples were withdrawn and subjected to catalytic activity assay as described above. The reaction was quenched by 500 μl of 0.5 M EDTA. The supernatant was analyzed in a Mightysil RP-18 column by HPLC with detection at 254 nm. Separation of dTTP and dTDP-glucose was achieved by 100 mM KH_2PO_4 buffer as a mobile phase at flow rate of 1 ml min^{-1} [16]. Purified pro-stRmlA and cell free extract of stRmlA were considered for this study of thermostability.

2.5. Effect of metal ions and chemical denaturants

The enzyme activity was carried out using 20 mM concentration of metal ions and various concentrations of denaturating agents like urea and guanidine chloride for 20 min as described in Section 2.4.

2.6. Effect of pH

The enzyme activity of stRmlA was estimated at various pH (2–13) under standard assay conditions to determine the optimum pH. The pH stability of the enzyme was measured by incubating purified pro-stRmlA for 20 min at 70 °C in Tris–HCI buffer of desired pH.

2.7. Substrate specificity

The substrate specificity of the purified enzyme for mannose-1-phosphate and galactose-1-phosphate was studied using the routine assay described in Section 2.4.

2.8. Comparison of primary structure of thermophilic and mesophilic GPTTs

The CodonW program was used to compare the patterns of relative synonymous codon usage (RSCU) of various GPTTs [17]. The amino acid contents of GPTTs were determined at the internet address "http://www.justbio.com".

2.9. Structural modeling of thermophilic and mesophilic GPTTs

Secondary structures of thermophilic and mesophilic GPTTs were analyzed by the PREDIATOR program [18]. stRmlA was modeled using the modeler "6v4" from the server using the template "1gor" [19]. The template was searched using mGenThreader and the corresponding alignment was carried out using TITO steps; however, it was modified manually. The energy-minimized model of stRmlA was prepared by eliminating some residues from C-terminal and focused only on active site residues. The active site residues were assigned from the model as well as from the multiple alignments with other counterparts. dTTP and G-1-P were extracted from the PDB database, and were docked at the active site using the AUTODOCK3 program [20]. The ligands and macromolecules were prepared for docking experiments as per the usual procedure using the autodock tools. The possible interaction of dTTP and G-1-P was analyzed. To compare the nature of residues at the active site between mesophilic and thermophilic enzymes, all the residues within 6 Å around the substrate were selected. The polar and apolar areas of the modeled structures of stRmlA, MtmD (*S. agillaceus*), RmlA (*A. thermoaerophilus*), and RffH (*E. coli* K12) were calculated using the GETAREA program [21].

3. Results

3.1. Purification and thermal stability of stRmlA

In order to purify the stRmlA, the soluble fraction of the cell free extract of *E. coli* BL21 (DE3)/pSNP110 was incubated either with 50 mM dTTP or dTMP at 85 °C for 30 min. The purpose of the heat treatment is not only to activate the stRmlA, but also to degrade and precipitate the contaminated proteins of *E. coli*. Thereafter, the supernatant containing purified enzyme was collected by centrifugation, and examined as a single band on the SDS-PAGE gel. It is interesting to note that upon heat treatment, 38 kDa of stRmlA is converted to a 25 kDa pro-stRmlA. The purified enzyme has a concentration of 23.88 mg/ml, eight-fold decrease in concentration of cell free extract.

The reaction mixture including either purified pro-stRmlA or its cell free extract shows the similar catalytic activity towards bioconversion of G-1-P to dTDP-glucose. It is well known that the binding of dTTP does not require Mg^{2+} and G-1-P. pro-stRmlA purified in the presence of dTTP and Mg^{2+} shows a decrease in activity compared to the enzyme purified only in the presence of dTTP. This implies Mg^{2+} has no effect on the thermal stability of pro-stRmlA.

Similarly, in order to examine the effects of other compounds on the thermal stability of stRmlA, its cell free extract was incubated at 85 °C for 30 min with uridine 5′-monophosphate (UMP), thymidine, G-1-P, glucose, galactose-1-phosphate, and mannose-1-phosphate. However, none of them provides thermal stability in the presence dTTP or dTMP. Moreover, we also extended a similar study on mesophilic RffH from *E. coli* K12 but found that dTTP and dTMP had no effect on its stability. Thus, we can conclude that thermal stability of stRmlA depends solely on dTTP and

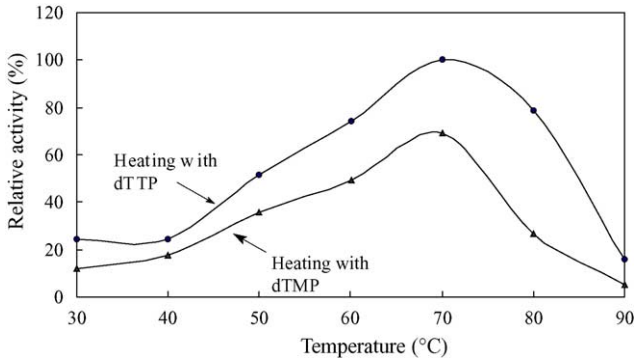

Fig. 1. Thermal stability of stRmlA in the presence of dTTP and dTMP.

dTMP. When the stRmlA was examined at various temperatures, the activity of the enzyme was increased with the change of temperature from 30 to 70 °C and it exhibited the maximum activity at 70 °C (Fig. 1), which is the same as the growth temperature of *T. caldophilus*.

3.2. Effect of metal ions and chemical denaturants

stRmlA requires a single bivalent metal ion to become active. As shown in Fig. 2, the order of effectiveness is $Mg^{2+} > Zn^{2+} > Cu^{2+} > Co^{2+} > Fe^{2+} > Ca^{2+} > Fe^{3+}$ and Ni^{2+} ions. The optimum concentration of Mg^{2+} was 20 mM. stRmlA showed the catalytic activity on a wide range of metal ions in contrast to its mesophilic counterparts. Moreover, stRmlA showed its activity even in the presence of strong denaturants like 50 mM guanidine (pH 7.0) and 50 mM urea.

3.3. Effect of pH

The optimum pH of stRmlA for the formation of dTDP-glucose was 11.5, in contrast with 7.5 for its mesophilic counterpart RffH from *E. coli* K12. stRmlA showed catalytic activity in a range of pH from 2.0 to 13.0. It showed a decrease in catalytic activity at the neutral point, i.e., at pH of 7.0. In contrast, mesophilic RmlA from *E. coli* K12 exhibited activity between pH of 6.0 and 10.0, with a maximum activity around 8.0–8.5 [3].

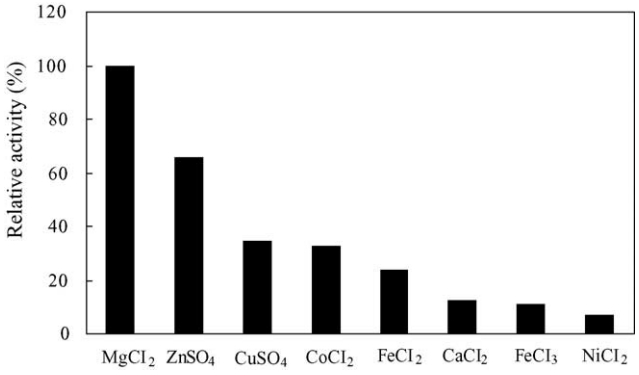

Fig. 2. Effect of metal ions on catalytic activity of stRmlA.

3.4. Substrate specificities

The substrate specificity of the purified enzyme was examined in each forward reaction of dTTP. While G-1-P is an stRmlA substrate, galactose-1-phosphate and mannose-1-phosphate have no activity with stRmlA.

3.5. Codon usage comparisons of thermophilic and mesophilic GPTTs

We analyzed the patterns of synonymous codon usage of stRmlA and various other GPTTs. The significant increment of relative synonymous codon frequencies of 13 codons (CCC, GGC, GUG, CUC, AUC, UAC, CGC, AAC, ACC, UCC, GUC, UUC and GAC) was observed in *strmlA* (Table 1). There was a relatively lower of 10 codons (GCU, GCA, GGA, CCG, UCA, AUU, CAA, CAU, AAA and CGU) in *strmlA* than other counterparts. We found that highly expressed thermophilic (*strmlA*) and mesophilic GPTTs displayed a more biased codon usage than genes with lower expression (putative GPTTs) as shown in Table 1. Some codons are extremely rare in the *strmlA*. Most notably UUU, CCU, GAU, GUU, UUA, CUU, CUA, CCA, UCU, AGU, AUA, ACU, AAC, UAU, GAU and AGA are absent in *strmlA*.

For further analyses of RCSU, *desIII* (GC%; 65) from *S. venezuelae*, *rmlA* (GC%; 65) from *Pseudomonas aeruginosa*, and *strmlA* (GC%; 67) were selected to avoid the effect of GC content variations among GPTTs. We found that only CCC, CUC, and UCC codons are unique in *strmlA*. Among these, we observed a several fold increase in CCC codons encoding proline (with a corresponding decrease in CCG codons) in *strmlA*. There are only CUC codons for leucine (no UUA, CUU and CUA codons) in *strmlA*. UCC codons encoding for serine are found at a higher frequency in *strmlA*. These codons (CCC, CUC and UCC) are also distributed uniformly in *strmlA*. We believe that such codons affect the translation speed of mRNA, the protein folding path, and the resulting structure of the folded protein.

Further analyses revealed that GC rich coding sequences are higher not only in thermophilic *strmlA* but also in mesophilic counterparts *desIII* and *rmlA*; however, a previous study based on genome analyses of thermophiles and mesophiles showed that AT rich coding sequences are significantly higher in thermophiles [22]. Thus we conclude that GC rich coding sequences may not be the reason for thermostablity.

3.6. Amino acid comparisons of thermophilic and mesophilic GPTTs

Analyses of amino acids of stRmlA with various other GPTTs showed that nonpolar amino acids such as alanine,

Table 1
Comparison of relative synonymous codon usage of putative and highly expressed mesophilic and thermophilic GPTTs

Amino acid	Codon	[a]Putative mesophiles (mean)	[b]Putative thermophiles (mean)	[c]Functional mesophiles (mean)	Thermophilic stRmlA
Ala	GCU	0.46	**0.91**	0.66	0.12
	GCC	1.91	1.49	1.63	**2.42**
	GCA	0.51	**0.91**	0.63	0.12
Gly	GGA	0.64	**1.22**	0.62	0.35
	GGC	1.94	0.95	1.77	**2.24**
Val	GUU	0.48	**1.23**	0.75	0
	GUC	1.65	0.65	1.33	**2.34**
	GUA	0.44	**0.75**	0.53	0
	GUG	1.43	1.37	1.39	**1.66**
Leu	UUA	0.59	**0.74**	0.57	0
	CUU	0.38	**0.92**	0.68	0
	CUC	1.56	2.05	1.1	**3.4**
	CUA	0.24	0.77	0.26	0
Phe	UUC	1.39	1.13	1.13	2
	UUU	0.61	**0.87**	**0.87**	0
Pro	CCU	0.4	**0.88**	0.47	0
	CCC	1.39	1.46	0.5	**3.5**
	CCA	0.58	**0.92**	0.79	0
	CCG	1.63	0.73	**2.23**	0.5
Ser	UCU	0.64	**1.09**	0.82	0
	UCC	1.58	0.9	0.86	**2.73**
	UCA	0.66	**0.99**	0.43	0.27
	AGU	0.42	0.94	1.2	0
	AGC	1.32	1.42	1.2	**1.64**
Lle	AUU	0.68	**0.7**	1.14	0.1
	AUC	1.94	0.88	1.65	2.9
	AUA	0.37	**1.42**	0.21	0
Thr	ACU	0.48	**1.20**	0.42	0
	ACC	1.99	0.75	1.92	**2.89**
	ACG	0.96	**1**	0.88	0.67
Asn	AAU	0.64	0.85	**0.93**	0
	AAC	1.36	1.15	1.07	2
Gln	CAG	1.28	1.5	1.44	**1.71**
Tyr	UAU	0.83	0.91	**0.99**	0
	UAC	1.17	1.09	1.01	**2**
His	CAU	0.54	0.68	**0.88**	0.33
	CAC	1.46	1.32	1.13	**1.67**
Asp	GAU	0.54	**1.36**	0.98	0
	GAC	1.46	0.64	1.02	**2**
Lys	AAA	0.86	0.86	**1**	0.25
	AAG	1.14	1.14	1	**1.75**
Arg	CGU	0.81	0.26	**1.58**	0.5
	CGC	2.04	0.39	2.63	**3**
	AGA	0.28	**1.83**	0.21	0

The values shown are the relative frequencies of synonymous codons within each codon group. Only those codons that show significant differences are listed here.

[a] Putative 18 mesophilic GPTTs.
[b] Putative 7 thermophilic GPTTs.
[c] Functional mesophilic (highly expressed) GPTTs.

isoleucine and valine are higher in stRmlA (Table 2). Such amino acids are supposed to enhance the protein hydrophobicity, which is another notable feature of thermophilic protein. Polar and charged amino acids like aspartate and histidine are observed more frequently in stRmlA, and are crucial to increase salt bridges and H-bonds in stRmlA. Mesophilic GPTTs have substantially more thermolabile residues, methionine and glutamine, than stRmlA.

Table 2
Number of amino acids in mesophiles and thermophiles GPTTs

Amino acids (abbreviation)	Mesophiles[a] (mean)	Thermophilic stRmlA	Difference	*t*-value
Alanine (A)	78	**92.4**	14.4	−9.202
Cysteine (C)	**6.5**	5.6	−0.9	−2.211
Aspartate (D)	56.5	**86.8**	30.3	−8.999
Glutamate (E)	**68.8**	58.8	−10	−7.206
Phenylalanine (F)	32.1	25.2	−6.9	−4.672
Glycine (G)	97.9	95.2	−2.7	−9.265
Histidine (H)	16.1	**33.6**	17.5	−5.642
Isoleucine (I)	65.4	**81.2**	15.8	−8.636
Lysine (K)	**42.2**	22.4	−19.8	−4.254
Leucine (L)	**108.9**	84	−24.9	−8.592
Methionine (M)	**21.1**	11.2	−9.9	−2.998
Asparagine (N)	**29.6**	14	−15.6	−3.309
Proline (P)	**47.6**	44.8	−2.8	−6.315
Glutamine (Q)	37	19.6	−17.4	−3.975
Arginine (R)	58.5	**67.2**	8.7	−7.825
Serine (S)	57.1	**61.6**	4.5	−7.468
Threonine (T)	44.1	**50.4**	6.3	−6.767
Valine (V)	78.2	**114.8**	36.6	−10.345
Tryptophan (W)	8.5	8.4	−0.1	−2.733
Tyrosine (Y)	**42.4**	16.8	−25.6	−3.543

The values shown are the numbers of amino acids within each gene. The numbers were scaled to a total of 1000 for each gene, to offset the effects of variations in gene size. Mean values for the thermophilic and mesophilic genes are shown. The differences between the mean values are indicated with the *t*-test.

[a] Sixteen mesophilic genes considered for this comparison.

3.7. Structural comparisons of thermophilic and mesophilic GPTTs

stRmlA has higher amounts of β-strands than other highly expressed mesophilic counterparts. This contributes to thermal stability by increasing the number of intermolecular hydrogen bond (Table 3). The modeling of stRmlA and docking of dTTP at the active site showed that the general conserved residues for its activity were situated at the preferred interactive distance from the substrate (Fig. 3). The reason why stRmlA showed its activity only after heating with dTTP or dTMP, but not during heating in the absence of dTTP or dTMP, is a mystery. Although the exact reason remains unclear, it is apparent from our experiments that thymidyl and phosphate groups of the dTTP and dTMP could play a role in enzyme stability because, as discussed earlier, only thymidine is unable to provide thermal stability in stRmlA.

The residues at the active site, similar to other mesophilic counterparts, are conserved in stRmlA. As shown in Fig. 3, hydrophobic pocket formed by the residues Phe141, Leu168, Val169 could provide the interaction of methyl group of thymidine. Similarly, Arg13, a conserved residue from the motif (G)GXGXR(L), forms a strong H-bond with α-phosphate group of dTTP followed by the formation of a similar H-bond by Asp 107 with the terminal phosphate group (β-phosphate group). The interaction is further strengthened by the formation of H-bond by the backbones of Gly9 and Gly11 with β-phosphate group. Similarly, the G-1-P, a substrate of stRmlA was docked at the active site and the interactions were analyzed. The results showed that the conserved residues for sugar binding were found at the appreciable position.

Table 3
Comparison of secondary structure mesophilic and thermophilic GPTTs

GPTT	Strain	Sheet	Helix	Coil
Functional				
stRmlA	*Thermus caldophilus*	32.5	17.6	49.7
RmlA	[a]*A. thermoaerophilus*	21.8	31.6	46.5
RmlA[b]	*Salmonella enterica*	23.9	30.4	45.5
RffH[b]	*Escherichia coli K12*	24.5	31.3	44
RfbA[b]	*Escherichia coli K12*	23.2	30.9	45.9
RmlA[b]	*Mycobacterium tuberculosis*	24.3	31.9	43.7
RfbA[b]	*Porphyromonas gingivalis*	21	30	49
RmlA[b]	*Pseudomonas aeruginosa*	20	31.3	48.6
Putative				
RffH	*Photorhabdus luminescens*	23.5	31.3	45
CloV	*Streptomyces roseochromogenes*	21.6	33.1	45.2
RfbA	*Streptococcus agalactiae*	22.4	30.7	46.7
RfbA	*Clostridium perfringens*	23.8	31.3	45.4
RfbA	*Neisseria meningitidis*	25	31.9	43
anG1	*Streptomyces nanchangensis*	22.5	30.4	47

[a] *Aneurinibacillus thermoaerophilus* DSM 10155.

[b] Functional mesophilic (highly expressed) GPTTs. Rest is putative mesophilic GPTTs.

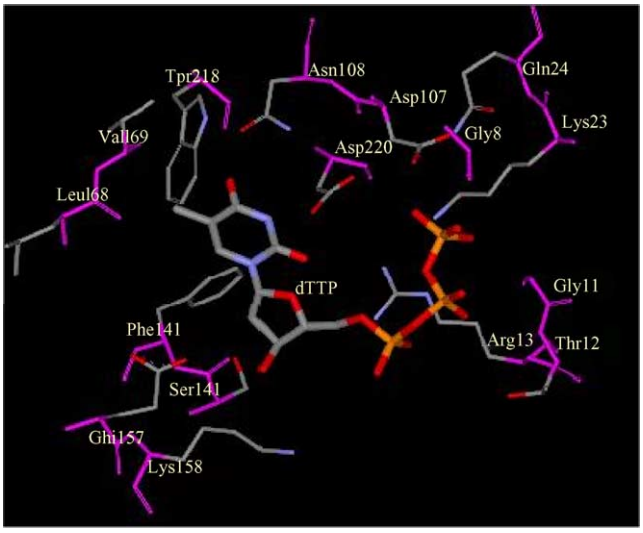

Fig. 3. Modeling of stRmlA to investigate the interaction of G-1-P at active sites.

4. Discussion

The cell free extract of stRmlA is thermally unstable however; the reaction mixture including cell free extract of stRmlA shows the catalytic activity on a wide range of temperatures (up to 90 °C). Being a thermophilic enzyme, it

Table 4
Number of polar and apolar functional groups present in exposed surface areas

GPTT	Strain	Polar (Å^2)	Apolar (Å^2)	Remarks
stRmlA	*Thermus caldophilus*	16.2	35.5	Thermophile shows activity up to 90 °C
RmlA	[a]*A. thermoaerophilus*	22.3	32.0	Mesophile shows activity up to 55 °C
MtmD	*Streptomyces agillaceus*	20	32.9	Mesophile shows highest identity (73%) with stRmlA
RffH	*Escherichia coli* K12	15.5	22.2	Mesophile shows activity at 37 °C

[a] *Aneurinibacillus thermoaerophilus* DSM 10155.

should be stable even after heat treatment, but it is not true for this enzyme. It retains thermophilic activity only after heat treatment in the presence of dTTP or dTMP. On comparing these two, the catalytic activity of stRmlA purified by incubating with 50 mM dTMP is lower than incubated with dTTP.

Interestingly, stRmlA (38 kDa) converted to pro-stRmlA (25 kDa) after heat treatment. We postulated that stRmlA will break somewhere between Leu^{235} and Val^{258} after heat treatment. It is supported by multiple alignment of stRmlA (356 amino acids) with other functional mesophilic and thermophilic GPTTs studied. Catalytically important residues (**R**13, **K**23, **D**107, **K**158 and **D**220) are present within 220 amino acid of stRmlA. RmlA from *Aneurinibacillus thermoaerophilus* DSM 10155 demonstrated catalytic activity up to 55 °C [23], but it has only 257 amino acids. It clearly supports the degradation of stRmlA. Other thermophilic proteins that show size reduction after heat treatment strengthened our prediction [24]. Unlike other GPTTs, it showed catalytic activity on a wide range of pH (2–13) in the presence of various metal ions. There is no effect of 50 mM urea and guanidine chloride on the catalytic activity of stRmlA.

Codon usage comparisons, amino acid composition analysis, and structural modeling of various GPTTs probe the molecular basis for their unusual thermostability. Some codons (CCC, CUC and UCC) are of significantly higher frequency and uniformly distributed in *strmlA* They might control the protein folding path and structure of the folded protein. We have no further evidences to support molecular basis for a connection between the CCC, CUC and UCC codons and thermostability. A further investigation on this area is needed to explore this statistical correlation and thermostability. We also found some other significance codons in stRmlA, which will be useful for designing the degenerate primers to amplify GPTT.

As predicted by the CodonW program, it has a lower aromaticity index than its functional mesophilic counterparts (data not shown). This indicates that stRmlA has fewer numbers of aromatic amino acid residues (Cys, Met, Trp, Tyr, Phe). These results are also in agreement with other findings obtained by amino acid analyses (Table 2). We also observed a lower number of deamidated amino acids like asparagine and glutamine in stRmlA. These two amino acids are easily deamidated at elevated temperatures, so their reduction aids thermostability [25]. It has been suggested that the methionine and aromatic amino acid residues are oxidized rapidly so their reduction favors thermostability [26].

Understanding the structural basis for the enhanced stability of proteins from thermophilic organisms relative to their mesophilic counterparts is a challenging problem; however, the increasing availability of high-resolution crystal structure from both thermopiles and mesophiles provides an important insight for determining the structural basis of thermal stability.

Obviously, there might be several reasons for thermal stability. Results suggest that stRmlA has higher numbers of β-strand than other highly expressed mesophilic counterparts, which contributes to thermal stability by increasing the amount of intermolecular hydrogen bonding.

The nature of the residues around the substrates (dTTP and G-1-P) and the exposed area of the whole protein were investigated. It is obvious that thermal stability is due to the residues at the active site, but may also be enhanced by exposed residues. In Table 4, we clearly see that the ratio of apolar to polar residues in stRmlA is higher than in mesophilic counterparts, which is consistent with theoretical predictions. This is the reason why this enzyme has no effect on a wide range of pH. It is important to recall that there will be unfavorable electrostatic interactions with charged or polar residues exposed to the surface area of the protein for adverse variations in pH. Finally, we reached the conclusion that thymidyl as well as phosphate group of dTTP or dTMP could play a role in enzyme stability. We proposed that dTTP or dTMP enters the active site and binds tightly to inhibit the defolding of the protein, which in turn, increases its thermal stability.

Acknowledgements

The work was supported by the 21C Frontier Microbial Genomics and Application Center Program, Ministry of Science & Technology (Grant MG02-0301-004-2-3-1), Republic of Korea.

References

[1] Blankenfeldt W, Asuncion M, Lam JS, Naismith JH. The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidylyltransferase (RmlA). EMBO J 2000;19:6652–63.
[2] Barton WA, Lesniak J, Biggins JB, Jeffrey PD, Jiang J, Rajashankar KR, Thorson JS, Nikolov DB. Structure, mechanism and engineering of a nucleotidylyltransferase as a first step toward glycorandomization. Nat Struct Biol 2001;8:545–51.

[3] Sivaraman J, Sauve V, Matte A, Cygler M. Crystal structure of *Escherichia coli* glucose-1-phosphate thymidylyltransferase (RffH) complexed with dTTP and Mg^{2+}. J Biol Chem 2002;277:44214–9.
[4] Zuccotti S, Zanardic D, Rosano C, Sturla L, Tonetti M, Bolognesi M. Kinetic and crystallographic analyses support a sequential-ordered bi bi catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidylyltransferase. J Mol Biol 2001;313:831–43.
[5] Perutz MF, Raidt H. Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. Nature 1975;255:256–9.
[6] Argos P, Rossmann MG, Grau UM, Zuber H, Frank G, Tratschin JD. Thermal stability and protein structure. Biochemistry 1979;18:5698–703.
[7] Creighton TE. Prote: structures and molecular properties. 2nd ed. New York: Freeman & Company; 1997.
[8] Gregory AC, Singer, Donal AH. Thermophilic prokaryotes have characteristic patters of codon usage, amino acid composition and nucleotide content. Gene 2003;23:39–47.
[9] Liu Q, Feng Y, Zhao X, Dong H, Xue Q. Synonymous codon usage bias in *Qryza sativa*. Plant Sci 2004;167:101–5.
[10] Sudeep PG, Absar A, Mala BR. A novel thermostable xylanase from *Thermomonospora* sp.: influence of additives on thermostability. Bioresour Technol 2001;78:221–4.
[11] Kumar S, Tsai CJ, Nussinov R. Factors enhancing protein thermostability. Prot Eng 2000;13:179–91.
[12] Haki GD, Rakshit SK. Developments in industrially important thermostable enzymes: a review. Bioresour Technol 2003;89:17–34.
[13] Chakravarty S, Varadarajan R. Elucidation of determinants of protein stability through genome sequence analysis. FEBS Lett 2000;470:65–9.
[14] Kreil DP, Ouzounis CA. Identification of thermophilic species by the amino acid compositions deduced from their genomes. Nucleic Acids Res 2001;29:1608–15.
[15] Xiao L, Honig B. Electrostatic contributions to the stability of hyperthermophilic proteins. J Mol Biol 1999;289:1435–44.
[16] Parajuli N, Lee DS, Lee HC, Liou K, Sohng JK. Cloning, expression and characterization of glucose-1-phosphate thymidylyltransferase (*strmlA*) from *Thermus caldophilus*. Biotechnol Lett 2004;26:437–42.
[17] Peden JF, "http://www.molbiol.ox.ac.uk/cu/" distributes the program Codon W, 1997.
[18] Frishman D, Argos P. Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. Protein Eng 1996;9:133–42.
[19] Sali, Potterton L, Yuan F, Vlijmen VH, Karplus M. Evaluation of comparative protein modeling by MODELLER. Proteins 1995;23:318–26.
[20] Morris GM, Goodsell DS, Huey R, Olson AJ. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. J Comput Aided Mol Des 1996;10:293–304.
[21] Fraczkiewicz R, Braun W. GETAREA program. J Comp Chem 1998;19:319–33.
[22] Lao PJ, Forsdyke DR. Thermophilic bacteria strictly obey Szybalski's transcription direction rule and politely purine-load RNAs with both adenine and guanine. Genome Res 2000;10:228–36.
[23] Graninger M, Kneidinger B, Bruno K, Scheberl A, Messner P. Homologs of the Rml enzymes from *Salmonella enterica* are responsible for dTDP-beta-L-rhamnose biosynthesis in the Gram-positive thermophile *Aneurinibacillus thermoaerophilus* DSM 10155. Appl Environ Microbiol 2002;68:3708–15.
[24] Wu J, Bian Y, Tang B, Chen X, Shen P, Peng Z. Cloning and analysis of WF146 protease, a novel thermophilic subtilisin-like protease with four inserted surface loops. FEMS Microbiol Lett 2004;230:251–8.
[25] Hess JM, Kelly RM. Influence of polymolecular events on inactivation behavior of xylose isomerase from *Thermotoga neapolitana* 5068. Biotechnol Bioeng 1999;62:509–17.
[26] Berlett BS, Stadtman ER. Protein oxidation in aging, disease, and oxidative stress. J Biol Chem 1997;272:20313–6.