

Supplementary Information

Conservation and Role of Electrostatics in Thymidylate Synthase

Divita Garg^{a,b,c,1}, Stephane Skouloubris^{d,e}, Julien Briffotiaux^d, Hannu Myllykallio^d, Rebecca C. Wade^{a,f,g}

^aMolecular and Cellular Modeling Group, Heidelberg Institute for Theoretical Studies (HITS), Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany; ^bInstitute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany; ^cMunich Center for Integrated Protein Science, Biomolecular NMR Spectroscopy, Department Chemie, Technische Universität München, Lichtenbergstrasse 4, 85747 Garching, Germany; ^dLaboratoire d'Optique et Biosciences, INSERM U1182, Centre National de la Recherche Scientifique Unité Mixte de Recherche 7645, Ecole Polytechnique, Palaiseau Cedex, France; ^eUniversity of Paris-Sud, Orsay, F-91405; ^fCenter for Molecular Biology (ZMBH), Heidelberg University, 69120 Heidelberg, Germany; ^gInterdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Heidelberg, Baden-Württemberg, Germany

¹ Present address: The Scripps Research Institute, Department of Molecular and Experimental Medicine, 10550 North Torrey Pines Road, MEM-131, La Jolla, CA 92037, USA

The Supplementary Information consists of

- 1 PDF document containing Tables S1 and S2, Figures S1-S4, and a list of the proteins studied, and
- Two coordinate files in pqr format: P04818.pqr for human TS, and P0A886.pqr for *E. coli* TS. The coordinate files have hydrogen atoms added, atomic partial charges and radii for the electrostatic potential calculations, and give the coordinate system for defining the region for the PIPSA analysis defined in Figure S1. The human and *E. coli* TS structures were used as templates for homology modeling. The residue numbering scheme in these files corresponds to those of the corresponding PDB files, 1HVY (human) and 2G8O (*E.coli*).

Table S1. Values of the Hodgkin similarity index for the electrostatic potential at the active site of TS of various organisms modeled against the crystal structure of human TS (PDB ID 1HVY). Human and *E. coli* TS were chosen as representatives of eukaryotes and prokaryotes, respectively.

	P04818 (human)	P0A886 (<i>E. coli</i>)	Q8D2N4 (W.g.b)	P59427 (B.b.p)	Q8K9C3 (B.s.g)	W.g.b_ mut
P04818	1	0.849	-0.080	-0.088	-0.002	0.635
P0A886	0.849	1	-0.466	-0.320	-0.254	0.389
Q8D2N4	-0.080	-0.466	1	0.794	0.851	0.363
P59427	-0.088	-0.320	0.794	1	0.930	0.552
Q8K9C3	-0.002	-0.254	0.851	0.930	1	0.593
W.g.b_ mut	0.635	0.389	0.363	0.552	0.593	1

Wigglesworthia glossinidia brevipalpis (W.g.b.), *Buchnera aphidicola* subsp. *Baizongia pistaciae* (B.b.p), and *Buchnera aphidicola* subsp. *Schizaphis graminum* (B.s.g). W.g.b_mut=W.g.b TS quadruple mutant K23G/K82E/K86E/K257G

Table S2. Primer sequences used to mutate the pQE80L_EcoliWT to code the four single mutants of *E. coli* TS.

Mutation		Primer Sequence
G23K	Forw.	cagaaaaacgaccgtaccAAAaccggaacgctttccat
	Rev.	atggaaagcgttccgggTTTggtacgggctgtttttctg
E82K	Forw.	aatgtcaccatctgggacAAAatgggcccgatgaaaacg
	Rev.	cgttttcatcggcccaTTTgtcccagatgggtgacatt
E86K	Forw.	tgggacgaatgggcccgatAAAaacggcgacctcgggc
	Rev.	gcccgagggtcgccggtTTTatcggcccatcgtccca
G257K	Forw.	aaggctacgatccgcacccgAAAattaaagcgccgggtggctat
	Rev.	atagccaccggcgctttaatTTTcggatgcggatcgtagcctt

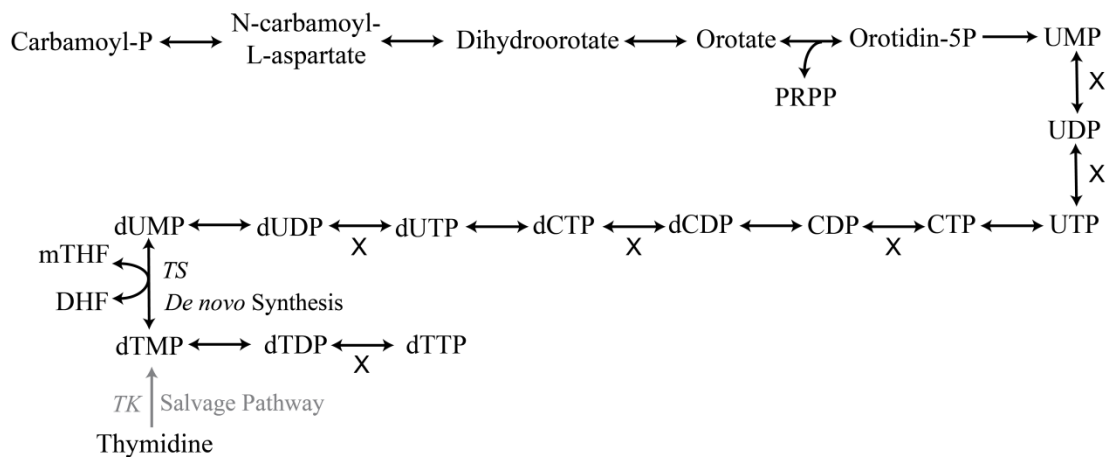


Figure S1. Representative thymidine synthesis pathway and thymidine salvage pathway. Black arrows represent the steps that are catalyzed by proteins whose coding sequences are present in the W.g.b. genome, whereas crosses denote the steps whose enzymes are lacking coding sequences in the *Buchnera* genomes. The grey arrow represents the step missing in both W.g.b. and *Buchnera* due to the lack of a coding sequence for thymidine kinase. The *de novo* synthesis of dTMP catalyzed by thymidylate synthase (TS) and the salvage pathway catalyzed by thymidine kinase (TK) are labeled.



Fig S2. Sequence alignment of *E. coli*, W.g.b. and human TS proteins. The sequences are from UniProt (*E. coli*: P0A886; W.g.b.: Q8D2N4; human: P04818) and numbered accordingly. The background colors of the letters indicate residue type. The Small Domain (SD) and the Interface Loop (IL) are marked by horizontal blue and orange lines, respectively.

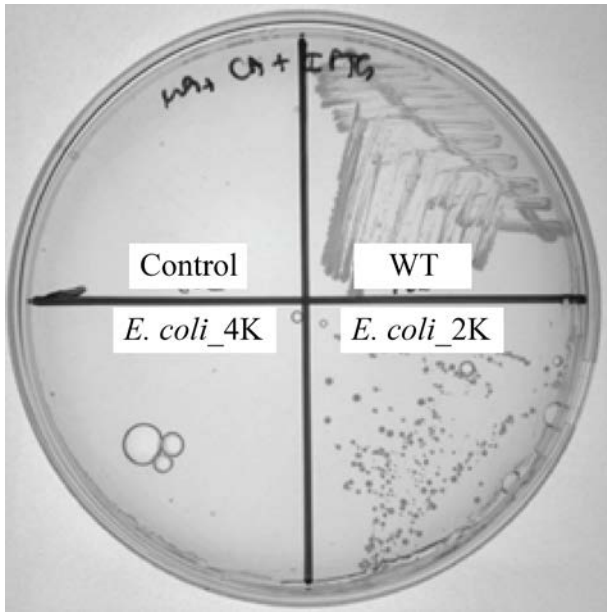
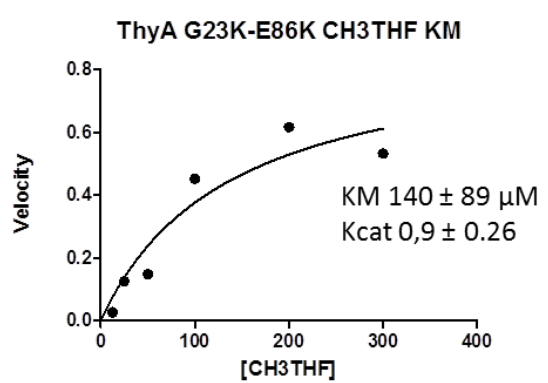
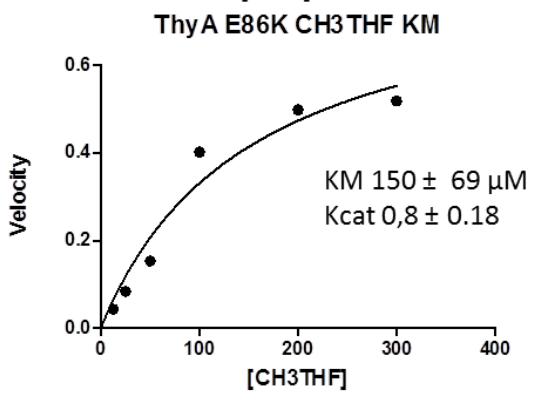
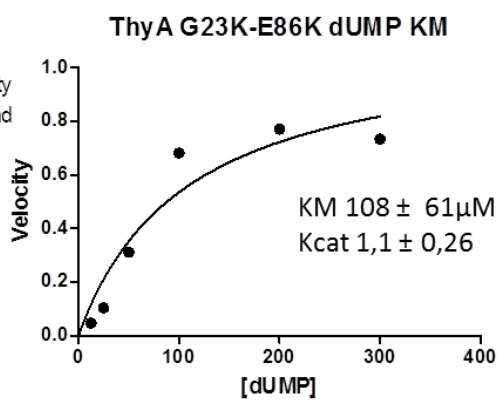
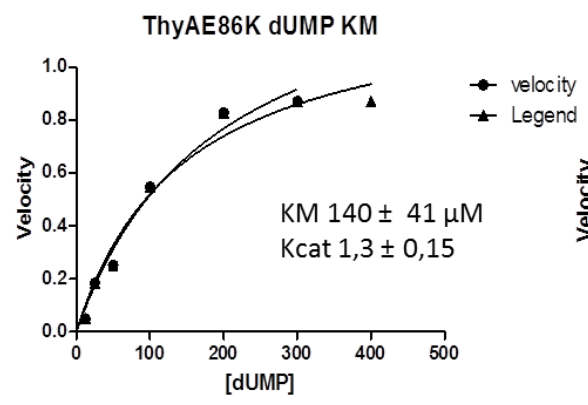
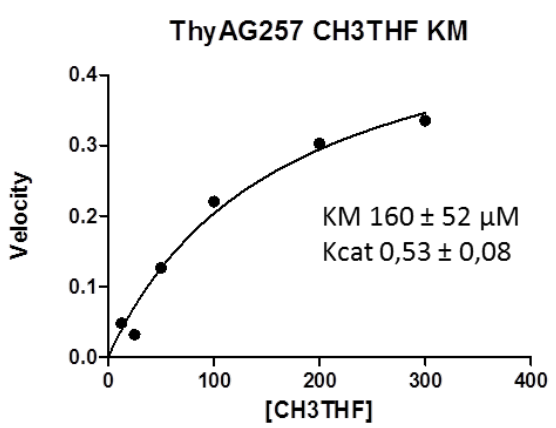
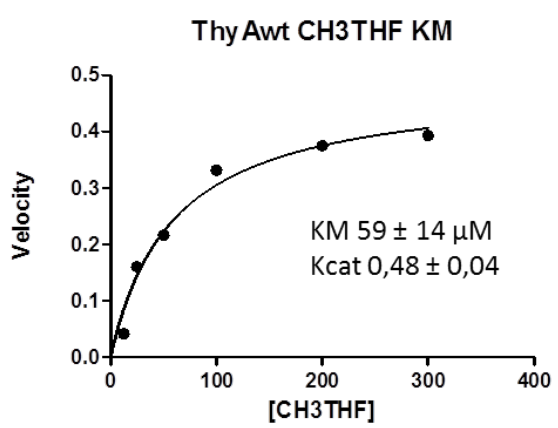
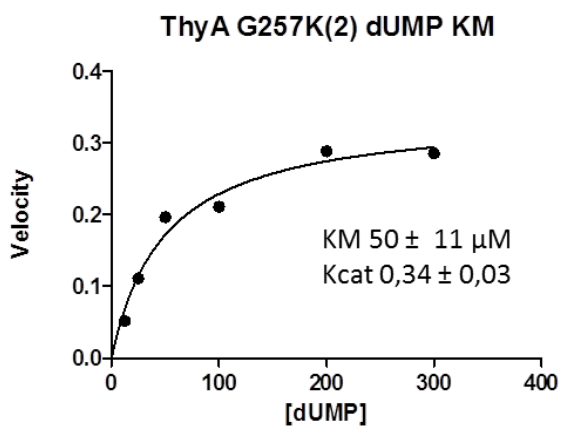
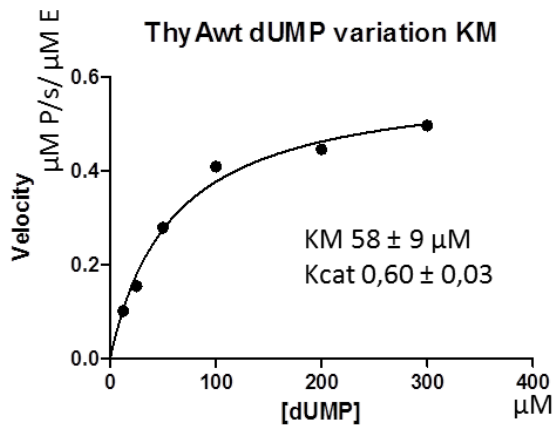


Fig S3. Cell complementation tests for ThyA-negative *E. coli* FE013 (pQE80L) as a control, and for pQE80L expressing the WT TS (*E. coli*_WT), the *E. coli* double mutant G23K/E86K (*E. coli*_2K), and the *E. coli* quadruple mutant G23K/E82K/E86K/G257K TS (*E. coli*_4K). See Table 2.



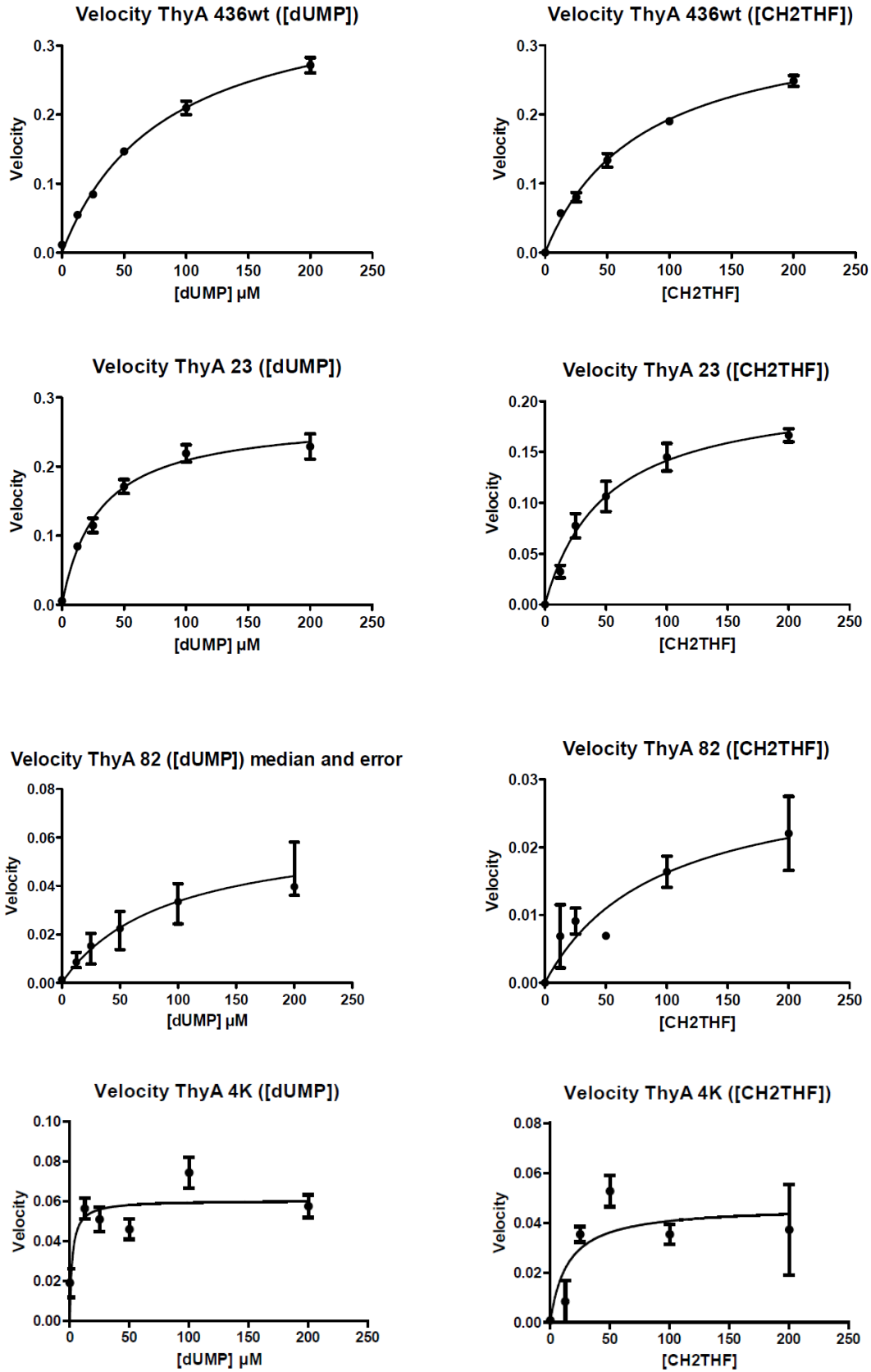


Fig S4. Measured kinetic activities for *E. coli* WT and mutant TS enzymes used to derive K_M and K_{cat} values given in Table 3. Error bars on measurements were derived from three independent measurements.

List of Uniprot IDs for the 110 TS proteins analysed, together with the organism name and taxonomy.

*The sequence for TS of tsetse fly was downloaded from the webpage of Prof. Serap Aksoy, Yale School of Public Health, U.S.A.

Eukaryota

P04818	<i>Homo sapiens</i>	Metazoa	Chordata
P45352	<i>Rattus norvegicus</i>	Metazoa	Chordata
P07607	<i>Mus musculus</i>	Metazoa	Chordata
O96650	<i>Ascaris suum</i>	Metazoa	Ecdysozoa
O76511	<i>Drosophila melanogaster</i>	Metazoa	Ecdysozoa
*	Tsetse fly	Metazoa	Ecdysozoa
O62584	<i>Encephalitozoon cuniculi</i>	Fungi	Microsporidia
P13100	<i>Pneumocystis carinii</i>	Fungi	Dikarya
Q9P4T7	<i>Agaricus bisporus</i>	Fungi	Dikarya
P45351	<i>Cryptococcus neoformans</i>	Fungi	Dikarya
P06785	<i>Saccharomyces cerevisiae</i>	Fungi	Dikarya
O81395	<i>Zea mays</i>	Viridiplantae	Streptophyta
Q2QRX6	<i>Oryza sativa</i>	Viridiplantae	Streptophyta
P51820	<i>Glycine max</i>	Viridiplantae	Streptophyta
P45350	<i>Daucus carota</i>	Viridiplantae	Streptophyta
Q05762	<i>Arabidopsis thaliana</i>	Viridiplantae	Streptophyta
Q27828	<i>Paramecium tetraurelia</i>	Alveolata	Ciliophora
O02604	<i>Plasmodium vivax</i>	Alveolata	Apicomplexa
P13922	<i>Plasmodium falciparum</i>	Alveolata	Apicomplexa
Q27713	<i>Plasmodium berghei</i>	Alveolata	Apicomplexa
P20712	<i>Plasmodium chabaudi</i>	Alveolata	Apicomplexa
Q07422	<i>Toxoplasma gondii</i>	Alveolata	Apicomplexa
Q27793	<i>Trypanosoma cruzi</i>	Euglenozoa	Kinetoplastida
Q27783	<i>Trypanosoma brucei</i>	Euglenozoa	Kinetoplastida
P07382	<i>Leishmania major</i>	Euglenozoa	Kinetoplastida
P16126	<i>Leishmania amazonensis</i>	Euglenozoa	Kinetoplastida
Q23695	<i>Crithidia fasciculata</i>	Euglenozoa	Kinetoplastida

Prokaryota

Q6MID2	<i>Bdellovibrio bacteriovorus</i>	Proteobacteria	Deltaproteobacteria
Q8PCE7	<i>Xanthomonas campestris</i>	Proteobacteria	Gammaproteobacteria
Q8PP46	<i>Xanthomonas axonopodi</i>	Proteobacteria	Gammaproteobacteria
Q5GWB5	<i>Xanthomonas oryzae</i>	Proteobacteria	Gammaproteobacteria
Q9I6F1	<i>Pseudomonas aeruginosa</i>	Proteobacteria	Gammaproteobacteria
Q83BG2	<i>Coxiella burnetii</i>	Proteobacteria	Gammaproteobacteria
Q603S2	<i>Methylococcus capsulatus</i>	Proteobacteria	Gammaproteobacteria
Q5X119	<i>Legionella pneumophila</i>	Proteobacteria	Gammaproteobacteria
Q5WSU5	<i>Legionella pneumophila</i>	Proteobacteria	Gammaproteobacteria
Q5ZRL3	<i>Legionella pneumophila</i>	Proteobacteria	Gammaproteobacteria
Q8EH94	<i>Shewanella oneidensis</i>	Proteobacteria	Gammaproteobacteria
Q87BT4	<i>Xylella fastidiosa</i>	Proteobacteria	Gammaproteobacteria

Q9PB13	<i>Xylella fastidiosa</i>	Proteobacteria Gammaproteobacteria
P0A886	<i>Escherichia coli</i>	Proteobacteria Gammaproteobacteria
Q8ZHV1	<i>Yersinia pestis</i>	Proteobacteria Gammaproteobacteria
Q667F9	<i>Yersinia pseudotuberculosis</i>	Proteobacteria Gammaproteobacteria
Q6REU8	<i>Xenorhabdus nematophilus</i>	Proteobacteria Gammaproteobacteria
Q7N8U4	<i>Photorhabdus luminescens</i>	Proteobacteria Gammaproteobacteria
P48464	<i>Shigella flexneri</i>	Proteobacteria Gammaproteobacteria
Q5PEN6	<i>Salmonella paratyphi-a</i>	Proteobacteria Gammaproteobacteria
Q8ZMA9	<i>Salmonella typhimurium</i>	Proteobacteria Gammaproteobacteria
Q8Z412	<i>Salmonella typhi</i>	Proteobacteria Gammaproteobacteria
Q6D8I6	<i>Erwinia carotovora</i>	Proteobacteria Gammaproteobacteria;
P57515	<i>Buchnera aphidicola</i>	Proteobacteria Gammaproteobacteria
P59427	<i>Buchnera aphidicola</i>	Proteobacteria Gammaproteobacteria
Q8K9C3	<i>Buchnera aphidicola</i>	Proteobacteria Gammaproteobacteria
Q8D2N4	<i>Wigglesworthia glossinidia brevipalpis</i>	Proteobacteria Gammaproteobacteria
P44420	<i>Haemophilus influenzae</i>	Proteobacteria Gammaproteobacteria
Q5R064	<i>Idiomarina loihiensis</i>	Proteobacteria Gammaproteobacteria
Q6FER7	<i>Acinetobacter sp</i>	Proteobacteria Gammaproteobacteria
Q2NRH3	<i>Sodalis glossinidius</i>	Proteobacteria Gammaproteobacteria
O33380,	<i>Neisseria gonorrhoeae</i>	Proteobacteria Betaproteobacteria
Q9JT57	<i>Neisseria meningitidis</i>	Proteobacteria Betaproteobacteria
Q82WU3	<i>Nitrosomonas europaea</i>	Proteobacteria Betaproteobacteria
Q9RAM7	<i>Methylobacillus flagellatus</i>	Proteobacteria Betaproteobacteria
Q5P233	<i>Azoarcus sp.</i>	Proteobacteria Betaproteobacteria
Q8Y0U6	<i>Ralstonia solanacearum</i>	Proteobacteria Betaproteobacteria
Q7NZ95	<i>Chromobacterium violaceum</i>	Proteobacteria Betaproteobacteria
P67042	<i>Brucella melitensis</i>	Proteobacteria Alphaproteobacteria
P67043	<i>Brucella suis</i>	Proteobacteria Alphaproteobacteria
Q89G35	<i>Bradyrhizobium japonicum</i>	Proteobacteria Alphaproteobacteria
Q98KH9	<i>Rhizobium loti</i>	Proteobacteria Alphaproteobacteria
Q6N447	<i>Rhodopseudomonas palustris</i>	Proteobacteria Alphaproteobacteria
Q92NQ5	<i>Rhizobium meliloti</i>	Proteobacteria Alphaproteobacteria
Q8UDS3	<i>Agrobacterium tumefaciens</i>	Proteobacteria Alphaproteobacteria
Q9A6H0	<i>Caulobacter crescentus</i>	Proteobacteria Alphaproteobacteria
Q6FZ91	<i>Bartonella quintana</i>	Proteobacteria Alphaproteobacteria
Q6G2S8	<i>Bartonella henselae</i>	Proteobacteria Alphaproteobacteria
Q5YPL7	<i>Nocardia farcinica</i>	Actinobacteria Actinobacteridae
Q6A761	<i>Propionibacterium acnes</i>	Actinobacteria Actinobacteridae
Q6AFI0	<i>Leifsonia xyli subsp. xyli</i>	Actinobacteria Actinobacteridae
Q8G3T9	<i>Bifidobacterium longum</i>	Actinobacteria Actinobacteridae
Q73VZ2	<i>Mycobacterium paratuberculosis</i>	Actinobacteria Actinobacteridae
Q9CBW0	<i>Mycobacterium leprae</i>	Actinobacteria Actinobacteridae
P67044	<i>Mycobacterium tuberculosis</i>	Actinobacteria Actinobacteridae
P67045	<i>Mycobacterium bovis</i>	Actinobacteria Actinobacteridae
Q6NIF2	<i>Corynebacterium diphtheriae</i>	Actinobacteria Actinobacteridae
Q8FR47	<i>Corynebacterium efficiens</i>	Actinobacteria Actinobacteridae
Q8NS38	<i>Corynebacterium glutamicum</i>	Actinobacteria Actinobacteridae

Q7MTB5	<i>Porphyromonas gingivalis</i>	Bacteroidetes	Bacteroidetes
Q8A639	<i>Bacteroides thetaiotaomicron</i>	Bacteroidetes	Bacteroidetes
Q64PV5	<i>Bacteroides fragilis</i>	Bacteroidetes	Bacteroidetes
D0JBG7	<i>Blattabacterium sp. subsp. Blattella germanica</i>	Bacteroidetes	Flavobacteria
D0J8U9	<i>Blattabacterium sp. subsp. Periplaneta americana</i>	Bacteroidetes	Flavobacteria
Q7UID0	<i>Rhodopirellula baltica</i>	Planctomycetes	Planctomycetacia
Q5WDS1	<i>Bacillus clausii</i>	Firmicutes	Bacillales
Q59212	<i>Bacillus licheniformis</i>	Firmicutes	Bacillales
Q65J44	<i>Bacillus licheniformis</i>	Firmicutes	Bacillales
Q9ANR7	<i>Bacillus mojavensis</i>	Firmicutes	Bacillales
P42326	<i>Bacillus subtilis</i>	Firmicutes	Bacillales
Q5D189	<i>Bacillus subtilis</i>	Firmicutes	Bacillales
Q9K7B5	<i>Bacillus halodurans</i>	Firmicutes	Bacillales
P54081	<i>Bacillus amyloliquefaciens</i>	Firmicutes	Bacillales
Q67JQ1	<i>Symbiobacterium thermophilum</i>	Firmicutes	Lactobacillales
Q5KZ25	<i>Geobacillus kaustophilus</i>	Firmicutes	Bacillales
Q97EV3	<i>Clostridium acetobutylicum</i>	Firmicutes	Clostridia
Viruses			
P07606	<i>Bacillus phage phi3T</i>	dsDNA viruses	Caudovirales
P00471	<i>Enterobacteria phage T4</i>	dsDNA viruses	Caudovirales
P90463	<i>Human herpesvirus 8</i>	dsDNA viruses	Herpesvirales
P06854	<i>Saimiriine herpesvirus 2</i>	dsDNA viruses	Herpesvirales
P12462	<i>Herpesvirus ateles</i>	dsDNA viruses	Herpesvirales
Q89940	<i>Equine herpesvirus 2</i>	dsDNA viruses	Herpesvirales
P09249	<i>Varicella-zoster virus</i>	dsDNA viruses	Herpesvirales