An Alternative Flavin-Dependent Mechanism for Thymidylate Synthesis

Hannu Myllykallio,[1]* Gerard Lipowski,[2] Damien Leduc,[1] Jonathan Filee,[1] Patrick Forterre,[1] Ursula Liebl[2]*

Although deoxythymidylate cannot be provided directly by ribonucleotide reductase, the gene encoding thymidylate synthase ThyA is absent from the genomes of a large number of nonsymbiotic microbes. We show that ThyX (Thy1) proteins of previously unknown function form a large and distinct class of thymidylate synthases. *ThyX* has a wide but sporadic phylogenetic distribution, almost exclusively limited to microbial genomes lacking *thyA*. ThyX and ThyA use different reductive mechanisms, because ThyX activity is dependent on reduced flavin nucleotides. Our findings reveal complexity in the evolution of thymidine in present-day DNA. Because ThyX proteins are found in many pathogenic microbes, they present a previously uncharacterized target for antimicrobial compounds.

All deoxythymidine 5′-monophosphate (dTMP) in bacteria and eukarya is thought to be formed either de novo by thymidylate synthase (ThyA)–dependent methylation of deoxyuridine 5′-monophosphate (dUMP) or by thymidine kinase (Tdk)–dependent salvage of thymidine compounds from the growth medium (*1*). ThyA uses tetrahydrofolate (H_4folate) as a reductant and forms dihydrofolate (H_2folate) as a product of the methylation reaction (*2*). Because reduced folate derivatives are essential for a variety of biological processes, H_2folate formed by ThyA is rapidly reduced to H_4folate by dihydrofolate reductase (DHFR). This functional, and often physical, coupling of ThyA and DHFR proteins was thought to be essential for de novo thymidylate synthesis in virtually all actively dividing cells.

The hyperthermophilic anaerobic archaeon *Pyrococcus abyssi,* which lacks thymidine kinase, incorporates label from extracellular uracil, but not from thymidine, into its DNA (*3*). This implies that *P. abyssi* must synthesize thymidylate de novo. However, iterative similarity searches (*4*) of the three completed *Pyrococcus* genomes did not reveal any candidate genes for *thyA* or for their distantly related putative archaeal homologs (*5*). When analyzing additional genomes, we detected many other examples of archaea and bacteria that lacked known pathways for formation of dTMP (see Table 1 for a nonexhaustive listing). Many organisms missing *thyA* also lacked recognizable dihydrofolate reductase genes (Table 1). Together, these in silico observations suggested that, in many microbes, alternative enzymes might function in thymidylate metabolism.

A gene of unknown function (*6*) had previously been shown to complement a thymidine-requiring mutant of *Dictyostelium discoideum* to thymidine prototrophy. This gene was named *thy1*, but, to avoid confusion with the Thy-1 cell surface antigen, we will refer to this gene and its homologs as *thyX*. ThyA and ThyX (Thy1) lack any sequence similarity, and there was no functional evidence that ThyX is a thymidylate synthase. Using *thyX* from *D. discoideum* (GenBank accession number g135829) as bait in manual similarity searches, we have found many additional *thyX* genes with a wide phylogenetic distribution in up to 30% of microbial genomes (Fig. 1) (*7*). Although the average pairwise sequence identity of the ThyX homologs is only 28.4% (*8*), sequence alignments revealed a specific sequence motif $RHRX_7S$ ("ThyX motif") common to this family of proteins (fig. S1) (other positively charged amino acid residues can substitute for the arginine residues) (*9*). With the exception of *Corynebacterinea* genomes, *thyX* and *thyA* genes have mutually exclusive phylogenetic patterns (*10*), on the basis of which we predict that ThyX has substituted for the unrelated ThyA protein.

We tested directly whether *P. abyssi* or *Helicobacter pylori thyX* (GenBank accession numbers g14916853 and g7464006, respectively) can functionally complement growth defects of an *Escherichia coli* strain specifically impaired in thymidylate synthase activity. Using an arabinose inducible promoter, these proteins were produced in *E. coli* strain χ2913 carrying a well-defined deletion

Table 1. Nonexhaustive list of genomes apparently lacking *thyA* and *tdk* genes. Clinically relevant bacteria are indicated in bold type. N.A., not analyzed.

Species	*ThyA**	*DHFR*	*Tdk*	*thyX*	Interest and comments
Bacteria					
Helicobacter pylori	–	–	–	+	Implicated in ulcer formation
Campylobacter jejuni	–	–	–	+	Leading cause of food poisoning
Rickettsia prowazekii	–	–	–	+	Causative agent of typhus
Borrelia burgdorferi	–	–	–	+	Implicated in Lyme disease
Treponema pallidum	–	–	–	+	Causative agent of syphilis
Chlamydia (three species)	–	+	–	+	Obligate intracellular pathogens
***Mycobacterium tuberculosis*‡**	+	+	–	+	Tuberculosis
Archaea					
Pyrococcus abyssi	–	–	–	+	Hyperthermophile
Pyrococcus horikoshii	–	–	–	+	Hyperthermophile
Sulfolobus solfaraticus	–	–	–	+	Hyperthermophile
Eukarya					
Dictyostelium discoideum	N.A.	N.A.	+	+	Not completely sequenced
Viruses					
Bacterial and eukaryotic DNA viruses (five species)	–	N.A.	N.A.	+	

*Iterative PSI-BLAST tool (*4*) was used for extensive similarity searches using various ThyA or ThyX sequences as baits. Hits with an expected value $<1 \times 10^{-5}$ were considered significant. The cutoff value for recruitment of alignments into successive iterations was 0.02. †*thyA*, gene encoding thymidylate synthase required for de novo synthesis of dTMP; *thyX*, a gene family implicated in de novo synthesis of thymidylate; *tdk*, thymidine kinase required for salvage of exogenous thymidine. ‡*Corynebacterinea* sp. genomes contain both *thyX* and *thyA* but were included in the list (see text).

[1]Institut de Génétique et Microbiologie, CNRS UMR 8621, Université Paris-Sud, Orsay, France. [2]Institut National de la Santé et de la Recherche Médicale (INSERM) U451–Laboratory of Optics and Biosciences, Ecole Polytechnique–ENSTA, Palaiseau, France.

*To whom correspondence should be addressed. E-mail: hannu.myllykallio@igmors.u-psud.fr (H.M.) and ursula.liebl@polytechnique.fr (U.L.)

in the *thyA* gene (Fig. 2A). In the absence of thymidine, *E. coli* χ2913 transformants, carrying the *H. pylori thyX* gene on plasmid, formed yellowish colonies on solid minimal media only in the presence of arabinose (Fig. 2B). The level of growth for individual colonies in our complementation tests was approximately 50% of that observed in the presence of thymidine. The growth was inhibited by trimethoprim, a specific inhibitor of dihydrofolate reductase, but only in the presence of a relatively high dose of this drug [100 μg/ml (table S1)]. Changing Ser^{107} of *H. pylori* ThyX [the last residue in the conserved ThyX motif (fig. S1)] into Ala and ochre (stop) codons abolished complementation (table S1). These genetic data show directly that *H. pylori thyX* can functionally replace *E. coli thyA* in dTMP synthesis. In similar experiments, *P. abyssi thyX* did not complement *E. coli* χ2913 to thymidine prototrophy, presumably reflecting either the incapability of this "hyperthermophilic" protein to function under "mesophilic" conditions or the presence of chemically modified folates in *Pyrococcus* sp. (*11*).

We purified *H. pylori* ThyX carrying a histidine-tag at its carboxy-terminus from cell-free extracts of *E. coli* strain χ2913. The obtained protein preparations (purity >95%) typically contained 1 to 2 mg/ml of protein, with an apparent molecular mass of ≈31 kD on SDS–polyacrylamide gel electrophoresis (SDS-PAGE) gels (the expected molecular mass of *H. pylori* ThyX is 31.5 kD) (Fig. 3A), and were bright yellow in color. Superdex 200 (Phamacia, Sweden) size-exclusion chromatography revealed a native molecular mass of 111 kD for this protein, which suggests that its active form could correspond to a homotetramer (*12*). Spectroscopic analyses of the isolated (oxidized) protein revealed absorbance characteristics typical of a flavoprotein (Fig. 3B), with broad peaks at 447.5 and 375 nm, which were absent in the dithionite-reduced enzyme. Similar absorption characteristics were found for the cofactor after its release from the protein by heat denaturation at 80°C for 5 min.

The loss of tritium from [5-^{3}H]dUMP during the formation of dTMP allows the quantification of thymidylate synthase activity after removal of the radioactive nucleotides from the reaction mixtures (*13*). We found that ThyX catalyzes in vitro the release of tritium from [5-^{3}H]dUMP in a protein concentration– and CH_2H_4folate-dependent manner (*12*), revealing the first biochemical activity for ThyX proteins. However, the activity detected under conditions described for ThyA proteins was too low to explain the results of our complementation tests. In further experiments, we found that including oxidized flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD) and reduced pyridine nucleotides in the assays substantially increased the tritium release activity of *H. pylori* ThyX (Table 2). The simplest explanation for this observation is that an electron flow from reduced pyridine nucleotides via flavin nucleotides is necessary for catalytic activity of ThyX proteins. To exclude the possibility that *H. pylori* ThyX is a uridine 5′ monophosphate (UMP) using methyl transferase, we have shown that UMP

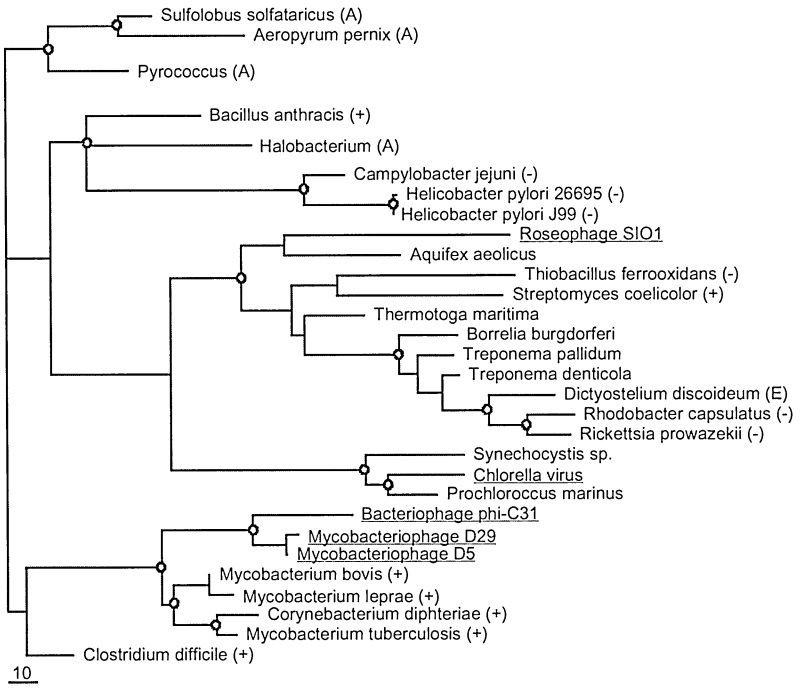

Fig. 1. An unrooted phylogenetic tree based on ThyX homologs retrieved by iterative similarity searches was obtained by maximum likelihood (ML) analyses with the program PROTML [quick-add search with 500 replicates, JJT-F (*16*) amino acid substitution model (*17*)]. The tree shown is the best ML tree obtained. Bootstrap values were computed using the resampling of estimated likelihood (RELL) method (values higher than 90% are indicated by closed circles). Similar results were also obtained using the neighbor joining algorithm. The scale bar represents the number of substitutions per 100 sites. Additional, more distant, candidate *thyX* genes were detected in *Chlamydia* and *Thermoplasma* species but were excluded from the analysis shown here. A, archaea; E, eukarya; + and –, Gram-positive and Gram-negative bacteria, respectively. Viral sequences are underlined.

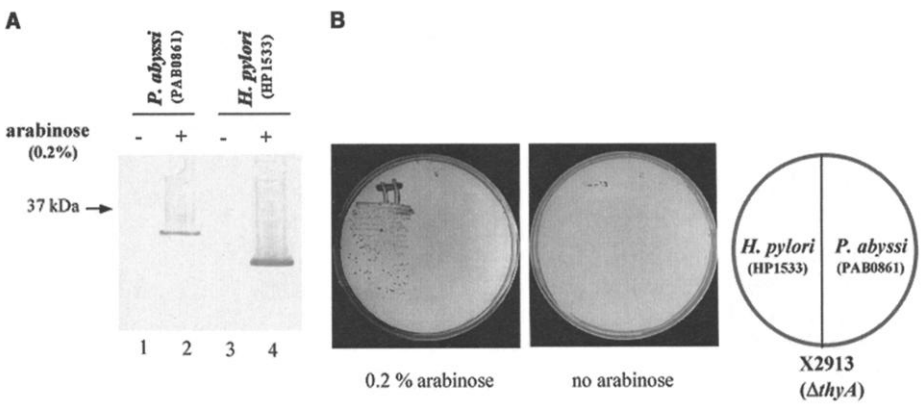

Fig. 2. (**A**) Expression analyses of *P. abyssi* (strain Orsay) and *H. pylori* (strain 26695) ThyX proteins. Western immunoblot analyses using whole-cell lysates of noninduced (–arabinose) and induced (+0.2% L-arabinose) cell cultures of *E. coli* χ2913 (Δ*thyA*) with monoclonal antibodies to Gly-Lys-Pro-Ile-Pro-Asn-Pro-Leu-Leu-Gly-Leu-Asp-Ser-Thr (V5) epitope (V5) are shown. (**B**) The ability of *P. abyssi* and *H. pylori thyX* to permit thymidine-independent growth of *E. coli* thymidine auxotroph χ2913 (Δ*thyA*) was scored after 3 to 4 days in the presence or absence of 0.2% L-arabinose on minimal M9 agar lacking thymidine (*18*). Complementation of χ2913 (Δ*thyA*) to thymidine prototrophy was brought about by arabinose induction of *H. pylori* ThyX. X2913, *E. coli* χ2913.

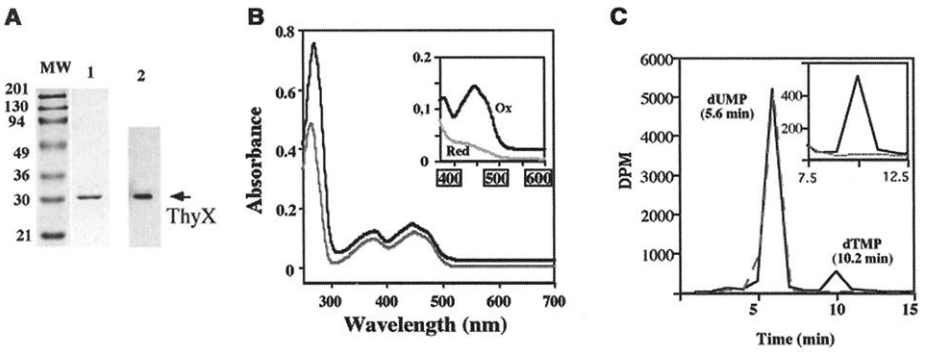

Fig. 3. Biochemical analyses of *H. pylori* ThyX. (**A**) 12% SDS-PAGE and immunoblot analyses of isolated *H. pylori* ThyX protein. 1.5 (lane 1) and 0.3 (lane 2) μg of pure protein was detected using Coomassie Brilliant Blue staining or monoclonal antibodies to V5 epitope (Invitrogen) with enhanced chemiluminescence detection, respectively. The expected molecular weight of *H. pylori* ThyX is 31.5 kD (**B**) Spectroscopic analyses *of H. pylori* ThyX indicate that it is a flavoprotein. Absolute spectra of 10 μM isolated enzyme (black line) and the cofactor after its release from protein (gray line). The inset shows spectra for the oxidized enzyme (Ox) and the dithionite reduced protein (Red). (**C**) dTMP-forming activity of *H. pylori* ThyX protein. Positions of radioactive peaks were determined by collecting 1-mL fractions, followed by the determination of radioactivity using scintillation counting (*3*). The complete reaction is indicated by a smooth line; the control reaction omitting ThyX protein is indicated by a dashed line.

Table 2. Optimized assay conditions for *H. pylori* ThyX protein.

Assay condition	Nmol of ^{3}H released/mg of protein (60 min incubation)
H. pylori ThyX	
Complete*	63.0 (100%)
–protein	0.7 (1.1%)
–H_4folate	0.8 (1.2%)
–NADH,–NADPH,–FMN	2.2 (3.4%)
–FMN	2.70 (4.3%)
–NADPH	37.3 (59.2%)
–NADH	19.0 (30.1%)
+500 μM dUMP	9.3 (14.8%)*
+500 μM UMP	64.4 (102.2%)

*Complete assay mixtures contained 50 mM Tris-HCl at pH 7.0, 1 mM CH_2H_4folate preparation, 10 mM $MgCl_2$, 2 mM NADPH, 1 mM NADH, 0.5 mM flavin mononucleotide (FMN), and 9 μM [5-^{3}H]dUMP (specific activity, 1.7357 Ci/mmol). CH_2H_4folate preparation was obtained by incubating 2 mM H_4folate, 96 mM 2-mercaptoethanol, 42 mM formaldehyde, and 50 mM Tris-HCl in dark for 30 min. Flavin adenine dinucleotide can efficiently replace FMN in the reaction (*11*). Under these reaction conditions, 20 μM dUMP and 100 μM CH_2H_4folate were sufficient to saturate tritium-release activity of *H. pylori* ThyX in a 60-min incubation. A maximal specific activity measured using time-course experiments with ThyX corresponds to 0.015 μmol of ^3H_2O formed per min per mg of protein.

cannot compete with dUMP in the reaction (Table 2). In addition, tritium-release activity of *H. pylori* ThyX is inhibited by micromolar concentrations of dTMP (*12*). Moreover, with the use of [6-^{3}H]dUMP and high-pressure liquid chromatography (Fig. 3C), we have analytically shown that *H. pylori* ThyX possesses a dTMP-forming activity (≈75 nmol/mg protein in a 60-min incubation). These results have established that the physiologically relevant activity of *H. pylori* ThyX is that of a dUMP using thymidylate synthase.

Analogous to CH_2H_4folate- and $FADH_2$-dependent ribothymidyl synthase of *Streptococcus faecalis* (*14*), our data suggest that CH_2H_4folate functions in the ThyX catalyzed reaction only as a one-carbon donor, whereas electrons required for the formation of the methyl moiety originate from reduced-flavin nucleotides. Similar to *S. faecalis* ribothymidyl synthase, ThyX catalysis should then result in the formation of H_4folate as a reaction product. Because of the intrinsic instability of CH_2H_4folate and reduced flavin nucleotides in the presence of molecular oxygen, we have not yet directly tested this proposal. However, the proposed reaction mechanism for ThyX is supported by the absence of dihydrofolate reductase genes in many ThyX-containing organisms (Table 1), whereas all bacteria and eukarya with *thyA* contain DHFR (*12*). The existence of different reaction mechanisms for the two classes of thymidylate synthases is also supported by the observation that ThyX proteins lack sequence motifs that are thought to be essential for catalysis and substrate binding in ThyA proteins (*2*).

ThyA and/or *thyX* genes are present in all completed genome sequences, which suggests that these two proteins are the only thymidylate synthases. The lack of any sequence similarity between ThyA and ThyX proteins, together with the differences in their enzymatic mechanisms, suggests an independent origin for the two distinct thymidylate synthases. This suggestion was confirmed by the very recently determined structure of a eubacterial ThyX homotetramer (see note added in proof). The peculiar sporadic distribution of *thyX* in the three domains of life (Fig. 1) likely results from several independent lateral gene transfer or non-orthologous gene replacement events (e.g., between α-proteobacteria and eukarya or from bacteriophages to mycobacteria). These multiple gene–transfer events prevent a definitive conclusion regarding the evolutionary origin of the two pathways for thymidylate synthesis, but indicate the hitherto unnoticed complexity in the evolution of thymidine-containing genetic material (*15*). Notably, *thyX* is present in a number of human pathogenic bacteria, but absent in the human genome, making this class of thymidylate synthases an attractive target for specifically inhibiting microbial growth.

Note added in proof: The structure of the *Thermotoga maritima* ThyX homotetramer (1KQ4) in complex with a flavin adenine dinucleotide has been solved very recently by the Joint Center for Structural Genomics (www.jcsg.org). It reveals that ThyA and ThyX proteins are not structurally related, thus confirming the independent origin for the two distinct classes of thymidylate synthases.

References And Notes

1. J. Neuhard, R. A. Kelln, in *Escherichia coli and Salmonella*, F. C. Neidhardt, Ed. (American Society for Microbiology, Washington, DC, 1996) pp. 580–599.
2. C. W. Carreras, D. V. Santi, *Annu. Rev. Biochem.* **64**, 721 (1995).
3. H. Myllykallio *et al.*, *Science* **288**, 2212 (2000).
4. A. Altschul *et al.*, *Nucleic Acids Res.* **25**, 3389 (1997).
5. R. Aurora, G. D. Rose, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2818 (1998).
6. J. L. Dynes, R. A. Firtel, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7966 (1989).
7. Our similarity searches of 140 genome sequences indicated that *thyX* and *thyA* can be found in 27 and 73% of microbial genomes, respectively.
8. Pairwise sequence identity of 39 ThyX homologs was determined with the program Psi-Blast (*4*), using *H. pylori* ThyX (HP1533) as bait.
9. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; H, His; R, Arg; and S, Ser.
10. M. Y. Galperin, E. V. Koonin, *Nature Biotechnol.* **18**, 609 (2000).
11. R. H. White, *J. Bacteriol.* **173**, 1987 (1991).
12. H. Myllykallio, D. Leduc, unpublished data.
13. D. Roberts, *Biochemistry* **5**, 3546 (1966).
14. A. S. Delk, D. P. Nagle, J. C. Rabinowitz, *J. Biol. Chem.* **255**, 4837 (1980).
15. A. Poole, D. Penny, B.-M. Sjöberg, *Nat. Rev. Mol. Cell Biol.* **2**, 147 (2001).
16. D. T. Jones, W. R. Taylor, J. M. Thornton, *Comput. Appl. Biosci.* **8**, 275 (1992).
17. J. Adachi, M. Hasegawa, *Mol. Phylogenet. Evol.* **6**, 72 (1996).
18. M. L. Michaels, C. W. Kim, D. A. Matthews, J. H. Miller, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3957 (1990).

19. We thank D. W. Santi for *E. coli* strain χ2913. H.M. acknowledges financial support from the European Molecular Biology Organization, the Fondation des Treilles, and CNRS. Supported by the Association pour la Recherche sur le Cancer (ARC) and by the programme "Recherche Fondamentale en Microbiologie et Maladie Infectieuse et Parasitaires" (PRFMMIP) from the French Ministry of Science and Education.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1072113/DC1
Materials and Methods
Fig. S1
Table S1

21 March 2002; accepted 23 April 2002
Published online 23 May 2002;
10.1126/science.1072113
Include this information when citing this paper.