Crystal Structure of Bacterial Inorganic Polyphosphate/ATP-glucomannokinase
INSIGHTS INTO KINASE EVOLUTION*

Received for publication, July 19, 2004, and in revised form, September 7, 2004
Published, JBC Papers in Press, September 17, 2004, DOI 10.1074/jbc.M408126200

Takako Mukai‡, Shigeyuki Kawai‡, Shigetarou Mori‡, Bunzo Mikami§, and Kousaku Murata¶

From the ¶Department of Basic and Applied Molecular Biotechnology, Division of Food and Biological Science, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan and the §Laboratory of Food Quality Design and Development, Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan

Inorganic polyphosphate (poly(P)) is a biological high energy compound presumed to be an ancient energy carrier preceding ATP. Several poly(P)-dependent kinases that use poly(P) as a phosphoryl donor are known to function in bacteria, but crystal structures of these kinases have not been solved. Here we present the crystal structure of bacterial poly(P)/ATP-glucomannokinase, belonging to Gram-positive bacterial glucokinase, complexed with 1 glucose molecule and 2 phosphate molecules at 1.8 Å resolution, being the first among poly(P)-dependent kinases and bacterial glucokinases. The poly(P)/ATP-glucomannokinase structure enabled us to understand the structural relationship of bacterial glucokinase to eucaryotic hexokinase and ADP-glucokinase, which has remained a matter of debate. These comparisons also enabled us to propose putative binding sites for phosphoryl groups for ATP and especially for poly(P) and to obtain insights into the evolution of kinase, particularly from primordial poly(P)-specific to ubiquitous ATP-specific proteins.

ATP participates universally in the transfer and storage of free energy in biological systems as the most common phosphoryl donor for kinases. Some Gram-positive bacteria such as Arthrobacter sp. strain KM and Mycobacterium tuberculosis possess inorganic polyphosphate (poly(P))1-dependent kinases that use poly(P) instead of ATP as the phosphoryl donor (1–4). Poly(P) is a biopolymer of several (up to thousands) orthophosphate residues linked by a high energy phosphoanhydride bond approximately equivalent to that of ATP (5). Poly(P)-dependent kinases consist of poly(P)/glucokinase (GK) (1), poly(P)/ATP-GK (2), poly(P)/ATP-glucomannokinase (GMK) (3), and poly(P)/ATP-NAD kinase (4); the first three catalyze phosphorylation of glucose, the first step of glycolysis, and the last catalyzes that of NAD, the last step in NADP biosynthesis. Note that poly(P)-GK is poly(P)-specific, although poly(P)/ATP-type enzymes use poly(P) and ATP. Although the primary structures of these poly(P)-dependent kinases were determined (1–4), the lack of a crystal structure has prevented us from clarifying the poly(P)-utilizing mechanism of these unique kinases. Poly(P)-dependent kinases have ATP-specific partners, and a knowledge of the crystal structure of poly(P)-dependent kinase may aid in understanding the structural relationship of poly(P)-dependent kinase to the ubiquitous ATP-specific kinase, i.e. to understand structural determinants enabling poly(P)/ATP-type and poly(P)-specific type kinases to use poly(P), ATP-specific kinase to reject poly(P) and poly(P)-specific kinase to reject ATP. Such understanding would lend insights into the evolutionary relationship of poly(P)-dependent kinase to ubiquitous ATP-specific kinase, where relationships are also of interest, since poly(P) could be formed and participate in ATP synthesis under ancient prebiotic conditions and serve as a possible ancient energy carrier preceding ATP (5).

Among poly(P)-dependent kinases, GMK, isolated from the Gram-positive bacterium Arthrobacter sp. strain KM, is a monomer with 30 kDa that phosphorylates glucose and mannose with preference for glucose through the use of poly(P) and ATP (3). The primary structure of GMK shows high homology with those of poly(P)- and poly(P)/ATP-GKs and other Gram-positive bacterial GKs but little with those of Gram-negative bacterial GKs and eucaryotic hexokinases (HKs), except for a few conserved motifs (3, 6). HK shows a broad specificity for hexose, whereas GK has high specificity for glucose (7, 8). Although crystal structures of eucaryotic HKs from yeast (9–13), human (14–19), rat (20), and parasite Schistosoma mansoni (20) have been solved, no crystal structure has been reported for Gram-positive or negative bacterial GKs. Hence, the crystal structural and evolutionary relationships between bacterial GK and eucaryotic HK have remained matters of debate (7, 8). Crystal structures of archaeal ADP-GKs from Thermococcus litoralis (21), Pyrococcus horikoshii (22), and Pyrococcus furiosus (23) definitively show that ADP-GK is distinguished from eucaryotic HK. ADP-GKs use ADP, whereas eucaryotic HK and bacterial GK, except for poly(P)-dependent GK, are regarded as using ATP but not poly(P).

Here we show the crystal structure of GMK complexed with one glucose and two phosphate (PO43−) molecules, the first among poly(P)-dependent kinases and bacterial GKs. Crystal structural analysis for GMK provides evidence for close crystal structural and evolutionary relationships between bacterial...
**Structure of Inorganic Polyphosphate/ATP-glucomannokinase**

### EXPERIMENTAL PROCEDURES

**Purification, Crystallization, and X-Ray Diffraction**—GMK was expressed from the GMK gene at NdeI-XhoI sites of pET-21b (Novagen) as a recombinant-enzyme in *Escherichia coli* BL21(DE3) cells (Novagen) in LB medium by inducing isopropyl-β-d-thiogalactopyranoside (0.4 mM) as described elsewhere (4) but at 20 °C for 24 h with sufficient aeration. The expressed recombinant GMK was purified using the same procedure as for native-GMK (3), except that, after ammonium sulfate precipitation, the supernatant was directly applied to a butyl Toyopearl (Tosoh) column and eluted using the gradient in potassium phosphate, pH 7.0) with 3 mM glucose from 10 to 600 mM. Active fractions contained the purified GMK and were homogeneous on SDS-PAGE (24). GMK was crystallized by mixing 3 l of protein solution (10 mg/ml GMK and 10 mM glucose in 10 mM potassium phosphate, pH 7.0) with 3 µl of mother liquor composed of 2.0M ammonium sulfate, 2% (v/v) polyethylene glycol 400, 0.1 M Hepes, pH 8.0.

Crystals were soaked in several heavy atom derivative solutions composed of 1 mM UO₂Ac₂, 10 mM AgNO₃, 10 mM GdCl₃, 1 mM HgCl₂, 10 mM K₂Pt(CN)₄, 2 mM TmCl₂, and 1 mM EuCl₃ for 20–60 min at 20 °C using CuKα radiation generated by a MAC Science length of 0.9 Å at the BL-41XU station of SPring-8 in Hyogo, Japan. The expressed recombinant GMK was purified using the same procedure as for native-GMK (3), except that, after ammonium sulfate precipitation, the supernatant was directly applied to a butyl Toyopearl (Tosoh) column and eluted using the gradient in ammonium sulfate from 30 to 0%. After dialysis, eluted GMK was purified by a hydroxylapatite (Nacalai Tesque) column using the gradient in potassium phosphate from 10 to 600 mM. Active fractions contained the purified GMK and were homogeneous on SDS-PAGE (24). GMK was crystallized by hanging-drop vapor diffusion method as described elsewhere (25). The solution of a crystallization drop was prepared on a siliconized coverslip by mixing 3 µl of protein solution (10 mg/ml GMK and 10 mM glucose in 10 mM potassium phosphate, pH 7.0) with 3 µl of mother liquor composed of 2.0M ammonium sulfate, 2% (v/v) polyethylene glycol 400, 0.1 M Hepes, pH 8.0. Crystals were soaked in several heavy atom derivative solutions composed of 1 mM UO₂Ac₂, 10 mM AgNO₃, 10 mM GdCl₃, 1 mM HgCl₂, 10 mM K₂Pt(CN)₄, 2 mM TmCl₂, and 1 mM EuCl₃ for 20–60 min at 20 °C. Derivative solutions were prepared in a mother liquor, but 0.1 M Tris-HCl, pH 7.0 (for UO₂Ac₂) and 0.1 M sodium acetate, pH 4.9 (for the remainder), were included instead of 0.1M Hepes, pH 8.0. Diffraction data for the native crystal of GMK up to 2.8 Å and derivative crystals at 10 mM potassium phosphate, pH 7.0) with 3 µl of mother liquor composed of 2.0M ammonium sulfate, 2% (v/v) polyethylene glycol 400, 0.1 M Hepes, pH 8.0.

**Structure Determination and Refinement**—The crystal structure of GMK was determined by multiple isomorphous replacement (MIR). The crystal structure of GMK was determined by multiple isomorphous replacement (MIR).

### RESULTS

**Structure Determination and Quality of Refined Model**—Recombinant-GMK purified to homogeneity showed the same molecular structure (30 kDa, monomer) and N-terminal sequence and almost the same kinetic parameters as those of native-GMK (3). GMK was crystallized in the presence of glucose, phosphate, and sulfate, and prismatic colorless crystals of the GMK grew to a maximum 0.1 mm in 2 weeks at 20 °C. The GMK crystal belongs to space group *P*2₁*₂*₂*₂* with unit cell dimensions of *a* = 65.21, *b* = 82.52, and *c* = 102.06 Å, and the solvent content was 49%, assuming 2 molecules in an asymmetric unit (25). The structure of GMK was determined by molecular dynamics using a CNS program package. Several rounds of restrained least squares refinement to a resolution of 1.8 Å followed by manual model building were conducted. Water molecules were incorporated when the difference in density was more than 3.0σ above the mean and the 2Fo − Fc map showed a density of more than 1.0σ. The final model was determined with an R-factor of 19.1% (free R-factor = 22.0%) at 50.0–1.8 Å resolution.

The stereochemical quality of the model was assessed using PROCHECK (31) and WHAT-CHECK (32) programs. Molecular models were prepared using MOLSCRIPT (33) and Raster3D (34) programs. Homology for the crystal structure was searched for in DALI (www.ebi.ac.uk/dali). Alignment of the primary structure was constructed by ClustalW (35). SCOP (36) was used to classify the crystal structure. Coordinates of crystal structures were taken from the Protein Data Bank (www.rcsb.org). Molecular models were superimposed by a fitting program implemented in TURBO-FRODO (Bio-Graphics). Coordinates and molecular topologies of pentapolyphosphate were generated using PRODRG (37).

| Compound  | UO₂Ac₂ | AgNO₃ | GdCl₃ | HgCl₂ | K₂Pt(CN)₄ | TmCl₂ | EuCl₃ |
|-----------|--------|-------|-------|-------|-----------|-------|-------|
| Soaking conditions | | | | | | | |
| Concentration (µl) | 1 | 10 | 10 | 1 | 10 | 2 | 1 |
| Soaking time (min) | 20 | 60 | 60 | 60 | 60 | 60 | 20 |
| Data used | Resolution limit (Å) | 3.1 | 2.7 | 2.8 | 3.2 | 3.0 | 3.1 | 3.5 |
| Phasing power | 0.62 | 0.88 | 0.75 | 0.42 | 0.84 | 0.81 | 0.65 |
| Rmerge | 0.735 | 0.637 | 0.757 | 0.764 | 0.714 | 0.713 | 0.757 |
| Rfactor | 0.118 | 0.108 | 0.091 | 0.125 | 0.093 | 0.105 | 0.081 |
| Number of sites | 2 | 4 | 2 | 2 | 8 | 2 | 2 |
| Average figure of merit | 0.489 | | | | | | |
| 0.825 (after solvent flattening) | | | | | | | |
| 0.883 (after averaging) | | | | | | | |

**TABLE I** Statistics for heavy atom derivatives of GMK

| Data collection | Wave length (Å) | 0.9 |
| Resolution limit (Å) | 50.0–1.8 (1.86–1.80) |
| Space group | *P*2₁*₂*₂*₂ |
| Unit cell dimensions (Å) | *a* = 65.21, *b* = 82.52, *c* = 102.06 |
| Molecules/asymmetric unit | 2 |
| Measured reflections | 253,953 (24,044) |
| Unique reflections | 50,787 (4,907) |
| Redundancy | 5.0 (4.9) |
| Completeness (%) | 97.7 (95.8) |
| Rmerge (%) | 5.9 (26.2) |
| Refinement statistics | |
| Final model | 506 residues, 373 water, 2 glucose, 4 phosphate |
| Resolution limit (Å) | 15.0–1.8 (1.86–1.80) |
| Used reflections | 49,826 (4,783) |
| Average R-factor (Å²) | 24.5 |
| Root mean square deviation | |
| Bond length (Å) | 0.0044 |
| Bond angle (deg) | 1.34 |
| R-factor (%) | 19.1 (21.2) |
| Free R-factor (%) | 22.0 (24.7) |

* Data in the highest resolution shells is given in parentheses.  
* This value is based on 10% of data excluded from refinement.
Fig. 1. Structures of GMK. a, stereo view of ribbon model of GMK with glucose (red), phosphate A (orange), and phosphate B (yellow). b, topology of secondary elements of GMK. The α-helix (H) and β-strand (SA and SB for N and C domains) are designated by numbers from the N-terminal. c, stereo view of a Cα backbone trace of superimposed structures of N-domain (residues 11–122) (green) and C-domain (residues 139–246) (yellow-green) of GMK.
MIR. Refinement statistics for MIR phasing and the final model are given in Tables I and II.

The electron densities of the main and side chains were generally well defined on the $2F_o - F_i$ map. The final $R$-factor was 19.1% (free $R$-factor = 22.0%) for 49,926 reflections with $F > 2.0 \sigma (F)$ at 50–1.8 Å of resolution. Interpretable electron density for GMK begins at residue 11 and ends at residue 263 for 1 molecule of GMK. N-terminal sequence analysis of recombinant-GMK showed that the first residue is missing, and the DNA sequence of the GMK gene exhibits a

**FIG. 2.** Substrate-binding sites. **a**, electron density map of glucose bound to GMK. Glucose oxygens are numbered. **b**, binding sites of GMK with glucose (red), phosphate A (orange), and phosphate B (yellow). Hydrogen bonds are indicated by dotted lines. **c**, binding sites of C-HKI with glucose (red) and ADP (orange) (PDB accession code 1DGK) (18). Interactions of residues with the β-phosphoryl group are not indicated since this group was not located at the proper site (18).
The C domain of H3/SB3:SB2:SB1:SB4:SB5/H4:H5:H6:H7:H8. N domain consisted of H1:H2/SA3:SA2:SA1:SA4:SA5/H9, and segments between SA5 and H3 and between SB5 and H9 (Fig. 1).

The refined model consisted of 506 residues, 373 water molecules, and one cis-peptide between Gln47 and Pro48 residues. The overall structure of GMK presented here is the first crystal structure of a eucaryotic HK and ADP-GK by comparing overall structures of bacterial HKs and poly(P)-dependent kinases, enabling us to understand the structural relationship of bacterial GK to eucaryotic HK and ADP-GK by comparing overall structures of enzymes and their substrate-binding sites. These comparisons also led us to propose putative binding sites of GMK for phosphoryl groups of ATP and for poly(P) and, finally, to obtain insights into the evolution of kinase as described below. Note

### Structure of Inorganic Polyphosphate

**Table III**

| Ligand atom | Protein atom (GMK-A/B) | Distance (GMK-A/B) | Element/domain |
|-------------|------------------------|-------------------|----------------|
| Glucose     |                        |                   |                |
| O1          | Glu180 OE1             | 2.6/2.6           | H4/C           |
| O2          | Glu180 OE2             | 3.1/3.2           | H4/C           |
|             | Glu180 OE1             | 2.7/2.7           | L-SB3:H4+/C    |
|             | His171 NE2             | 2.8/2.8           | L-SB3:H4/C     |
| O3          | Wat6 309/Wat 323       | 2.8/2.7           |                |
| O4          | Glu180 OE2             | 2.6/2.7           | L-SB3:H4/C     |
|             | Asn125 OD1             | 3.0/3.0           | SA5/N          |
| O5          | Asp155 OD1             | 3.1/3.0           | H3/C           |
| O6          | Arg42 NH2/Wat 466      | 2.9/2.9           |                |
| Phosphate A | Wat 384/Wat 443/Wat 354| 2.9/2.8/3.0       |                |
| O2          | Wat 471/Wat 466        | 2.8/2.6           |                |
| O3          | Thr151 N               | 3.0/3.0           | L-SB1:SB2/C    |
| Phosphate B | Wat 358/Wat 354        | 2.8/2.7           |                |
| O2          | Thr250 OG1             | 3.1/2.9           |                |
| O3          | Arg42 NH2/Wat 466      | 2.9/2.7           | SA3/N          |
| O4          | Thr250 OG1             | 3.0/2.8           | L-SB1:SB2/C    |
|             | Arg42 NE               | 2.8/2.8           | SA3/N          |

*a* Loop connecting SB3 and H4.

*b* Wat, water molecule.

Glucose- and Phosphates-binding Sites of GMK—Ligands (glucose and phosphates) were refined with fractional occupancies of 1.0. The average B-factor for glucose was 13.4 Å² and lower than the average for all atoms (24.6 Å²), but B-factors for phosphate-A (46.3 Å²), especially for phosphate-B (64.3 Å²), were higher than the average for all atoms. The electron density of the glucose O1 atom was observed for the β-anomer configuration (Fig. 2a). Selected contacts between GMK and ligands are listed in Table III. The O1 atom of phosphate B interacted with the Arg42 NH2 atom in GMK-A but with water 466 in GMK-B due to different directions in the side chain of Arg42 between GMK-A and GMK-B (data not shown). The binding sites of GMK with bound glucose, phosphate A, and phosphate B and their interactions with GMK are shown in Fig. 2b.

### DISCUSSION

The crystal structure of GMK presented here is the first among bacterial GKS and poly(P)-dependent kinases, enabling us to understand the structural relationship of bacterial GK to eucaryotic HK and ADP-GK by comparing overall structures of enzymes and their substrate-binding sites. These comparisons also led us to propose putative binding sites of GMK for phosphoryl groups of ATP and for poly(P) and, finally, to obtain insights into the evolution of kinase as described below. Note

**Figure 1**

A ribbon model of the overall crystal structure of GMK together with one bound glucose and two bound phosphates (phosphate A and B) in an asymmetric unit. Note that it was unclear if phosphate or sulfate was contained in GMK due to the similar molecular structures of the 2 anions and due to the uncertainty of the N domain (residues 123–263), 18% identity, and 48% similarity of the N domain on the C domain gave the root mean square deviation of 1.1 Å with 41 Ca within 2.0 Å (Fig. 1c). Taken together, we propose that the two domains stem from the duplication of one primordial domain.

From overall crystal structural features, GMK was classified into the “ASKHA (acetate and sugar kinase/hsc70/actin)” superfamily on classification by Cheek et al. (40). Accordingly, homology analysis using the crystal structure as a query on DALI showed that GMK resembled eucaryotic HKs such as human HK I (PDB accession code 1QHA) and yeast HK II (PDB accession code 2YHX).
that bacterial GK consists of Gram-positive and -negative bacterial GKs (6), where GMK belongs to the former (3), but no crystal structure of Gram-negative bacterial GK is known. Hence, we regard GMK as representative of bacterial GK.

Overall Structures of GMK, Eucaryotic HK, and ADP-GK—
The structural relationship of bacterial GK to eucaryotic HK and ADP-GK has not been understood. As mentioned above, however, the crystal structure of GMK appears to be related, at least to those of eucaryotic HK. We, therefore, attempted to compare GMK and eucaryotic HK in detail.

Eucaryotic HKs consist of isozymes (e.g. mammalian (human) HK I-VI) (7, 8). HK I (100 kDa) consists of N- and C-terminal halves (50 kDa), which have primary and tertiary structures, and substrate-binding sites, resembling each other and those of other 50 kDa eucaryotic HKs (human HK IV, also called GK, S. mansoni HK, and yeast HK PI and PII) (9–20). The C-terminal half of human HK I (C-HKI) has catalytic activity, but the N-terminal half is inactive (14–18). The tertiary structure of C-HKI (residues 525–913) is chosen as representative of eucaryotic HK for comparison with GMK. The tertiary structure of C-HKI consists of two domains, called large and small, between which glucose- and ADP-binding sites are formed, and shows “closed conformation” observed in eucaryotic HKs complexed with glucose (9–20) (Fig. 3a).

The secondary structure of GMK was homologous with that of C-HKI (Fig. 1b and Fig. 3b). The tertiary structure of GMK also appeared to be similar to that of C-HKI (Fig. 1a and Fig. 3a) and, accordingly, could be superimposed on that of C-HKI, giving the root mean square deviation of 1.35 Å with 132 Cα within 2.0 Å, especially the N domain of GMK on the small domain of C-HKI (Fig. 4). The primary structure of GMK corresponds to that of C-HKI (Fig. 5) when the two sequences are aligned based on superimposition of their substrate-binding sites (presented below), not as aligned by ClustalW. The primary structure of GMK is similar to that of C-HKI in this alignment (53% similarity, 16% identity over 258 residues), but several additional elements were found in C-HKI, suggesting that C-HKI evolved by acquiring elements from a GMK-like ancestor. The additional element corresponding to residues 753–818 in C-HKI was most apparent and corresponded approximately to the region previously designated a “flexible sub-

![Fig. 3. Structures of C-HKI.](a)
domain” (residues 766–811) (15). The flexible subdomain contains adenine binding residues (see below) (18), is found in other crystal structures of eucaryotic HKs (9–20), and appears to be involved in allosteric regulation of HK I by glucose 6-phosphate and phosphate through interaction with the N-terminal half (18, 41). Collectively, the close relationship of the overall tertiary structure of GMK to that of C-HKI, namely eucaryotic HK, was demonstrated, but the overall tertiary structure of GMK apparently differs from those of ADP-GKs (21–23) (Fig. 1a and Fig. 6).

Glucose-binding Sites of GMK, Eucaryotic HK, and ADP-GK—In C-HKI, residues and their atoms interacting with each of the oxygen atoms of glucose were comparable with those of GMK (17) (Table IV). The residues of GMK, Asn$^{122}$, Asp$^{123}$, Glu$^{168}$, and Glu$^{180}$ and their atoms corresponded well to those of C-HKI, Asn$^{656}$, Asp$^{657}$, Glu$^{708}$, and Glu$^{742}$ in tertiary structure (Fig. 2, b and c) and primary structure (Fig. 5). Among

![Small-domain](red) and Large-domain (blue)

**Fig. 4.** Stereo view of a Ca backbone trace of superimposed structures of GMK (light green) and C-HKI (light pink).

**Fig. 5.** Alignment of primary structures of GMK and C-HKI. Secondary elements are indicated based on Fig. 1b and Fig. 3b. Similarity was calculated by ClustalW using the primary structure of C-HKI, from which all additional elements in the structure of C-HKI were removed, and specified as dots and plus symbols (lower and higher similarities). Residues specified in the text are boxed in brown. Flexible subdomains (residues 766–811) in C-HKI are boxed in blue.

**Fig. 6.** Ribbon model of ADP-GK from *P. furiosus* with glucose (red) (PDB accession code 1UA4) (23).
these residues in C-HKI, the significance of Asp₁₂₃, Glu₁₈₀, and Glu⁷⁴₂ has been confirmed by site-directed mutagenesis (42, 43), and Asp₁₂₃ (corresponding to Asp¹²³ in GMK) is proposed as a catalytic base (42). The glucose-binding site of GMK, thus, resembled that of C-HKI, although glucose bound to GMK as a β-anomer (Fig. 2, a and b) and to C-HKI as an α-anomer (17) (Fig. 2c). The glucose-binding site of GMK showed no homology to that of *P. furiosus* ADP-GK (23) except for a similar spatial arrangement of Asp⁴⁴⁰ and Glu⁸₈ in ADP-GK to Asp¹²₃ and Glu¹₈₀ in GMK (data not shown). Collectively, the glucose-binding sites of eucaryotic HKs resembled that of GMK, but not that of ADP-GK, except for Asp⁴⁴⁰ (putative catalytic base) and Glu⁸₈ in ADP-GK (23).

Putative Binding Sites of GMK for Poly(P) and for Phosphoryl Groups of ATP—Phosphate A in GMK appeared to occupy a site near the α-phosphoryl group of ADP in a tertiary structural comparison (Fig. 2, b and c), whereas phosphate B is located slightly apart from phosphate A (Fig. 2b). We expected that phosphate A and B binding gave us information concerning the binding sites for poly(P) and ATP. Although the possibility remains that the phosphates may be sulfates, we assume that sulfate binding can also represent phosphate binding, since sulfate can bind to the phosphate-binding site of yeast HK PII (13). Two phosphoryl groups in pentapolyphosphate consisting of five phosphates could be superimposed on phosphate A and B such that the distance between the phosphorus atom in the terminal phosphoryl group of pentapolyphosphate and the oxygen atom of the 6-hydroxyl group of glucose was 3.0 Å (Fig. 7a), which is a suitable distance for phosphoryl transfer (18). Thus, we tentatively propose that phosphate A- and B-binding sites in GMK represent a pentapolyphosphate-binding site. Accordingly, poly(P), which consists of up to thousands of phosphates (5), could enter into the putative pentapolyphosphate-binding site, since this site is located at the side of the interdomain cleft (Fig. 7b).

Our pentapolyphosphate binding model also suggested that phosphate A is located at the binding site for the β-phosphoryl group of ATP (Fig. 7a). Note the β-phosphoryl group in C-HKI complexed with ADP is not located at the proper site (18) (Fig. 2c). In C-HKI the β-phosphoryl group is proposed to interact with N and OG₁ atoms of Thr⁶₈₀ (18). Thr⁶₈₀ corresponds to Thr¹₅₁ in GMK in primary and tertiary structures (Fig. 5 and Fig. 2, b and c), and N and OG₁ atoms of Thr¹₅₁ in GMK interacted with phosphate A (Table III). Furthermore, Asp⁵₃² in C-HKI is proposed to interact with β- and γ-phosphoryl groups via Mg²⁺ and water (18) and also corresponds to Asp¹₈ in GMK (Fig. 5 and Fig. 2, b and c). Taken together, these data suggest that phosphate A is located at the binding site for the β-phosphoryl group of ATP, and that the phosphoryl group-binding site of GMK is homologous with that of C-HKI. Our model of binding for phosphoryl groups of poly(P) and ATP

| Glucose atoms | GMK atoms | C-HKI atoms |
|---------------|-----------|-------------|
| O₁            | Glu¹₈₀ OE₁⁺ | Glu⁷⁴₂ OE₁⁺ |
| O₂            | Glu¹₈₀ OE₂   | Wat 457     |
|               | His⁷¹₁ NE₂   | Thr⁶₈₀ OG₁   |
|               | Glu¹₈₀ OE₁⁺ | Glu⁷⁴₂ OE₁⁺ |
|               | Wat 3₀₉      |             |
| O₃            | Asn⁵₂₂ OD₁⁻ | Asn⁶₈₃ OD₁⁻ |
|               | Glu¹₈₀ OE₂⁺ | Glu⁷⁴₂ OE₂⁺ |
| O₄            | Wat 3₀₉      | Asn⁶₈₃ OD₁⁻ |
|               | Asp⁵₃₂ OD₁⁻ | Asn⁶₈₃ OD₂⁻ |
| O₅            | Wat 3₁₄      | Lys⁶₂₁ NZ    |
| O₆            |             | Lys⁶₂₁ NZ    |
|               | Asp⁵₃₂ OD₂⁻ | Asp⁶₈₃ OD₂⁻ |
|               | Wat 3₅₈      | Wat 1₆₄     |

* Data are from Table III.
* Data are from Rosano et al. (17).
* Residues and atoms were homologous between GMK and C-HKI.

**Fig. 7.** Putative pentapolyphosphate-binding site in GMK. a, stereo view of binding sites of GMK with glucose (red), phosphate A (orange), phosphate B (yellow), and putatively bound pentapolyphosphate (yellow green). Interactions of residues with phosphate A and B are indicated by dotted lines. Two phosphoryl groups in pentapolyphosphate were superimposed on phosphate A and B with manual fitting. b, stereo view of a ribbon model of GMK with bound glucose and putative pentapolyphosphate as in a.
further indicates that phosphorylation sites for ATP and poly(P) are shared, in agreement with previous results for a competition plot for GMK (3).

Residues in GMK, Thr222 and Lys255, interacting with phosphate B (Fig. 2b) were conserved in the primary structures of C-HKI as Thr536 and Arg539 (Fig. 5) and in other eucaryotic HKs (13), Gram-negative and -positive bacterial GKs including poly(P)- and poly(P)/ATP-GKs (1, 3, 6), and even some proteins in theASKHA superfamily (40, 44). The sequence around residues highly conserved in these proteins is called a “phosphate-1” motif (44). The phosphate-1 motif is regarded as conserved in ATP-utilizing proteins in the ASKHA superfamily such as acetate and sugar kinases, heat shock protein, and actin, and plays a significant role in interaction with ATP (40, 44). Thr536, Arg539, and Asp532 in this motif of C-HKI are proposed to interact with phosphoryl groups of ATP via Mg2+ and water (18). Note that all crystal structure-solved proteins containing this motif except for GMK are ATP-specific, and we could find no crystal structures of these ATP-specific proteins showing the direct interaction of this motif with phosphate, suggesting that some structural elements in these ATP-specific proteins, not existing in GMK, possibly prevent this motif from binding to poly(P).

Where does adenine-ribose (adenosine) of ATP bind to GMK? Gly747 in C-HKI, interacting with ribose of ADP, appears to correspond slightly to Ala185 in GMK in primary structural alignment (Fig. 5). Thr784, Lys785, Ser788, and Thr863 (interacting with adenine of ADP) in C-HKI were not conserved in GMK, however (Fig. 5). Note that Thr784, Lys785, and Ser788 in C-HKI are located in the flexible subdomain not found in GMK. Although the primary structure of poly(P)-GK (1) was highly conserved with those of GMK and poly(P)/ATP-GK (2), poly(P)-GK could not use ATP. Identifying the correct adenosine-binding site of GMK would aid in understanding why poly(P)-GK rejects ATP. To specify adenine- and poly(P)-binding sites in GMK, a crystallographic study of GMK complexed with poly(P) or ATP is currently in progress.

Insights into Kinase Evolution—The GMK crystal structure presented here led us to assume evolutionary relationships of bacterial GK with eucaryotic HK and archaeal ADP-GK. Tertiary structural and evolutionary relationships between bacterial GK and eucaryotic HK are currently matters of debate (7, 8), whereas relationships between eucaryotic HK and archaeal ADP-GK are distinctly diverse (7, 8, 21–23). Our study showed the close structural relationship between bacterial GK and eucaryotic HK from the points of overall tertiary structure (Fig. 4), binding sites for substrates (glucose and phosphoryl group) (Fig. 2, b and c), and primary structure (Fig. 5), and hence, we propose that eucaryotic HK and bacterial GK diverged from a common poly(P)-specific primordial protein as described below. The structural relationship between bacterial GK and ADP-GK is apparently diverse (Fig. 1a and Fig. 6), and hence, the evolutionary relationship would also be diverse, agreeing with the accepted idea that the structural and evolutionary relationship between eucaryotic HK and archaeal ADP-GK is distinctly diverse (7, 8, 21–23). It appeared that the common ancestor of bacterial GK and eucaryotic HK evolved into eucaryotic HK by acquiring several elements, including a flexible subdomain (Fig. 5), to satisfy physiological criteria required in eucaryotes (organisms containing mitochondria), whereas ancestors evolved into bacterial GK without such acquisitions. Eucaryotic HK appears to evolve to be suitable for organisms containing mitochondria, since (i) eucaryotic amitochondriate protist (without mitochondria), Trichomonas vaginalis, carries only Gram-negative bacterial GK (6), not eucaryotic HK, and (ii) human HK I interacts with mitochondria, permitting direct exchange of adenine nucleotides between the mitochondrial matrix and the HK active site (45, 46).

Crystal structural and evolutionary relationships of poly(P)-dependent kinase with ubiquitous ATP-specific kinase are also interesting and significant but remain to be clarified due to the lack of crystal structural information on poly(P)-dependent kinase. Here we present the crystal structure of this type of kinase and propose the putative poly(P)-binding site of GMK, which contains the phosphate-1 motif well conserved in ATP-specific proteins. Taking into account that (i) adenine-interacting residues in C-HKI are located in the flexible subdomain not found in GMK (Fig. 5), (ii) our proposition that eucaryotic HK acquired the flexible subdomain, interacting with adenine, during evolution, and finally (iii), GMK takes a fold similar to those of ATP-specific ASKHA proteins containing the phosphate-1 motif, we propose that ATP-specific proteins containing the phosphate-1 motif may have evolved from a primordial poly(P)-specific GK-like protein and may acquire the ability to use ATP and lose that to use poly(P) during evolution, which agrees with the suggestion that poly(P) is an energy carrier predicted by ATP (5).
Structure of Inorganic Polyphosphate/ATP-glucomannokinase

4673–4680

36. Le Conte, L., Ailey, B., Hubbard, T. J., Brenner, S. E., Murzin, A., and Chothia, C. (2000) *Nucleic Acids Res.* 28, 257–259

37. Schuttelkopf, A. W., and van Aalten, D. M. (2004) *Acta Crystallogr. Sect. D* 60, 1355–1363

38. Luzzati, V (1952) *Acta Crystallogr.* 5, 802–810

39. Ramachandran, G. N., and Sasisekharan, V. (1968) *Adv. Protein Chem.* 23, 283–437

40. Cheek, S., Zhang, H., and Grishin, N. V. (2002) *J. Mol. Biol.* 320, 855–881

41. Fang, T. Y., Alechina, O., Aleshin, A. E., Fromm, H. J., and Honzatko, R. B. (1998) *J. Biol. Chem.* 273, 19548–19553

42. Arora, K. K., Filburn, C. R., and Pedersen, P. L. (1991) *J. Biol. Chem.* 266, 5359–5362

43. Zeng, C., Aleshin, A. E., Guanjin, C., Honzatko, R. B., and Fromm, H. J. (1998) *J. Biol. Chem.* 273, 700–704

44. Bork, P., Sander, C., and Valencia, A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 7290–7294

45. Polakis, P. G., and Wilson, J. E. (1985) *Arch. Biochem. Biophys.* 236, 328–337

46. Cesar, M. C., and Wilson, J. E. (1998) *Arch. Biochem. Biophys.* 350, 109–117