Structural and Enzymatic Investigation of the *Sulfolobus solfataricus* Uridylate Kinase Shows Competitive UTP Inhibition and the Lack of GTP Stimulation[†,‡]

Kristine Steen Jensen,[§] Eva Johansson,[||,⊥,#] and Kaj Frank Jensen*[,§]

Department of Biological Chemistry, Institute of Molecular Biology, University of Copenhagen, Sølvgade 83H, DK-1307 Copenhagen K, Denmark, Centre for Crystallographic Studies, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark, and European Synchrotron Radiation Facility, BP220 F38043 Grenoble Cedex, France

Received August 31, 2006; Revised Manuscript Received December 8, 2006

ABSTRACT: The *pyrH* gene encoding uridylate kinase (UMPK) from the extreme thermoacidophilic archaeon *Sulfolobus solfataricus* was cloned and expressed in *Escherichia coli*, and the enzyme (*Ss*UMPK) was purified. Size exclusion chromatography and sedimentation experiments showed that the oligomeric state in solution is hexameric. *Ss*UMPK shows maximum catalytic rate at pH 7.0, and variation of pH only influences the turnover number. Catalysis proceeds by a sequential reaction mechanism of random order and depends on a divalent cation. The enzyme exhibits high substrate specificity toward UMP and ATP and is inhibited by UTP, whereas CTP and GTP do not influence activity. UTP binds to the enzyme with a sigmoid binding curve, whereas GTP does not bind. The crystal structure of *Ss*UMPK was determined for three different complexes, a ternary complex with UMP and the nonhydrolyzable ATP analogue β,γ-methylene-ATP, a complex with UMP, and a complex with UTP to 2.1, 2.2, and 2.8 Å resolution, respectively. One UTP molecule was bound in the acceptor site per subunit, leading to the exclusion of both substrates from the active site. In all cases, *Ss*UMPK crystallized as a hexamer with the main fold shared with other prokaryotic UMPKs. Similar to UMPK from *Pyrococcus furiosus*, *Ss*UMPK has an active site enclosing loop. This loop was only ordered in one subunit in the ternary complex, which also contained an unusual arrangement of ligands (possibly a dinucleotide) in the active site and an altered orientation of the catalytic residue Arg48 relative to the other five subunits of the hexamer.

Uridylate kinase (UMPK[1]) plays a key role in pyrimidine biosynthesis by catalyzing the ATP dependent phosphorylation of UMP, thus forming UDP. UDP may be further phosphorylated to UTP, which is a substrate for RNA polymerase, a precursor for CTP synthesis, and a cofactor in sugar metabolism. Another fate of UDP is the reduction to dUDP as a step in the pathway for the synthesis of dTTP and dCTP.

The UMPKs of bacterial origin differ considerably from the UMPKs of eukaryotic origin. In general, the eukaryotic enzymes have a broader substrate specificity, being able to phosphorylate both UMP and CMP, than the enzymes of bacterial origin, which show a strong preference for UMP (*1*). Another difference between the two classes of enzymes is that the UMPKs of eukaryotic origin are monomeric proteins with sizes around 25 kDa, whereas the UMPKs of bacterial origin are hexamers of identical polypeptides with sizes around 25 kDa (*1*, *2*). The recent determination of the crystal structures of several UMPKs of prokaryotic origin have shown that the overall fold of the prokaryotic monomer differs significantly from the fold of the enzymes of eukaryotic origin (*3*, *4*).

In general, the catalytic mechanism for nucleoside monophosphate kinases is sequential with no obligatory order for binding substrates and departure of products (*1*). The phosphoryl transfer from ATP to UMP is believed to be of mainly associative nature with a penta-coordinated transition state (*5*), although the enzyme's ways of stabilizing the transition state and the geometry of the transition state remain unclear (*1*, *6*).

The regulation of the activity of bacterial UMPKs is complex, involving feedback inhibition by UTP and activation by GTP (*2*, *7*, *8*). The mechanism by which GTP

[†] This work was supported by grants from the Danish National Science Research Council and the European Community−Research Infrastructure Action under the FP6 program "Structuring the European Research Area". We also acknowledge the contribution of the Danish Natural Science Council to DANSYNC.

[‡] The coordinates (and structure factors) have been deposited in the Protein Data Bank with accession codes 2j4j (r2j4jsf), 2j4k (r2j4ksf), and 2j4l (r2j4lsf) for the *S. solfataricus* uridylate kinase in complex with UMP and AMPPCP, UMP, and UTP, respectively.

* To whom correspondence should be addressed. Tel: +45 3532 2020. Fax: +45 3532 2040. E-mail: kfj@mermaid.molbio.ku.dk.

[§] Department of Biological Chemistry, University of Copenhagen.

[||] Center for Crystallographic Studies, University of Copenhagen.

[⊥] European Synchrotron Radiation Facility.

[#] Present address: Diabetes Protein Engineering, Novo Nordisk A/S, Novo Allé 6, DK-2880 Bagsværd, Denmark.

[1] Abbreviations: UMPK, uridine 5′-monophosphate kinase (EC 2.7.4.22); *Ss*UMPK, *S. solfataricus* uridine 5′-monophosphate kinase; *Ss*UMPK−UMP, *Ss*UMPK in complex with uridine 5′-monophosphate; *Ss*UMPK−UMP−AMPPCP, *Ss*UMPK in complex with uridine 5′-monophosphate and β,γ-methylene adenosine 5′-triphosphate; *Ss*UMPK−UTP, *Ss*UMPK in complex with uridine 5′-triphosphate; buffer A, 25 mM Tris-HCl and 0.1 mM EDTA at pH 7.6; PCR, polymerase chain reaction; PEI-plates, poly(ethyleneimine)-impregnated cellulose thin-layer plates on plastic sheets.

10.1021/bi0618159 CCC: $37.00
Published on Web 02/13/2007

stimulates UMPKs, however, seems to differ from organism to organism. Although GTP appears to enhance k_{cat} for the enzymes of *E. coli* and *Bacillus subtilis* (*2*, *8*), it was proposed to reduce K_M for ATP in the case of the enzyme from *Streptococcus pneumoniae* (*7*). To our knowledge, no reports indicate the existence of allosteric regulation of the activity of eukaryotic UMP/CMP kinases.

Because UMPKs of eukaryotic and bacterial origin have proven to be so different from each other, we decided to investigate the enzyme from *S. solfataricus* as the model for UMPKs of archaeal origin. Before we completed our work, the crystal structure of UMPK from the archaeon *P. furiosus* was published (*3*), but to date no UMPK of archaeal origin has been characterized with respect to reaction kinetics, substrate specificity, or allosteric regulation.

In this article, we report that the UMPK from *S. solfataricus* is feedback inhibited by UTP but not stimulated by the purine nucleotide GTP. We also report the crystal structure of *Ss*UMPK in complex with UMP and the nonhydrolyzable ATP analogue β,γ-methylene-ATP (*Ss*UMPK−UMP−AMPPCP), with UMP (*Ss*UMPK−UMP), and with UTP (*Ss*UMPK−UTP) and discuss the structures in context with results obtained by kinetic studies.

MATERIALS AND METHODS

Materials. The nucleoside triphosphates were from Amersham Biosciences (Uppsala, Sweden). β,γ-Methylene adenosine 5′-triphosphate and other nucleotides were from Sigma (St. Louis, MO). [2-^{14}C]uridine 5′-monophosphate (UMP) and [2-^{14}C]uridine 5′-triphosphate (UTP) were from Moravek Biochemicals, [^{32}P]adenosine 5′-triphosphate[γP] (ATP) was from Amersham Pharmacia Biotech AB (Uppsala, Sweden). [8-^{14}C]adenosine 5′-diphosphate (ADP) was from ICN MP Biomedicals. Restriction enzymes and T4-DNA ligase were from Promega (Madison, WI), and VentR DNA polymerase was from New England Biolabs (Beverly, MA). Purified chromosomal DNA of the *S. solfataricus* P2 strain was a gift from Dr. Q. She (Institute of Molecular Biology, University of Copenhagen, Denmark). Yeast alcohol dehydrogenase and bovine liver catalase were from Boehringer (Mannheim, Germany). Poly(ethyleneimine)-impregnated cellulose thin-layer plates on plastic sheets (PEI-plates) were made as previously described (*9*). Dyematrex Gel Red A was from Millipore Corporation (Bedford, MA), diethylaminoethyl cellulose DE52 was from Whatman International, Ltd. (Maidstone, Kent, U.K.), Q6 was from BIO-RAD (Hercules, CA), and pre-packed Sephadex G-25 (PD10) and Superose12 columns were from Amersham Biosciences (Uppsala, Sweden). Ultrafree-MC centrifugation filter devices (30,000 nominal molecular weight limit) were from Amicon, Millipore Corporation (Bedford, MA). Other fine chemicals were from Merck (Darmstadt, Germany) or Sigma (St. Louis, MO). The bacterial strain NF1830 (*E. coli* K-12, *recA1/F′lacI^{q} lacZ::*Tn5) has been described previously (*10*).

[^{32}P]UDP[βP] was prepared from UMP and [^{32}P]ATP[γP] using *Ss*UMPK to catalyze the reaction. The synthesis was carried out by the incubation of 5 μM [^{32}P]ATP[γP] (40 Ci/mmol) and 100 μM UMP with 42.5 ng of *Ss*UMPK at 60 °C for 20 min. The reaction was stopped by the addition of HCOOH to a final concentration of 0.3 M. [^{32}P]UDP[βP] was isolated by chromatography on PEI-plates in 1 M HCOOH and 0.5 M LiCl. In this system, ATP, ADP, UDP, and UMP are separated. [^{32}P]UDP[βP] was identified by exposure of an X-ray film. The lane containing [^{32}P]UDP[βP] was cut out and linked to a piece of Whatman 3 paper. Then [^{32}P]UDP[βP] was eluted on the paper by development in 2 M NH_3. Finally, [^{32}P]UDP[βP] was isolated from the paper by centrifugation in a Sorvall SS34 rotor. The final yield of [^{32}P]UDP[βP] was 2.11 μmol in 1200 μL, which corresponds to 84.5% conversion of ATP.

Construction of Expression Vectors. The open reading frame of the *pyrH* gene of *S. solfataricus* was amplified by standard polymerase chain reaction (PCR) techniques using Vent polymerase and chromosomal DNA of the *S. solfataricus* strain P2 as the template. Two expression vectors, pKSJ1A and pKSJ4B, were constructed by cloning the DNA products into the plasmid vector pUHE23-2 (*11*) so that transcription was brought under the control of the isopropyl-β-D-thiogalactoside-inducible $P_{A1/04}$ promoter. The two vectors only deviated from each other with respect to the choice of start codon for *pyrH*, pKSJ1A beginning with two methionine codons (ATGATG) and pKSJ4B starting with a single methionine codon. To construct the plasmid pKSJ1A, the oligo nucleotide 5′CCGAATTCAGGAGAGAAATAAtgatgaacattattttaaagataag was used as an upstream primer for the PCR reaction, and to construct pKSJ4B, the oligo nucleotide 5′CCGAATTCAGGAGAGAAATAatgaacattatttaaagataag was used instead. The downstream primer was in both cases the oligo nucleotide 5′ACGGGATCCTTAtcacactggctctattattgatgaaacc. The upstream primers contained an *Eco*RI restriction site and a ribosome-binding site, and the downstream primer contained a stop codon and a *Bam*HI restriction site. Lower case letters indicate the parts of the primers homologous to the *S. solfataricus* DNA sequences. The resulting PCR fragments were digested with endonucleases *Eco*RI and *Bam*HI and ligated into plasmid pUHE23-2, previously cut with the same restriction endonucleases and treated with alkaline phosphatase. The ligation mixture was transformed into the host *E. coli* strain NF1830, which overproduces the LacI repressor protein, keeping transcription of the cloned genes repressed, and plated on Luria−Bertani broth agar plates (*12*) supplemented with ampicillin (100 μg/mL). Selected colonies were grown for plasmids preparation and production of *Ss*UMPK. Sequencing of the chosen plasmids by standard techniques revealed that the *pyrH* gene on pKSJ1A was as expected from the published genome sequence (*13*), whereas the *pyrH* gene on pKSJ4B had suffered from a PCR generated but silent mutation, changing the triplet encoding Leu83 from TTG to TTA.

Cell Growth and Purification. Strain NF1830 transformed with pKSJ4B was grown in 5 L of Luria−Bertani broth (*12*) supplemented with ampicillin (100 μg/mL) at 37 °C with vigorous aeration. At $OD_{436} = 0.5$, isopropyl-β-D-thiogalactoside was added to a final concentration of 0.5 mM to induce transcription of the cloned *pyrH* gene, and growth was continued to stationary phase overnight. The cells (18 g) were harvested by centrifugation at 5000 rpm for 20 min in a Sorvall SS34 rotor and washed in 0.9% NaCl, centrifuged, re-suspended in 100 mM Tris-HCl and 2 mM EDTA at pH 7.6, and disrupted by ultrasonic treatment. The extract was cleared by centrifugation at 5000 rpm, giving a crude extract. The crude extract was heated for 10 min at 70 °C, and the denatured proteins were removed by centrifugation at

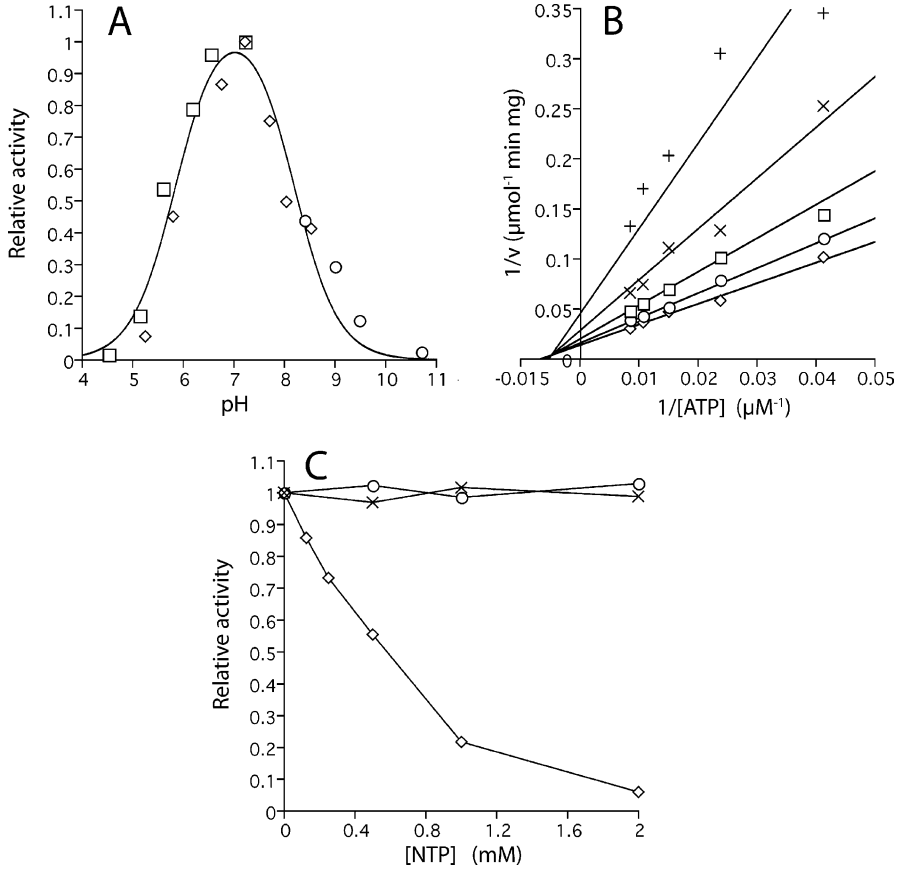

FIGURE 1: Kinetic properties of *Ss*UMPK. (A) pH optimum. Succinate/phosphate buffer (□); phosphate buffer (◊); glycine buffer (○). pH values are pH at 50 mM buffer at 60 °C. Assays were performed with 100 μM UMP, 500 μM ATP, and 10 mM $MgCl_2$. The graph was obtained from the data points by application of least-squares methods according to the equation parameter $P/(1 + 10^{pKa-pH} + 10^{pH-pKb})$. The parameters describing the variation of k_{cat} with respect to pH are $pK_a = 5.8 \pm 0.1$ and $pK_b = 8.2 \pm 0.1$. (B) Reciprocal reaction velocities as a function of the reciprocal concentrations of ATP measured at different fixed concentrations of UMP. [UMP] = 64 μM (◊); [UMP] = 32 μM (○); [UMP] = 16 μM (□); [UMP] = 8 μM (×); [UMP] = 4 μM (+). All data were fitted to the rate equation for a sequential reaction mechanism: $v = V_{max}/(1 + (K_{M,A}/[A]) + (K_{M,B}/[B]) + (K_{iA}K_{M,B}/[A][B]))$, in which v is initial reaction velocity, and $V_{max} = k_{cat}[E]_t$ is the maximum velocity at a given enzyme concentration, $[E]_t$ is the total enzyme concentration given as monomers, k_{cat} is the turnover number, [A] is the concentration of the first substrate to bind, [B] is the concentration of the second substrate to bind, $K_{M,A}$ is the K_M for substrate A, $K_{M,B}$ is the K_M for substrate B, and K_{iA} is the dissociation constant for the enzyme−substrate A complex. The parameters describing the data are $V_{max} = 45 \pm 2$ μmol min^{-1} mg^{-1} ($k_{cat} = 19$ s^{-1}), $K_{M,ATP} = 81 \pm 13$ μM, $K_{M,UMP} = 14 \pm 2$ μM, $K_{i,ATP} = 98 \pm 27$ μM, and $K_{i,UMP} = 17 \pm 5$ μM. (C) Effect of nucleoside triphosphates (NTPs) on initial steady-state rates. Addition of GTP (○); addition of CTP (×); addition of UTP (◊). The assays were performed with 10 mM $MgCl_2$, 27 μM UMP, and 250 μM ATP in an otherwise standard assay. The parameters and plots were obtained by using the Biosoft program Ultrafit.

10,000 rpm for 20 min. *Ss*UMPK was then precipitated by the addition of solid $(NH_4)_2SO_4$ to give 60% saturation. The solution was stirred at 4 °C for 60 min. The precipitated protein was collected by centrifugation, redissolved, and dialyzed against 25 mM Tris-HCl and 0.1 mM EDTA at pH 7.6 (buffer A). The fraction was applied to a 30 cm long column of 1.6 cm diameter of DEAE-cellulose equilibrated with buffer A. *Ss*UMPK was eluted by a NaCl gradient (0−300 mM) and appeared from the column around 160 mM NaCl, and fractions of 10 mL were collected. The fractions containing the most *Ss*UMPK were pooled. Following dialysis against buffer A, the enzyme solution was applied to a 15 cm long column of 1.6 cm diameter of Dyematrex Gel Red A equilibrated with buffer A and eluted in a salt gradient (0−1 M NaCl). *Ss*UMPK was eluted as a broad peak and dialyzed against buffer A. Finally, the enzyme was applied to an anion exchange Q6 column and eluted by a steep salt gradient (0−300 mM NaCl, volume 30 mL). *Ss*UMPK appeared as a sharp nucleotide-free peak with a tail still containing some bound nucleotide. Fractions containing the major peak of protein were pooled, dialyzed against buffer A with a final content of 50% (v/v) glycerol, and stored at −18 °C.

Enzyme Assays. A radioactive assay was used to measure *Ss*UMPK activity in the direction of UDP synthesis because the more commonly used photometric assay using auxiliary enzymes (*2*) is incompatible with high temperature. The standard assay was performed at 60 °C and the reaction mixture (50 μL) contained 50 mM Na_2HPO_4/NaH_2PO_4 at pH 7.1, 10 mM $MgCl_2$, 100 μM UMP (5 Ci/mol), and 500 μM ATP. The mixture was preheated for 5 min at 60 °C. Reactions were started by the addition of 5−10 ng of the enzyme in a 10 μL volume (time = 0). Samples of 10 μL were withdrawn from the reaction mixture after 1, 2, and 5 min and applied instantaneously to PEI-plates. The plates were developed in 0.9 M CH_3COOH and 0.3 M LiCl to separate UDP from UMP. The formation of radioactive UDP was determined by counting in a Phosphor Imager (Cyclone Storage Phosphor Screen, Packard Bioscience) or Instant Imager (Canberra Packard, Merriden, CT). When different substrate and enzyme concentrations were used, the specific details are given in the legends to Figures. To test the

specificity for phosphate acceptor, [2-^{14}C]UMP was replaced by other non radioactive nucleoside monophosphates (CMP, dUMP, AMP, GMP, and TMP) at a concentration of 100 μM, ATP was replaced by 500 μM [^{32}P]ATP[γP] (0.2 Ci/mol), and the enzyme concentration was enhanced 10, 100, and 1000 times. Otherwise, the assays were identical to the standard assay. The specificity of the enzyme toward phosphoryl donors was tested by replacement of ATP with 500 μM dATP, GTP, CTP, and UTP, respectively, in the standard assay. Aside from enhancement of the enzyme concentration 10, 100, and 1000 times, the assays were performed similar to the standard assay.

pH Optimum. The influence of pH on enzyme function (Figure 1A) was analyzed by measuring the activity at 60 °C using 100 μM UMP and 500 μM ATP as substrates in 50 mM buffers of various pH containing 10 mM $MgCl_2$. In the pH interval 4.5−7.2, Na_2HPO_4/succinic acid buffers were used, in the pH interval 5.2−8.5, phosphate buffers were used, and in the pH interval 8.4−10.7 glycine/NaOH buffers were used. The final pH of the buffers diluted to 50 mM was measured at 60 °C. Apparent K_M and V_{max} values were determined at pH 5.5, 7.2, and 8.5 in sodium phosphate buffers. Five different concentrations of UMP (12.5, 25, 50, 100, and 200 μM) were used in combination with 500 μM ATP, and five concentrations of ATP (31.25, 62.5, 125, 250, and 500 μM) were used in combination with 200 μM UMP.

Ligand-Binding Experiments. The radioactive ligands (4−200 μM with specific activity of 2−3 Ci/mol) were incubated with 30−50 μM purified *Ss*UMPK in a total volume of 200 μL in the presence of 10 mM $MgCl_2$ and 50 mM NaH_2PO_4/Na_2HPO_4 at pH 7.1. After assembly at room temperature, the mixtures were heated at 60 °C for 10 min and cooled again at room temperature. Samples (150 μL) of each mixture were transferred to Ultrafree-MC centrifugal filter devices (30,000 nominal molecular weight limit) and centrifuged at 4000*g* for 6 min. The semipermeable filter allows the free ligand to pass through during centrifugation, but retains enzyme−ligand complexes. The concentrations of free and bound ligand were determined by measuring the radioactivity in the noncentrifuged samples and in the liquid that passed through the filter as previously described (*14*). Before use in the ligand-binding experiment, the purified *Ss*UMPK was gel-filtrated on a PD-10 column to remove glycerol. The enzyme concentration used for calculating the binding stoichiometry (Figure 2) may be slightly overestimated, leading to underestimation of the binding stoichiometry because some protein may be lost during the gel-filtration step.

Size Analysis. The size of native *Ss*UMPK was analyzed by sedimentation through 5−20% sucrose gradients at 5 °C using a Beckman SW41 swinging bucket rotor at 39,000 rpm for 23 h as described by Martin and Ames (*15*) and modified as described previously (*16*). The 12 mL sucrose gradients were made up in buffer A. *E. coli* orotate phosphoribosyltransferase (M_r = 47,100) (*17*), beef liver catalase ($S_{20,w}$ = 11.3, M_r = 250.000), and yeast alcohol dehydrogenase ($S_{20,w}$ = 7.5, M_r = 150,000) were included as markers. Following centrifugation, the gradients were tapped in 40 equal sized fractions. The distances of migration from the meniscus were determined by assays of enzyme activity. In addition, gel filtration was performed on a Superose 12 column equilibrated with buffer A and calibrated with yeast alcohol dehydrogenase (M_r = 150,000), dihydroorotate dehydrogenase A from *Lactococcus lactis* (M_r = 68,376) (*18*), and carbonic anhydrase (M_r = 29,000). Protein elution was monitored by UV-absorption measurements.

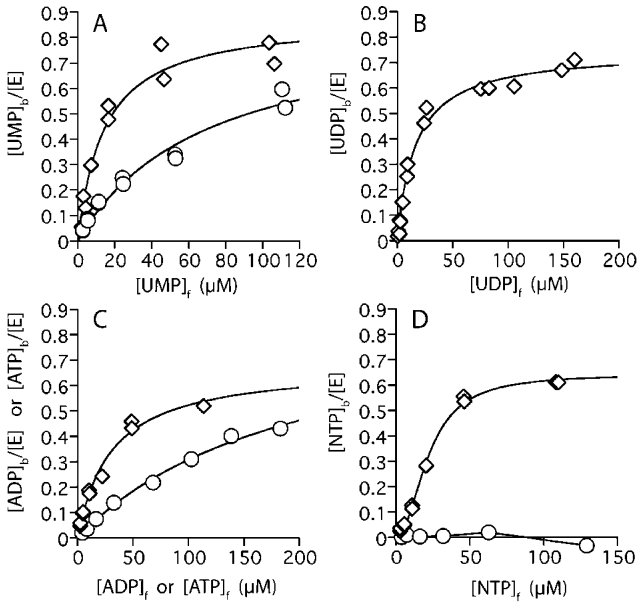

FIGURE 2: Ligand-binding curves. Binding isotherms of ligands to *Ss*UMPK in 50 mM phosphate buffer containing 10 mM $MgCl_2$. (A) UMP binding (◊); UMP binding in the presence of 100 μM UTP (○). (B) UDP binding. (C) ATP binding (◊); ADP binding (○). (D) UTP binding (◊); GTP (○). The UMP, UMP + UTP, ATP, UDP, and ADP binding data were fitted to the equation $[L_b] = (B_{max}[L_f])/(K_D + [L_f])$, whereas the data for the binding of UTP were fitted to the equation $[L_b] = (B_{max}[L_f]^{nH})/(L_{0.5}^{nH} + [L_f]^{nH})$, where $[L_b]$ is the concentration of the bound ligand, $[L_f]$ is the concentration of the free ligand, B_{max} is the maximal concentration of the bound ligand at a given enzyme concentration, K_D is the dissociation constant, $L_{0.5}$ is the concentration of the ligand given half-maximal saturation, and n_H is the Hill coefficient. The parameters describing the data are $K_{D,UMP}$ = 14 ± 3 μM (N = 0.88 ± 0.05); apparent $K_{D,UMP}$ = 70 ± 16 μM in the presence of 100 μM UTP (N = 0.88 ± 0.11); $K_{D,UDP}$ = 17 ± 2 μM (N = 0.75 ± 0.02); $K_{D,ATP}$ = 30 ± 4 μM (N = 0.68 ± 0.04); $K_{D,ADP}$ = 201 ± 39 μM (N = 0.93 ± 0.11); $L_{0.5,UTP}$ = 22 ± 1 μM, and $n_{H,UTP}$ = 2.1 ± 0.1 (N = 0.64 ± 0.02). The parameter N in parenthesis refers to the binding stoichiometry at saturation, which may be underestimated because of the uncertainty of enzyme concentration (Materials and Methods). The plots and parameters were generated by using the Biosoft program Ultrafit.

Crystallization. Solubility screening (*19*) and Crystal Screen I and II from Hampton Research (*20*) were used for initial crystallization screening at room temperature using the hanging drop vapor diffusion technique. Protein solutions (2 μL) in a 10 mM Tris-HCl buffer at pH 7.6 with 4.6 mg/mL *Ss*UMPK, 5 mM $MgCl_2$, and 2 mM UMP, 2 mM UTP, or 2 mM UMP and 2 mM AMPPCP, were mixed with 2 μL of the screening solutions. The crystals used for structure determination were optimized from Crystal Screen II (condition 11) with a protein solution including UMP, AMPPCP, and $MgCl_2$ and the final mother liquor consisting of 0.8 M 1,6-hexanediol, 5 mM $CoCl_2$, and 0.1 M sodium acetate at pH 4.6. These rod shaped crystals appeared after 2−3 days of incubation over 1 mL of mother liquor and grew to a final size of 0.3 × 0.1 × 0.05 mm^3. The second crystal form was grown with UMP and $MgCl_2$ and optimized from Crystal Screen II (condition 34), resulting in a final mother liquor consisting of 0.65 M sodium acetate, 100 mM $CdCl_2$, and

Table 1: Diffraction Data and Refinement Statistics[a]

	*Ss*UMPK−UMP −AMPPCP	*Ss*UMPK−UMP	*Ss*UMPK−UTP
	crystal parameters		
space group	$P2_12_12_1$	$P2_1$	$C2$
a (Å)	86.2	74.8	224.8
b (Å)	126.4	136.0	79.0
c (Å)	135.0	77.9	223.5
β (°)	90	113.2	96.9
molecules in asymmetric unit	6	6	12
Matthews coefficient (Å^3/Da)	2.44	2.41	3.32
solvent content (%)	46.7	46.1	60.8
	data collection statistics		
radiation source	MAX-lab, I711	MAX-lab, I711	MAX-lab, I911−3
temperature (K)	100	100	100
no. of measured reflections	581411 (86072)	243789 (33403)	397894 (51253)
no. of unique reflections	86630 (12494)	72006 (10407)	92163 (12182)
resolution range (Å)	91.29−2.10 (2.21−2.10)	22.54−2.20 (2.32−2.20)	40.52−2.8 (2.95−2.8)
R_{sym} (%)[b]	9.8 (50.2)	10.7 (65.2)	8.0 (54.5)
completeness (%)	99.9 (99.9)	99.3 (98.7)	94.5 (85.9)
mean $I/\sigma(I)$	6.0 (1.4)	12.0 (2.3)	12.7 (1.9)
mulitiplicity	6.7 (6.9)	3.4 (3.2)	4.3 (4.2)
	refinement statistics		
no. of reflections (total)	86521	71829	91357
no. of reflections (working set)	82168	68466	86764
no. of reflections (test set)	4353	3363	4593
number of atoms (total)	10875	8555	19680
resolution (Å)	25.00−2.10 (2.15−2.10)	21.67−2.20 (2.34−2.20)	29.63−2.80 (2.98−2.80)
R_{factor} (%)[c]	21.4 (27.0)	21.88 (28.1)	24.62 (38.8)
R_{free} (%)[d]	24.1 (30.8)	25.12 (29.6)	27.88 (41.8)
average B-factor (Å^2)	28.6	50.1	29.9
bond length rmsd[e] from ideal (Å)	0.013	0.006	0.0085
bond angle rmsd from ideal (deg)	1.59	1.2	1.46
subunit differences[f]	0.09 ≤ rmsd ≤ 0.37	0.53 ≤ rmsd ≤ 2.08	0.22 ≤ rmsd ≤ 1.19
Ramachandran plot (% allowed, generously allowed, and disallowed residues)	98.9, 0.1, 1.0	99.3, 0.4, 0.3	98.5, 1.2, 0.4
refinement program	REFMAC	CNS	CNS

[a] The values in parentheses are data for the highest resolution shell. [b] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where the sums are over all symmetry related reflections of intensity I. [c] $R_{factor} = \Sigma_{work} ||F_{obs}| - k|F_{calc}||/\Sigma_{work}F_{obs}$. [d] $R_{free} = \Sigma_{test} ||F_{obs}| - k|F_{calc}||/\Sigma_{test}F_{obs}$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively, k is the scale factor, and the sums are over all reflections in the working set and test set, respectively. [e] rmsd = root-mean-square deviation. [f] The differences between subunits were calculated with the program packages CCP4 (*35*) and CCP4i (*36*), and rmsd's are calculated for all Cα-atoms in the subunits. The relatively high limit of rmsd (2.08 Å) between subunits in the *Ss*UMPK−UMP structure is caused by the Cd-ions, which adhere to different positions in the different protein chains of the hexamer.

0.1 M Hepes at pH 7.5. These crystals appeared overnight and grew to a final size of 0.15 × 0.2 × 0.2 mm^3. Finally, the third crystal form was grown with UTP and $MgCl_2$ and optimized from Crystal Screen I (condition 34) with final mother liquor containing 1.8 M sodium formate and 0.1 M sodium acetate at pH 4.8. These crystals grew to a size of 0.3 × 0.3 × 0.2 mm^3 within one week.

X-ray Data Collection and Processing. Diffraction data for all crystals were collected under cryogenic conditions (100 K) at MAX-Laboratory, University of Lund, Sweden. The data from UMP- and the UMP−AMPPCP-containing crystals were collected at beamline I711 (*21*) on a MAR Research CCD detector (165 mm), whereas a MAR Research CCD detector (225 mm) on beamline I911-3 was used for data collection on UTP-containing crystals (*22*). The crystals of *Ss*UMPK in complex with UMP−AMPPCP were cryoprotected in a solution of 20% ethylene glycol added to the mother liquor, and for both the UMP- and UTP-containing crystals, 25% glycerol was added to the mother liquor. Auto indexing, data reduction, and scaling were performed with MOSFLM (*23*) and SCALA (*24*).

Structure Solution and Refinement. The molecular replacement technique as implemented in MolRep (*25*) was used for the structure determination of the *Ss*UMPK in complex with UMP and AMPPCP. The two subunits in the asymmetric unit of *P. furiosus* UMPK (Protein Data Bank code 2BMU) stripped from water molecules and ligands were used as the search model. The correct solution generated a full hexamer with three copies of the search model. This solution was used as the starting model for rigid body refinement, with the six chains as separate rigid bodies, prior to simulated annealing using strict 6-fold NCS in CNS (*26*). Simultaneously, prime and switch with 6-fold NCS in RESOLVE (*27*) was performed. The resulting models and difference electron density maps were used for model building using the graphics program O (*28*). This procedure was repeated, and a clear electron density was seen for UMP, AMPPCP, and a cobalt ion in the active site. Furthermore, a cobalt ion coordinated to three histidine residues, His104 from three subunits related by a 3-fold rotation axis, was detected. These molecules and ions were included in the model, and after a few cycles of refinement and model building, all amino acid residues except from 172 to 184 were located. The strict NCS relationships from the CNS refinement were loosened, and refinement was continued in REFMAC5 (*29*). As the strict NCS relationships were removed, it became obvious that in

one of the six subunits the conformation of the molecules bound in the active site was different, and furthermore, in this subunit the loop containing residues 172−184 that closes around the active site was ordered. This loop was model built, and water was added using ARP-wARP (*30*). After cycles of positional refinement with REFMAC5 using restrained NCS at stretches of amino acid residues of similar conformations in the six chains and model building in O, difference electron density maps of the sixth active site made it possible to identify the content. The molecules that fitted the electron density best was a dinucleotide, probably formed by UMP and AMPPCP, together with one magnesium ion and one water molecule.

The structure of *Ss*UMPK in complex with UMP was also determined using the molecular replacement technique with MolRep (*25*). In this case, a hexamer of *Ss*UMPK in complex with UMP and AMPPCP, stripped from ligands, was used as the search model. Rigid body refinement in CNS (*26*) allowing the six protein chains of the hexamer to move separately was followed by simulated annealing with strict 6-fold NCS and individual *B*-value refinement. Clear difference electron density was seen for UMP, at the same position as in the complex with UMP and AMPPCP, and UMP was added to the model. The strict NCS relationships were released, and several rounds of model building in O and positional and *B*-value refinement in CNS were performed. One magnesium ion coordinated to UMP in subunit A and 18 cadmium ions adhering to the enzyme at different positions outside the active site were included in the model. Refinement was continued in CNS using restrained NCS relationships, and water was added also using CNS (*26*).

Finally, the structure of *Ss*UMPK−UTP was determined in a way similar to that for the UMP complex structure. Difference electron density was seen for UTP in all 12 subunits. The uracil ring of UTP was at the same position as the uracil ring of UMP in the two other structures, and UTP was added to the model in all subunits. The strict NCS relationsships were released, and several rounds of model building in O and positional and *B*-value refinement in CNS were performed. Mg^{2+} ions were added to the model in the active site of seven subunits, and additional rounds of model building in O and positional and *B*-value refinement in CNS were performed.

Details from structure refinement are found in Table 1. The quality of the models was checked with PROCHECK (*31*) and WHATIF (*32*) as refinement progressed.

RESULTS

Cloning, Expression, and Purification. The *pyrH* gene encoding UMPK of *S. solfataricus* is identical to the open reading frame SSO976 in the genomic sequence (*13*). This open reading frame starts with a GUG codon followed by an AUG codon. In principle, both of these codons might serve as initiation codons, generating both a 227 residue protein ($M_r = 25104$) and a 226 residue protein ($M_r = 24973$) differing by one methionyl residue at the N-terminus. Expression vectors producing both the shorter version (pKSJ4B) and the longer version (pKSJ1A) were constructed by cloning PCR amplified copies of the gene into plasmid pUHE23-2 (*11*) so that transcription was brought under the control of the LacI regulated $P_{A1/04}$ promoter as described in Materials and Methods. Both versions of *Ss*UMPK were purified from cultures of strain NF1830 transformed with the plasmids and grown to stationary phase under inducing conditions. The purification procedure included heating at 70 °C, ammonium sulfate precipitation, and column chromatography on Dyematrex Gel Red A and the anion exchange materials DE52 and Q6. Minor amounts (≈2%) of a smaller 15 kDa protein, presumably an *Ss*UMPK degradation product, co-purified with *Ss*UMPK. When extracted from the cells, the *Ss*UMPK contained some bound nucleotide, detected by acid precipitation of protein samples as previously described (*14*). The slow release of this nucleotide during purification resulted in broad protein peaks during elution from the anion exchange columns because the nucleotide-bound forms of the enzyme adhered more strongly to the column material than the free enzyme did. However, when eluted from the Q6 column, which constituted the last step in the purification procedure, the main peak of UMPK was essentially free of the nucleotide, whereas the discarded tail fractions still contained some nucleotide. Because we did not detect any significant differences between the two versions of *Ss*UMPK during purification and initial kinetic characterization, we concentrated on studying the shorter version encoded by pKSJ4B.

General Properties of SsUMPK. The molecular size of the native enzyme was determined by gel-filtration chromatography and sedimentation analyses as described in Materials and Methods. Both analyses gave an M_r for the native protein close to 140 kDa (140 kDa from gel filtration and $S_{20,w} = 6.9$, corresponding 137 kDa from sedimentation), indicative of a hexameric globular structure of the protein in solution.

The dependency of enzyme activity on pH when assayed in the presence of Mg^{2+} (10 mM) using UMP (100 μM) and ATP (500 μM) as substrates is shown in Figure 1A. The rise and decline of activity as a function of pH is properly described by the dissociation of protons from two functional groups with $pK_a = 5.8 \pm 0.1$ and $pK_b = 8.2 \pm 0.1$. The variation of activity as a function of pH resulted mainly from changes in the turnover number (k_{cat}) as seen from the determination of saturation at three distinct values of pH. Thus, V_{max} was 5 ± 1 μmol min^{-1} mg^{-1} at pH 5.5, 45 ± 2 μmol min^{-1} mg^{-1} at pH 7.2, and 12 ± 2 μmol min^{-1} mg^{-1} at pH 8.5, whereas the $K_{M,UMP}$ (measured at 500 μM ATP) and the $K_{M,ATP}$ (measured at 200 μM UMP) changed less than 50%.

Catalysis was strictly dependent on the presence of a divalent cation. The ions Mn^{2+} and Co^{2+} at concentrations of 10 mM were as efficient as Mg^{2+}, the ion chosen for the standard assays, as promoters of enzyme activity. Cd^{2+} also promoted activity, but this ion caused severe precipitation problems in the phosphate buffer used for the assays.

As seen for the bacterial enzyme (*2*), *Ss*UMPK was very selective for UMP as the phosphoryl group acceptor. A minor reaction was also observed with dUMP as the substrate (3.7% of the activity detected with UMP), but no activity was seen with CMP, dTMP, or the purine nucleotides (AMP, IMP, XMP, and GMP). The specificity for the phosphoryl group donor was tested with UMP as the acceptor substrate. ATP was by far the most effective phosphoryl donor, although a very weak activity (less than 1% of ATP activity) was seen with dATP, GTP, CTP, and UTP.

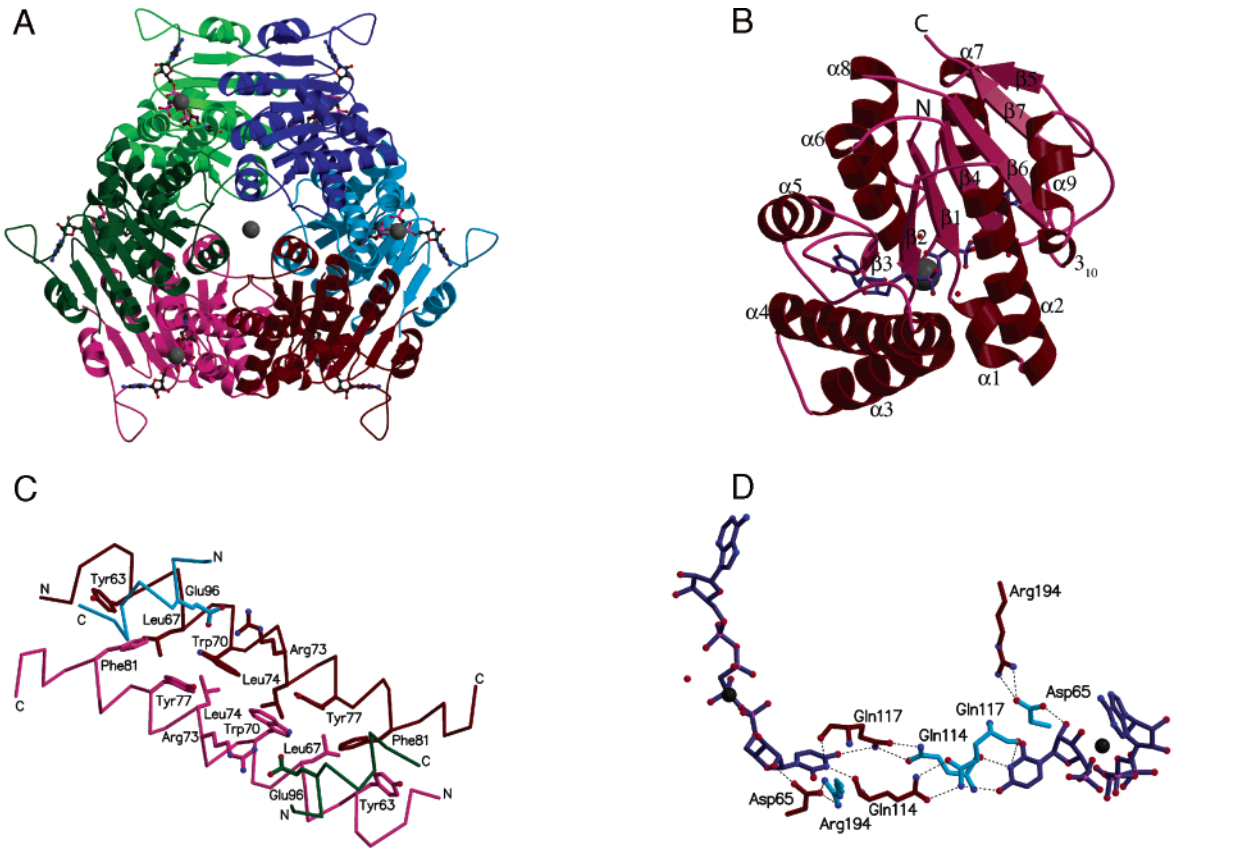

FIGURE 3: *Ss*UMPK structure. (A) *Ss*UMPK−UMP−AMPPCP hexamer viewed down the three fold noncrystallographic axis. Each monomer is colored differently: subunit A, light blue; subunit B, dark blue; subunit C, light green; subunit D, dark green; subunit E, pink; and subunit F, red. (B) *Ss*UMPK−UMP−AMPPCP monomer F with marked secondary structure elements. (C) E/F subunit interface. Hydrophobic interactions between residues of the longest α-helix (helix 4) dominate the interface. (D) A/F subunit interface. Hydrophilic interactions dominate, and the hydrogen bonds between residues of the two neighboring active sites are present. The Figure was prepared using Molscript (*33*) and Raster 3D (*34*).

The initial velocity kinetics was investigated to elucidate the reaction type using different concentrations of both ATP and UMP as substrates at a fixed concentration of Mg^{2+}(10 mM). The results are shown in Figure 1B in the form of double reciprocal plots of reaction velocity as a function of ATP concentration at different, constant concentrations of UMP. The observed pattern of intersecting straight lines agrees with the notion that *Ss*UMPK catalyzes the phosphoryl group transfer by a sequential bi−bi reaction mechanism and, hence, as a nucleophilic substitution reaction while both substrates are bound to the enzyme. The kinetic parameters were $V_{max} = 45$ μmol mg^{-1} min^{-1} ($k_{cat} = 19$ s^{-1}), $K_{M,\ ATP} = 81$ μM, and $K_{M,UMP} = 14$ μM. Furthermore, the K_M values for the two substrates equal the calculated dissociation constants for the enzyme−substrate complexes (the K_{ia} values listed in the legend to Figure 1B), indicating that equilibrium conditions are obeyed. Neither GTP nor CTP influenced the activity of the enzyme, but UTP inhibited the reaction (Figure 1C). UTP acted as a competitive inhibitor against both substrates, and furthermore, a high concentration of UMP (500 μM) abolished the inhibition of UTP at low ATP concentrations (not shown). This indicates that UTP binds to the acceptor site (the UMP site) with β- and γ-phosphates overlapping the donor site because UTP inhibition is expected to persist even at high UMP concentrations if UTP bound to the ATP site.

The ability of the native enzyme to bind reactants and effectors was analyzed in order to further elucidate the reaction order and the mode of action of effectors. All binding studies were performed in the presence of 10 mM $MgCl_2$. Some binding isotherms and parameters describing the binding interactions are shown in Figure 2. All four reactants (ATP, UMP, ADP, and UDP) bound to the free enzyme with hyperbolic binding kinetics, supporting the random mechanism suggested by the kinetic data. No binding of GTP could be detected in the tested range of concentrations, but the enzyme bound UTP in a cooperative manner with a Hill coefficient $n_H = 2$ and the same number of binding sites as seen for the substrates and products. The concentration of UTP needed to saturate half of the sites was 22 μM, and the addition of UTP (100 μM) to the mixtures inhibited the binding of UMP again, indicating that the two ligands share their site on the enzyme.

Crystal Structures. The enzyme was crystallized in the presence of different ligands or combinations of ligands, (a) with UMP alone, giving *Ss*UMPK−UMP crystals, (b) with UMP and the nonhydrolyzable ATP analogue β,γ-methylene-ATP, giving rise to *Ss*UMPK−UMP−AMPPCP crystals, and (c) with the inhibitor UTP, giving *Ss*UMPK−UTP crystals. Diffraction patterns were recorded using synchrotron radiation. The crystals had very different macroscopic appearances and belonged to different space groups. The *Ss*UMPK−UMP crystal, belonging to space group $P2_1$, diffracted X-rays to 2.2 Å resolution; the *Ss*UMPK−UMP−AMPPCP crystal belonged to space group $P2_12_12_1$ and diffracted to 2.1 Å resolution; and the *Ss*UMPK−UTP crystal diffracted to 2.8

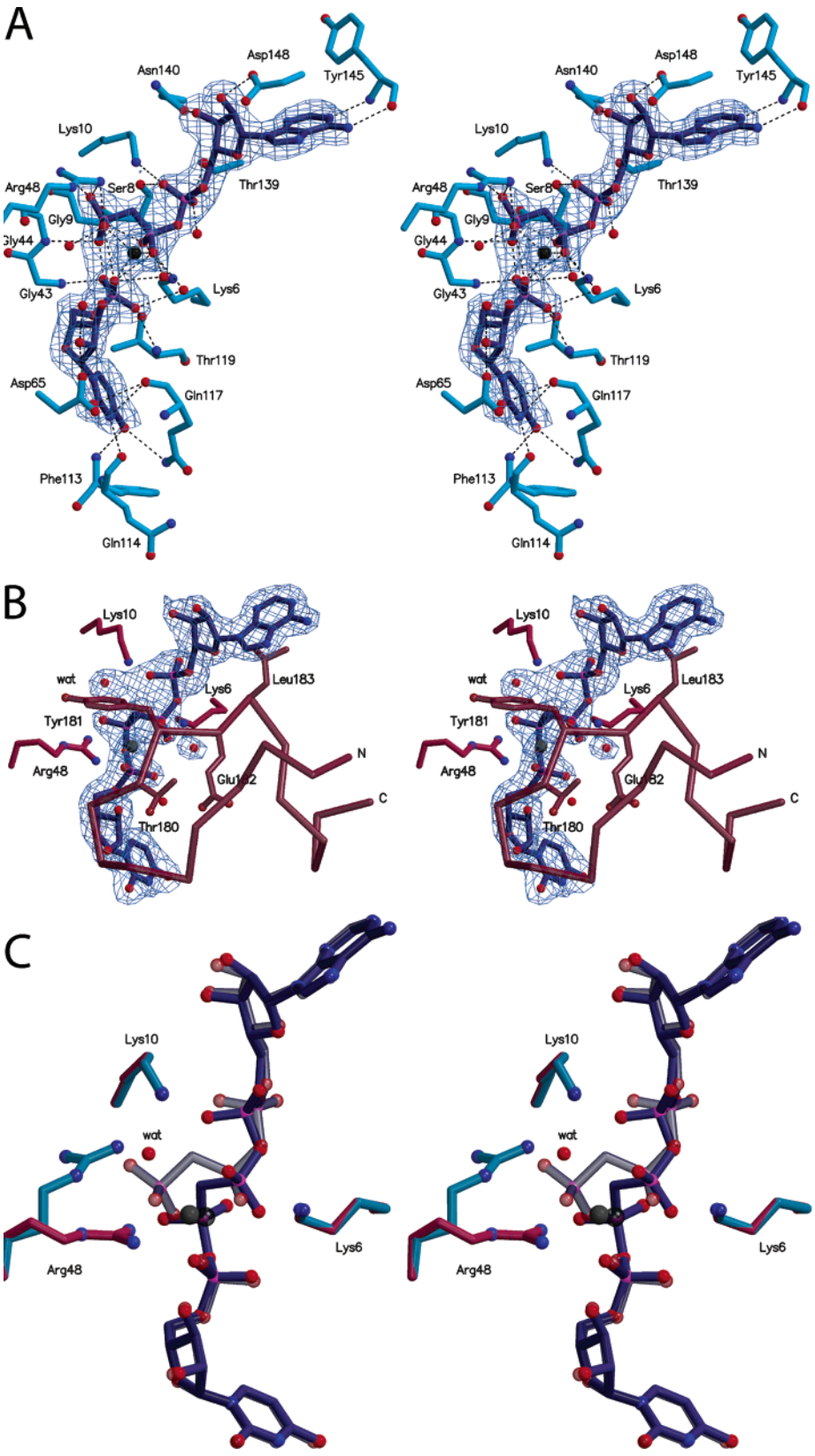

FIGURE 4: Stereoview of active site structures. (A) Active site of subunit A of *Ss*UMPK−UMP−AMPPCP. Hydrogen bonds are indicated as dashed lines. and the $2F_o - F_c$ electron density difference map (sigma level 1) for the ligands is displayed. (B) Active site of subunit F from *Ss*UMPK−UMP−AMPPCP with the modeled dinucleotide and the side chains of the positively charged amino acids shown along with the closed active site loop shown in C_α-trace with the side chains of Thr180, Tyr181, Glu182, and Leu183. The side chain of Tyr181 stacks with the side chain of Arg48, and the side chain of Leu183 stacks with the adenine ring. The $2F_o - F_c$ electron density difference omit map (sigma level 1) for the ligands is displayed. The map was constructed by performing simulated annealing of the final structure with CNS (*26*), while the modeled dinucleotide was omitted. (C) Overlay of the active sites of subunit A and F of *Ss*UMPK−UMP−AMPPCP. Only the positively charged amino acids and the ligands are included for clarity. The figure was prepared using Molscript (*33*) and Raster 3D (*34*).

Å resolution and belonged to space group *C*2. Data statistics are given in Table 1. The crystal structures were determined for all three types of crystals using molecular replacement methods. Details are given in Materials and Methods, and the refinement statistics are shown in Table 1. In all cases, the protein crystallized as a hexamer in agreement with the behavior of the protein in solution. In the complexes with UMP and UMP−AMPPCP, the enzyme crystallized with 6 molecules in the asymmetric unit, and in the complex with UTP, it crystallized with 12 molecules in the asymmetric unit. The structure of the *Ss*UMPK−UMP−AMPPCP hexamer is shown in Figure 3A. The individual subunits are very similar (Table 1), with the main differences in loops surrounding the active site. The structure of the *Ss*UMPK monomer shares the fold with other UMPKs of prokaryotic origin (*3*, *4*), containing seven β-strands, which constitute one central twisted β-sheet with strands 1, 2, 3, 4, and 6 parallel and strands 5 and 7 antiparallel to the other strands. The β-sheet is sandwiched between nine α-helices and one 3_{10}-helix (Figure 3B). There are two primary interfaces between the subunits of the hexamer; one with predominantly hydrophobic interactions (Figure 3C) and one with hydrophilic interactions (Figure 3D). In addition, there are contacts from each monomer to all the other five monomers in the hexamer.

The hydrophobic interaction face is between subunit A/B, C/D, and E/F and consists of stacking side chains in the longest α-helix, helix 4, from each monomer. The aromatic rings from Tyr63, Trp70, Tyr77, and Phe81 in helix 4 from one monomer stack with the same residues of helix 4 from the other monomer. This gives four ring-stacking couples per interaction face. Besides this, two leucine residues from each helix, Leu67 and Leu74, constitute the hydrophobic interactions between the two monomers (Figure 3C). The hydrophilic interface is between subunit A/F, B/C, and D/E. It involves hydrogen bonding between Gln117 and Gln114, Asp65, and Arg194 and between Arg73 and Glu96 from neighboring subunits (Figure 3D). Aside from subunit interactions, Gln117, Gln114, and Asp65 are involved in the binding of UMP (see below).

The *Ss*UMPK−UMP and *Ss*UMPK−UMP−AMPPCP structures contain Cd^{2+} and Co^{2+} ions, respectively, in addition to Mg^{2+}ions. In these two structures, two metal ions, besides the metal ion in the active sites, are situated on a noncrystallographic three fold axis (Figure 3A). Each is coordinated to three of the subunits. One binds to His104 of chains A, C, and E, and the other to His104 of B, D, and F. In addition, Tyr88 and His104 from the same subunit stack with each other, and a hydrogen bond is formed between the hydroxyl group of Tyr88 from subunit A to Gln100 from subunit C. Similar interactions are found for Tyr88 of the C-chain to Gln100 of the E-chain and Tyr88 of the E-chain to Gln100 of the A-chain. Similar interactions are seen for the B/D/F trimer.

In the *Ss*UMPK−UMP−AMPPCP structure, especially one loop called the active site loop, stretching from Gly172 to Leu184 and enclosing the active site, exhibits a large amount of disorder. It was only possible to model this loop in one subunit, namely, subunit F (Figure 4B). The reason for this arrangement may be that crystal packing causes fixation of this loop in subunit F because Gly172 and Ser175 of subunit F are within hydrogen bond distance from Asn16 and Asp18, respectively, of subunit D in a neighboring hexamer. Such interactions are not seen for the other five subunits in which the loops enclosing the active sites are free to move. Parts of the active site loops are ordered in two subunits in the *Ss*UMPK−UTP structure. In the *Ss*UMPK−UMP structure, the active site loop is ordered in three subunits, but here, it is found in a different position conflicting with phosphoryl donor site blocking for the binding of ATP.

In general, there are large similarities between the three complexes, but some differences exist. The loop, which stretches from residue140 to residue158 and contains amino acids involved in the binding of the phosphoryl donor, is found in only one position in all six subunits of the asymmetric unit of the *Ss*UMPK−UMP−AMPPCP structure but is found in many different positions in the two structures that do not contain the phosphoryl donor analogue. Helix 2 is well ordered and identically oriented in all subunits of the *Ss*UMPK−UMP−AMPPCP structure but found in several positions in the two other complexes. Furthermore, the helix could not be modeled in four subunits of the *Ss*UMPK−UTP structure. In these monomers, an electron density corresponding to the helix was apparent, but it was impossible to localize individual amino acids. These observations suggest that the ordering of helix 2 depends on the binding of a ligand in the phosphoryl donor site.

Active Site Structure and Substrate Specificity. The active site constitutes two nucleotide-binding sites, a nucleoside monophosphate site, and a nucleoside triphosphate site (Figure 4). In the *Ss*UMPK−UMP structure, UMP is bound in the active site of all subunits, and in one subunit, the ligand is coordinated to a Mg^{2+} ion. UTP is bound in all subunits in the *Ss*UMPK−UTP structure with highly disordered positions of the β- and γ-phosphoryl groups, reflected in high *B*-values. In seven subunits, the electron density of a Mg^{2+} ion coordinated to UTP is seen. In the *Ss*UMPK−UMP−AMPPCP structure, UMP and AMPPCP are bound in all active sites, together with a divalent metal ion. In four of the subunits, Co^{2+} is the predominant metal ion, but in subunit F and D, the predominant ion is Mg^{2+}. It is, however, possible that a mixture of Co^{2+} and Mg^{2+} occupies all different subunits in the crystal, but because it was not possible to refine the occupancies with the present resolution, the model was built with the metal ion of the highest occupancy.

As explained above, the F subunit in the *Ss*UMPK−UMP−AMPPCP structure differs from the other five subunits by having a well-ordered loop enclosing the active site. Furthermore, the active site of subunit F contained an unusual electron density, which best could be interpreted as a dinucleotide formed by the reaction between the γ-phosphate group of ATP and the 5′-phosphate group of UMP. The electron density around the ligand was reanalyzed by constructing a $2F_o - F_c$ omit map (*26*) and could still best be matched with the dinucleotide (Figure 4B). The position of the UMP part resembles that of UMP in the other five subunits, but the ATP part is repositioned. The γ-phosphate of ATP has changed place, resulting in a more linear arrangement of the four phosphate groups compared to that of the other five subunits. In addition, an electron density corresponding to a water molecule was seen at the position of a γ-phosphoryl oxygen in the other five subunits (Figure

Table 2: Potential Hydrogen Bonds of Ligands in the Active Site of *Ss*UMPK

ligand (atom)	distance[a]	residue/ligand (atom)
UMP (N3)	2.7 Å	Gln114 (O)
UMP (N3)	3.1 Å	Gln117 (O)
UMP (O4)	3.0Å	Phe113 (N)
UMP (O4)	3.1 Å	Gln117 (NE2)
UMP (O2′)	2.7 Å	Asp65 (OD1)
UMP (O2′)	3.1 Å	Wat (O)
UMP (O1P)	2.7 Å	Thr119 (OG1)
UMP (O1P)	2.8 Å	Thr119 (N)
UMP (O1P)	3.1 Å	Wat (O)
UMP (O1P)	2.2 Å	AMPPCP (OG1)
UMP (O2P)	2.8 Å	Gly43 (N)
UMP (O3P)	2.1 Å	Co229 (Co)
UMP (O3P)	3.1 Å	Wat (O)
UMP (O3P)	3.1 Å	AMPPCP (O2b)
UMP (O3P)	3.2 Å	AMPPCP (O2g)
AMPPCP (N1)	2.9 Å	Tyr145 (N)
AMPPCP (N6)	3.0 Å	Wat (O)
AMPPCP (N6)	3.0 Å	Tyr145 (O)
AMPPCP (O2′)	2.7 Å	Asp148 (OD1)
AMPPCP (O3′)	2.6 Å	Asn140 (ND2)
AMPPCP (O1α)	2.6 Å	Lys10 (NZ)
AMPPCP (O1α)	2.9 Å	Ser8 (OG)
AMPPCP (O2α)	2.3 Å	Wat (O)
AMPPCP (O2α)	2.6 Å	Thr139 (OG1)
AMPPCP (O1β)	2.9 Å	Lys6 (NZ)
AMPPCP (O1β)	3.2 Å	Gly9 (N)
AMPPCP (O2β)	2.3 Å	Co229 (Co)
AMPPCP (O2β)	2.8 Å	Glu182 (OE1)
AMPPCP (O2β)	2.9 Å	Wat (O)
AMPPCP (O2β)	3.1 Å	Wat (O)
AMPPCP (O2β)	3.1 Å	UMP (O3P)
AMPPCP (O1γ)	2.2 Å	UMP (O2P)
AMPPCP (O1γ)	3.1 Å	Gly44 (N)
AMPPCP (O2γ)	2.4 Å	Wat (O)
AMPPCP (O2γ)	2.4 Å	Co229 (Co)
AMPPCP (O2γ)	2.7 Å	Arg48 (NH2)
AMPPCP (O2γ)	3.2 Å	UMP (O3P)
AMPPCP (O3γ)	2.6 Å	Arg48 (NH2)
AMPPCP (O3γ)	3.0 Å	Arg48 (NH1)

[a] The distances are for subunit A of *Ss*UMPK−UMP-AMPPCP.

4C). Arg48, which binds to the γ-phosphate group of ATP, has, as the only residue, changed position, and the side chain of Tyr181 protruding from the active site loop stacks with Arg48 in subunit F, entering a position conflicting with the position of Arg48 in the other subunits (Figure 4B and C).

The high specificity for the uracil base is explained by the protein/nucleobase interactions (Figure 4A and Table 2). First of all, the binding site is very narrow whereby purine nucleotides are excluded from binding. Second, specific interactions with the uracil part of UMP are present. There is a hydrogen bond between N3 of UMP and the backbone carbonyl of Gln117. Also, O4 of UMP forms hydrogen bonds to the side chain amido group of Gln117 and to the backbone amide of Phe113. These hydrogen bonds cannot be established with CMP, explaining the lack of substrate activity of this nucleotide. The 2′-OH group of UMP forms a hydrogen bond to the side chain carboxylate of Asp65, and the 3′-OH of UMP is in hydrogen bond distance to water molecules. The interaction between 2′-OH and the protein

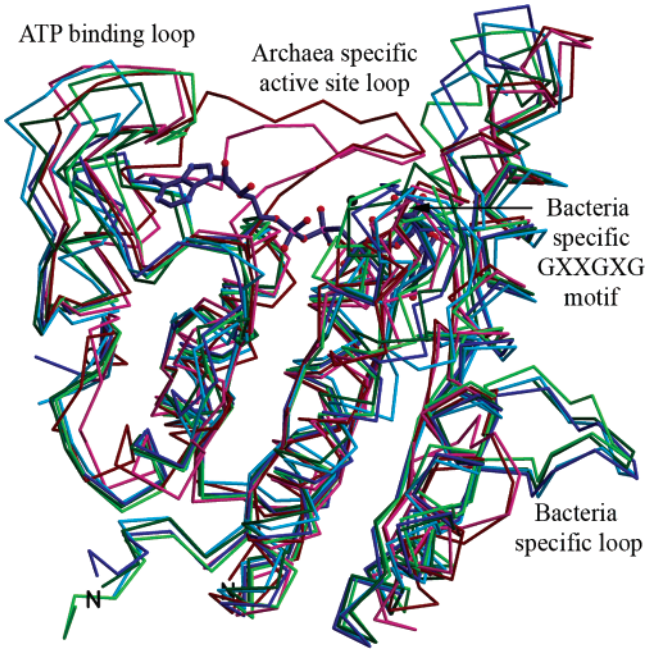

FIGURE 5: Superposition of prokaryotic UMPK Cα-chains. *Ss*UMPK−UMP−AMPPCP monomer F, red; *P. furiosus* (pdb entry 2BMU) subunit A, pink; *E. coli* (pdb entry 2BNE) subunit B, light green; *S. pyogenes* (pdb entry 1Z9D) subunit A, light blue; *N. meningitides* (pdb entry 1YBD) subunit A, dark green; and *H. influenzae* (pdb entry 2A1F) subunit A, dark blue. See text for further explanation. The criterion for monomer selection was highest number of modeled residues and lowest alphabetic chain name. The alignment was performed using default parameters in the program LSQMAN (*37*), and the Figure was prepared using Molscript (*33*) and Raster 3D (*34*).

explains the preference for UMP over dUMP as substrate. This hydrogen bond can also not be formed with dTMP, which in addition is excluded from the active site by the 5-methyl group because the distance between the C5 atom of UMP and the side chain methyl of Thr119 is only 3.6 Å. There is a hydrogen bond between the backbone amide of Gly43 and an oxygen atom of the phosphate group of UMP. The same oxygen is in hydrogen-bonding distance from a β- and a γ-phosphoryl oxygen of ATP. Another of the phosphoryl oxygen atoms of UMP is also in hydrogen-bonding distance from a β- and a γ-phosphoryl oxygen of ATP as well as coordinating the Co^{2+} ion in the active site. The last phosphoryl oxygen of UMP is in hydrogen bond distance to the side chain hydroxyl group of Thr119 and to the main chain amide of Thr119 as well as to a water molecule (Figure 4A).

The enzyme showed very strong preference for ATP as the phosphoryl donor. This specificity is also well accounted for in the structure because the distances between the exocyclic N6 of ATP and the carbonyl of Tyr145 and between N1 of ATP and the backbone amide of the same amino acid are optimal for hydrogen bond formation (Figure 4A and Table 2). There is no space for an amino or a keto group at C2 of the purine ring of ATP, explaining the poor donor activity of GTP. A hydrogen bond is formed between 2′-OH of ATP and the side chain of Asp148. A similar bond cannot be formed with the poor substrate dATP. Aside from this, the enzyme binds ATP by a hydrogen bond to the 3′-OH group and by several interactions with the triphosphate chain (Figure 4A and Table 2).

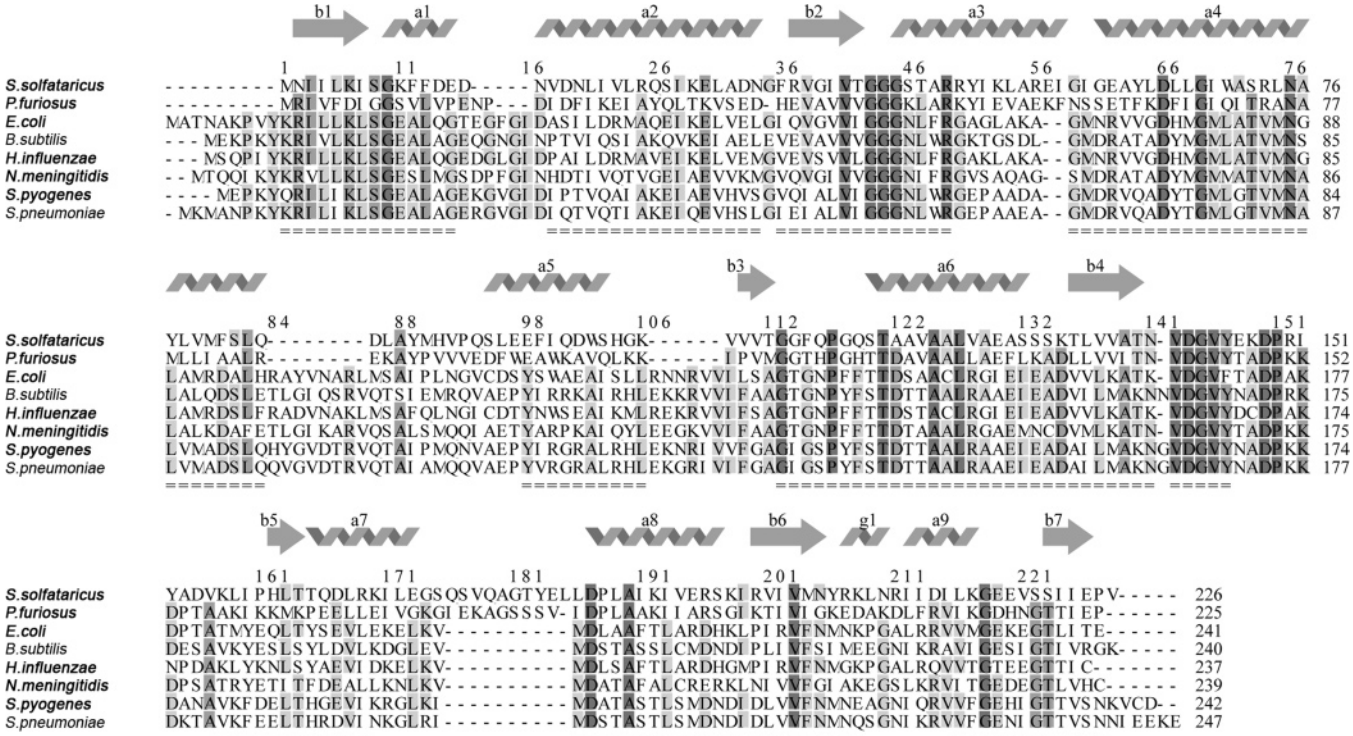

FIGURE 6: Structure based sequence alignment of hexameric uridylate kinases from *S. solfataricus* (subunit F), *P. furiosus* (pdb code 2BMU; subunit A), *E. coli* (pdb code 2BNE; subunit B), *H. influenzae* (pdb code 2A1F; subunit A), *N. meningitidis* (pdb code 1YBD; subunit A), *S. pyogenes* (pdb code 1Z9D; subunit A), *B. subtilis* (Swiss-Prot entry O31749), and *S. pneumoniae* (Swiss-Prot entry Q8DQ50). The secondary structure elements displayed above the sequences with twisted rods for α-helices and arrows for *β*-strands are according to *Ss*UMPK. The numbering of the secondary elements is the same as in Figure 3B. The residues that are identical in all amino acid sequences are shown in dark gray boxes, and the residues displaying seven identities are displayed in medium gray boxes, whereas the residues displaying five or six identities are displayed in light gray boxes. The sequences where the names of the organisms are shown in bold were structurally aligned using default parameters in the program LSQMAN (*37*). The aligned structures were used as input to the program STRUPRO, which assigns structurally equivalent residues on the basis of their Cα-positions, also with default parameters (*38*). The program package INDONESIA (http://alpha2.bmc.uu.se/~dennis/) was used for displaying the aligned structures and to include the sequences of *B. subtilis* and *S. pneumoniae* uridylate kinases. The signs (=) under the sequences represent amino acid residues with identical Cα-positions.

Structural Alignments. By superposition of a subunit of *Ss*UMPK with subunits of the UMPKs from *P. furiosus*, *E. coli*, *Streptococcus pyogenes*, *Neisseria meningitidis*, and *Haemophilus influenzae*, it is clear that *Ss*UMPK shares a fold with all of the other five enzymes but also that there are some interesting differences between the UMPK structures of bacterial and archaeal origin (Figure 5). The most conspicuous difference is that the loop enclosing the active site in the archaeal structures is missing from the four bacterial enzymes. Figure 6 shows a structure-based sequence alignment of the six enzymes together with the sequences of *B. subtilis* and *S. pneumoniae* UMPKs. The catalytic residues and residues involved in substrate binding are highly conserved, whereas there are large differences between the bacterial enzymes and the two archaeal enzymes in the region between the end of helix 4 and the beginning of *β*-strand 3. In the bacterial enzymes, this sequence forms a loop, which is absent in the archaeal enzymes (Figure 5). In addition, the archaeal enzymes differ considerably from the bacterial enzymes in the N-terminal region (Figure 6).

DISCUSSION

Reaction Mechanism. The kinetic investigation showed that the *Ss*UMPK catalyzes the phosphoryl group transfer by a sequential reaction mechanism, corresponding to the general picture for nucleoside monophosphate kinases (*1*). Both the kinetics and the ligand-binding studies are in agreement with a random equilibrium mechanism.

*Ss*UMPK contains only one metal ion in the active site. This ion is coordinated with phosphate groups of both ATP and UMP and is, together with Lys6, which binds to an oxygen of the *β*-phosphate of ATP, presumably involved in stabilizing and orienting the reactants on the enzyme (Figure 7). This arrangement is similar to the organization of the *E. coli* UMPK (*4*) but different from the enzyme from *P. furiosus*, which has two metal ions in the active site but lacks the positively charged lysine residue (*3*).

We do not believe that the coordination of Co^{2+} and Cd^{2+} ions to the enzyme on the noncrystallographic three fold axis is necessary for activity or for the correct folding of the enzyme because the overall fold is the same in the *Ss*UMPK−UTP structure, where no metal ions are bound at these sites.

The unusual electron density in the active site of subunit F of the *Ss*UMPK−UMP−AMPPCP structure is intriguing. The resolution limit (2.1 Å) does not allow an unambiguous interpretation of the density, but despite many attempts, we could only model it as a dinucleotide formed by covalent reaction between UMP and the *γ*-phosphoryl group of AMPPCP. It is generally believed that the phosphoryl transfer proceeds as a one-step concerted reaction involving a penta-coordinated phosphorus atom, the *γ*-phosphorus atom of ATP, with subsequent and concerted breakage of the *β*,*γ*-

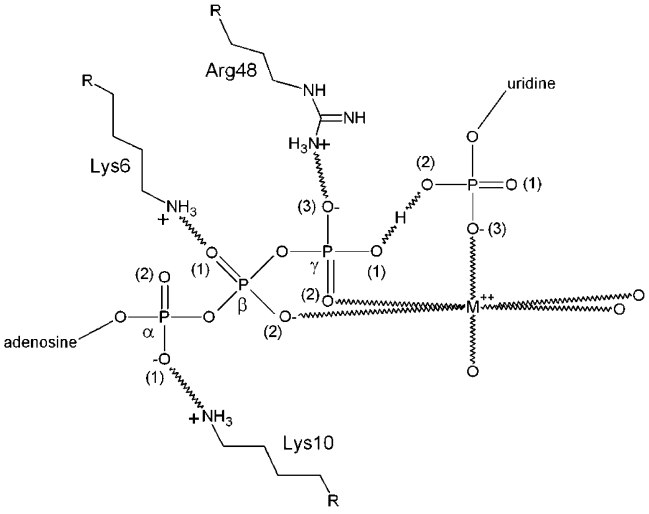

FIGURE 7: Schematic presentation of active site arrangement around the phosphoryl groups showing the two substrates, the metal ion, and positively charged amino acid side chains. The distance between phosphoryl oxygen(2) of UMP and γ-phosphoryl oxygen(1) of ATP is very short, indicative of the presence of a symmetric hydrogen bond. The metal ion is coordinated to both the β- and γ-phosphates of ATP and to the phosphoryl group of UMP. The pK_b value of 8.2 may reflect that UMP must be protonated at this position to be activated as a nucleophile, and the pK_a value of 5.8 may indicate that phosphoryl oxygen(3) of UMP needs a negative charge in order to carry out the attack on the electrophilic γ-phosphor atom of ATP. The divalent cation and the positively charged lysine residues presumably act to stabilize the highly negatively charged phosphoryl groups of both substrates, whereas Arg48 participates to correctly orient the substrates for catalysis.

phosphoanhydride bond and the formation of the α,β-phosphoanhydride bond in the UDP product (*5*, *6*). In the ATP analogue used for crystallization the β,γ-phosphorus atoms are linked by a methylene bridge that cannot be broken. Thus, it is not inconceivable that the reaction intermediate with a penta-cordinated phosphorus atom of AMPPCP, through an infrequent off-path reaction, might expel a water molecule when positioned in water-free surroundings in the active site of subunit F, where the loop, locked by crystal contacts, might prevent the exchange of ligands with the solvent.

The average distance between phosphoryl oxygen(2) of UMP and γ-phosphoryl oxygen(1) of ATP is only 2.3 Å in the other five subunits of the structure. The very short length of the hydrogen bond implies that the proton is equally shared between the two oxygen atoms (Figure 7). In turn, this suggests that phosphoryl oxygen(2) of UMP must be protonated to mediate catalysis, generating the pK_b value of 8.2 seen in the pH activity profile (Figure 1A). The formation of the very short hydrogen bond may act to increase the nucleophilic character of phosphoryl oxygen(3) of UMP and contribute to the development of a partial positive charge of the γ-phosphate of ATP and strengthen its ability as an electrophile. In addition, Arg48 may participate in the development of the partial positive charge on the γ-phosphate of ATP by formation of a hydrogen bond (average length 2.6 Å) with γ-phosphoryl oxygen(3). The pK_a value of 5.8 for the enzyme's productive interaction with the substrates is presumably due to the protonation of phosphoryl oxygen-(3) of UMP, which is expected to weaken the nucleophilic character of oxygen(3) (Figure 7).

In summary, substrate binding is followed by closure of the active site loop, resulting in movement of the catalytic residue Arg48 to a position where it stacks with Tyr181 and brings the γ-phosphate of ATP closer to the phosphoryl group of UMP, thus participating in stabilization of the transition state. After phosphoryl transfer has taken place, the loop opens and the products are released.

Regulation. The result that UTP acts as a competitive inhibitor against both the phosphoryl acceptor and donor agrees with the latest results from the study of the *E. coli* enzyme (*4*). However, we did not see any indication that the inhibitory effect of UTP was abolished by Mg^{2+}, as reported for the *E. coli* enzyme (*2*, *4*), because UTP bound to the enzyme in the presence of Mg^{2+} and because a Mg^{2+} was coordinated to UTP in the *Ss*UMPK−UTP structure. The finding that *Ss*UMPK lacks GTP stimulation was surprising because the enzyme shows large similarities to the investigated bacterial enzymes in sequence and 3D structure. Analysis of *E. coli* and *Salmonella* UMPK mutants has shown that a D201N (or D201G) mutation and also a D77N mutation lead to a loss of GTP stimulation (*39*, *40*). Both of these aspartate residues are conserved in *Ss*UMPK and correspond to Asp65 and Asp185 (Figure 6). Analysis of the surroundings of these aspartate residues revealed only minor structural differences between the enzymes, suggesting that other features are involved in the control of enzyme activity by GTP. The most striking differences between the two archaeal UMPKs on one side and the bacterial UMPKs on the other side are (a) that the loop that encloses the active site of the archaeal enzymes is absent from the bacterial enzymes, (b) that a loop between helix 4 and strand 3 in the bacterial enzymes lacks the enzymes of archaeal origin, and (c) that significant differences are found in the N-terminal regions of the proteins (Figures 5 and 6). Because *Ss*UMPK does not bind GTP, the binding site must either be missing or altered. It seems, therefore, likely that residues involved in the binding of GTP are found in the loops specific for the enzymes of bacterial origin. In a region near the N-terminal, the bacterial enzymes contain a GXXGXG motif, not present in the two archaeal enzymes (Figures 5 and 6). This motif is known to be involved in the binding of NTP in other proteins (*41*), and it is, therefore, likely to be involved in the binding of GTP. These extra residues in the bacterial enzymes form a loop preceding helix 2 in proximity to the corresponding loop in the neighboring subunit. Because UMPK from *P. furiosus* shares these differences from the bacterial enzymes with *Ss*UMPK, the lack GTP stimulation may be a general archaeal feature.

ACKNOWLEDGMENT

We thank Lise Schack for help with the purification of the enzyme, Flemming Hansen for assistance during diffraction data collection, and Martin Willemoës and Anders Kadziola for helpful discussions. We are also grateful to Vicente Rubio for sharing the structure of the *P. furiosus* UMPK prior to publication and for the beam time at the MAX-Laboratory, University of Lund, Sweden.

REFERENCES

1. Yan, H., and Tsai, M. D. (1999) Nucleoside monophosphate kinases: structure, mechanism, and substrate specificity, *Adv. Ezymol. Relat. Areas Mol. Biol. 73*, 103−134.

2. Serina, L., Blondin, C., Krin, E., Sismeiro, O., Danchin, A., Sakamoto, H., Gilles, A. M., and Bârzu, O. (1995) *Escherichia coli* UMP-kinase, a member of the aspartokinase family, is a hexamer regulated by guanine nucleotides and UTP, *Biochemistry 34*, 5066−5074.
3. Marco-Marin, C., Gil-Ortiz, F., and Rubio, V. (2005) The crystal structure of *Pyrococcus furiosus* UMP kinase provides insight into catalysis and regulation in microbial pyrimidine nucleotide biosynthesis, *J. Mol. Biol. 352*, 438−454.
4. Briozzo, P., Evrin, C., Meyer, P., Assairi, L., Joly, N., Bârzu, O., and Gilles, A. M. (2005) Structure of *Escherichia coli* UMP kinase differs from that of other nucleoside monophosphate kinases and sheds new light on enzyme regulation, *J. Biol. Chem. 280*, 25533−25540.
5. Schlichting, I., and Reinstein, J. (1997) Structures of active conformations of UMP kinase from *Dictyostelium discoideum* suggest phosphoryl transfer is associative, *Biochemistry 36*, 9290−9296.
6. Hutter, M. C., and Helms, V. (2000) Phosphoryl transfer by a concerted reaction mechanism in UMP/CMP-kinase, *Protein Sci. 9*, 2225−2231.
7. Fassy, F., Krebs, O., Lowinski, M., Ferrari, P., Winter, J., Collard-Dutilleul, V., and Salahbey Hocini, K. (2004) UMP kinase from *Streptococcus pneumoniae*: evidence for co-operative ATP binding and allosteric regulation, *Biochem. J. 384*, 619−627.
8. Gagyi, C., Bucurenci, N., Sirbu, O., Labesse, G., Ionescu, M., Ofiteru, A., Assairi, L., Landais, S., Danchin, A., Bârzu, O., and Gilles, A. M. (2003) UMP kinase from the Gram-positive bacterium *Bacillus subtilis* is strongly dependent on GTP for optimal activity, *Eur. J. Biochem. 270*, 3196−3204.
9. Randerath, K., and Randerath, E. (1967) Thin-layer seperation methods for nucleic acid derivatives, *Methods Enzymol. 12A*, 323−347.
10. Andersen, J. T., Poulsen, P., and Jensen, K. F. (1992) Attenuation in the *rph-pyrE* operon of *Escherichia coli* and processing of the dicistronic mRNA, *Eur. J. Biochem. 206*, 381−390.
11. Deuschle, U., Kammerer, W., Gentz, R., and Bujard, H. (1986) Promotors of *Escherichia coli*: a hierarchy of in vivo strength indicates alternate structures, *EMBO J. 5*, 2987−2994.
12. Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Habor Laboratory, Cold Spring Habor, NY.
13. She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C.-Y., Clausen, I. G., Curtis, B. A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P. M. K., Jong, I. H., Jeffries, A. C., Kozera, C. J., Medina, N., Peng, X., Thi-Ngoc, H. P., Redder, P., Schenk, M. E., Theriault, C., Tolstrup, N., Charlebois, R. L., Doolittle, W. F., Duguet, M., Gaasterland, T., Garrett, R. A., Ragan, M. A., Sensen, C. W., and Van der Oost, J. (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2, *Proc. Natl. Acad. Sci. U.S.A. 98*, 7835−7840.
14. Jensen, K. F., Arent, S., Larsen, S., and Schack, L. (2005) Allosteric properties of the GTP activated and CTP inhibited uracil phosphoribosyltransferase from the thermoacidophilic archaeon *Sulfolobus solfataricus*, *FEBS J. 272*, 1440−1453.
15. Martin, R. G., and Ames, B. N. (1961) A method for determining the sedimentation behavior of enzymes: application to protein mixtures, *J. Biol. Chem. 236*, 1372−1379.
16. Jensen, K. F., and Mygind, B. (1996) Different oligomeric states are involved in the allosteric behavior of uracil phosphoribosyltransferase from *Escherichia coli*, *Eur. J. Biochem. 240*, 637−645.
17. Poulsen, P., Bonekamp, F., and Jensen, K. F. (1984) Structure of the *Escherichia coli pyrE* operon and control of *pyrE* expression by a UTP modulated intercistronic attenuation, *EMBO J. 3*, 1783−1790.
18. Nielsen, F. S., Rowland, P., Larsen, S., and Jensen, K. F. (1996) Purification and characterization of dihydroorotate dehydrogenase A from *Lactococcus lactis*, crystallization and preliminary X-ray diffraction studies of the enzyme, *Protein Sci. 5*, 857−861.
19. Stura, E. A., Nemerow, G. R., and Wilson, I. A. (1992) Strategies in the crystallization of glycoproteins and protein complexes, *J. Cryst. Growth 122*, 273−285.
20. Jancarik, J., and Kim, S.-H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins, *J. Appl. Crystallogr. 24*, 409−411.
21. Cerenius, Y., Stahl, K., Svensson, L. A., Ursby, T., Oskarsson, A., Albertsson, J., and Liljas, A. (2000) The crystallography beamline I711 at MAX II, *J. Synchrotron Radiat. 4*, 203−208.
22. Ursby, T., Mammen, C. B., Cerenius, Y., Svensson, C., Sommarin, B., Fodje, M. N., Kvick, Å., Logan, D. T., Als-Nielsen, J., Thunnissen, M. M. G. M., Larsen, S., and Liljas, A. (2004) Macromolecular crystallography stations at Max-lab: the MAD station, *AIP Conf. Proc. 705*, 1241−1246.
23. Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data *Joint CCP4+ESR-EAMCB Newsletter on Protein Crystallography 26*.
24. Evans, P. R. (1997) CHOOCH - automatic analysis of flourescence scans and determination of optimal X-ray wavelengths for MAD and SAD, *CCP4 Newsletter on Protein Crystallography 33*, 22−24.
25. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *J. Appl. Crystallogr. 30*, 1022−1025.
26. Brunger, A. T., Adams, P. D., Clore, G. M., et al. (1998) Crystallography and NMR system (CNS): a new software system for macromolecular structure determination, *Acta Crystallogr., Sect. D 54*, 905−921.
27. Terwilliger, T. (2004) SOLVE and RESOLVE: automated structure solution, density modification and model building, *J. Synchrotron Radiat. 11*, 49−52.
28. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr, Sect. A 47*, 110−119.
29. Murshudov, G. N., Vagin, A. A., and Dobson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr, Sect. D 53*, 240−250.
30. Lamzin, V. S., Perrakis, A., and Wilson, K. S. (2001) The ARP/WARP Suite for Automated Construction and Refinement of Protein Models, *Crystallography of Biological Macromolecules*, Vol. F, Kluwer Academic Publishers, Amsterdam, The Nederlands.
31. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
32. Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program, *J. Mol. Graphics 8*, 52−56.
33. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
34. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic molecular graphics, *Methods Enzymol. 277*, 505−524.
35. Collaborative Computational Project. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr., Sect. D 50*, 760−763.
36. Potterton, E., Briggs, P., Turkenburg, M., and Dobson, E. (2003) A graphical user interface to the CCP4 program suite, *Acta Crystallogr., Sect. D 59*, 1131−1137.
37. Kleywegt, G. J., and Jones, T. A. (1997) Detecting folding motifs and similarities in protein structures, *Methods Enzymol. 277*, 525−545.
38. Kleywegt, G. J., and Jones, T. A. (1998) Databases in protein crystallography, *Acta Crystallogr., Sect. D 54*, 1119−1131.
39. Bucurenci, N., Serina, L., Zaharia, C., Landais, S., Danchin, A., and Barzü, O. (1998) Mutational analysis of UMP kinase from *Escherichia coli*, *J. Bacteriol. 180*, 473−477.
40. Sakamoto, H., Landais, S., Evrin, C., Laurent-Winter, C., Barzü, O., and Kelln, R. A. (2004) Structure-function relationships of UMP kinses from *pyrH* mutants of Gram-negative bacteria, *Microbiology 150*, 2153−2159.
41. Schulz, G. E. (1992) Binding of nucleotides by proteins, *Curr. Opin. Struct. Biol. 2*, 61−67.

BI0618159