Crystal Structure of YegS, a Homologue to the Mammalian Diacylglycerol Kinases, Reveals a Novel Regulatory Metal Binding Site*§

H. M. Amin Bakali‡§, Maria Dolores Herman‡, Kenneth A. Johnson‡, Amélie A. Kelly‡, Åke Wieslander§, B. Martin Hallberg†, and Per Nordlund‡

From the Division of Biophysics, Department of Medical Biochemistry and Biophysics, Karolinska Institute, SE-171 77 Stockholm, Sweden and Department of Biochemistry and Biophysics, Stockholm University, S-106 91 Stockholm, Sweden

The human lipid kinase family controls cell proliferation, differentiation, and tumorigenesis and includes diacylglycerol kinases, sphingosine kinases, and ceramide kinases. YegS is an Escherichia coli protein with significant sequence homology to the catalytic domain of the human lipid kinases. We have solved the crystal structure of YegS and shown that it is a lipid kinase with phosphatidylglycerol kinase activity. The crystal structure reveals a two-domain protein with significant structural similarity to a family of NAD kinases. The active site is located in the interdomain cleft formed by four conserved sequence motifs. Surprisingly, the structure reveals a novel metal binding site composed of residues conserved in most lipid kinases.

YegS is a cytosolic protein in Escherichia coli with significant sequence homology to mammalian lipid kinases such as diacylglycerol, ceramide, and sphingosine kinases (1, 2). YegS is annotated as a putative diacylglycerol kinase (DGK), 2 because it shares sequence similarity to the catalytic domain of the first reported DGKα (3). Diacylglycerol (DAG) is a lipid second messenger generated as a response to receptor activation and has important functions in the regulation of cell differentiation, proliferation, and carcinogenesis (4) Eukaryotic DGKs constitute a protein family that catalyzes the conversion of DAG to phosphatidic acid. A major cellular function of DAG is to activate protein kinase C (5); DGKs prevent this activity by converting DAG to phosphatidic acid. Related mammalian sphingosine and ceramide kinases, as well as the recently identified multiplipid kinase (MulK) (6) with a broader substrate specificity, are also members of the lipid kinase family. Both phosphorylated and nonphosphorylated sphingosine and ceramide adducts are important second messengers in cellular stress and apoptosis, but in particular, sphingosine-1-phosphate (7) and ceramide-1-phosphate (8) have attracted much attention recently. The sphingosine-1-phosphate-mediated pathways are proliferative, pro-inflammatory, and anti-apoptotic, and thus sphingosine kinase has been described as an oncogene and a potential neoplastic drug target. Several of the other lipid kinases have also been proposed as potential targets for therapeutic interventions (9–11).

Mammalian DGKs represent a large group of lipid kinases with ten known isoforms (12). These isoforms are subdivided into five groups, depending on the composition of accessory domains. In contrast, YegS does not contain any accessory domains. Mammalian DAG kinases have been shown to be regulated by second messengers, and Ca2+ has been demonstrated to be a potent activator of the DGKα isoform. The MulK broad specificity lipid kinase is also activated by Ca2+, but only at low Mg2+ concentrations (0.5 mM) (6), implying a complex dependence on divalent metals. The mechanism for the calcium-dependent effects of DGKs has been assumed to be mediated by the EF-hand domain, where Ca2+ binding induces membrane localization through the C1 domains and thereby enables access to the lipid substrates (13). In contrast, Ca2+ effects have also been shown to be mediated at the catalytic subunit of DGKs, where Ca2+ lowered the Km values for ATP binding (13). Furthermore, mutagenesis studies on the catalytic subunit of porcine DGKα have identified two conserved aspartic acid residues that are implied in Ca2+ regulation through an unknown mechanism (2). A structural basis for the regulation and action of the lipid kinases is lacking, as there is no three-dimensional structure available for any family member.

The function of YegS has not previously been established, but the homology to other lipid kinases implies that YegS might be involved in lipid phosphorylation. E. coli does contain a well characterized membrane-bound DGK (14). This DGK is a homotrimeric integral membrane protein in which each subunit has three predicted transmembrane helices (15). However, YegS is a soluble cytosolic protein. Numerous prokaryotic YegS homologues are available, but these are all from uncharacterized hypothetical proteins. Close homologues are found in the human pathogens Shigella flexneri, Salmonella typhimurium LT2, Yersinia pestis, and Mycobacterium tuberculosis. Recently, the latter gene, Rv2252, was shown to code for a DAG kinase.
that catalyzes the formation of phosphatidic acid but that is not essential for M. tuberculosis survival under laboratory conditions (16). In the present study we have shown that YegS is a lipid kinase with phosphatidylglycerol kinase activity, which has been previously undescribed. The structure of YegS reveals a two-domain protein with significant similarities to the fold and active site of NAD kinases. The structure also reveals a novel metal binding site coordinated by residues conserved in most lipid kinases. The mechanistic implications of the YegS structure for the human lipid kinases are discussed as well as the potential biological role for YegS homologues in prokaryotic lipid metabolism.

**EXPERIMENTAL PROCEDURES**

**Purification**—Two expression constructs of YegS were used in this study. The cloning and vector details for construct 1 were described earlier (17). Construct 2 is similar to construct 1 but lacks the non-cleavable C-terminal His tag of construct 1. Construct 2 was generated due to problems in reproducing the crystals from construct 1. Construct 1 protein was expressed and purified as described earlier (17). Construct 2 was purified from selenomethionine-labeled protein (18) using the same purification protocol as for construct 1, with the addition of a His tag cleavage step. The affinity-purified protein was exchanged into a buffer with detergent, and sonicated for 5 min in a water bath. After preincubating the assay mixture for 5 min at 25 °C, ATP was added and the reaction was stopped after 30 min at 120 °C. The plates were developed in acetone/toluene/water (91:30:8), and the radioactive lipids were visualized by autoradiography. For direct quantification, the amount of phosphate incorporated into the lipid phase was determined by scintillation counting. For assays employing Mg$^{2+}$ or Ca$^{2+}$, both EGTA and 5 mM MgCl$_2$ were omitted. To remove contaminating metals before the calcium effect measurements, the purified protein sample was incubated in 1 mM EGTA solution for 30 min at 4 °C. To remove EGTA, the protein solution was passed sequentially over two Sephadex G-25 PD-10 columns (GE Healthcare) that had been equilibrated in 40 mM Bis-Tris, pH 7.5. The protein samples as well as the other assay buffers/solutions were deionized using a Chelex 100 column (Bio-Rad) (21). Radioactive [$\gamma^{33}$P]ATP (9000 TBq/mmol) was purchased from PerkinElmer. Lipids were purchased from Avanti Polar Lipids and Sigma. Salts (MgCl$_2$, CaCl$_2$, and NaCl) and detergents were purchased from Sigma. Reference lipids on the developed TLC plates were visualized with iodine vapor.

**Mass Spectrometry**—For mass spectrometric analysis of a short chain substrate, 20 mM water-soluble phosphatidylglycerol (1,2-dihecanoyl-sn-glycero-3-[phospho-rac-(1-glycerol)] (sodium salt) (Avanti Polar Lipids) was incubated at room temperature for 1 h with 2 μg of enzyme in a buffer containing 40 mM Bis-Tris, pH 7.5, 1 mM ATP, 1 mM MgCl$_2$, 1 mM CaCl$_2$. Mass spectra were acquired in negative mode on an electrospray ionization quadrupole time-of-flight mass spectrometer (Q-ToF; Micromass). Reference sample had the same composition, but no protein was added. Tandem mass spectrometry was used to confirm the identity of the product. The parent ion (m/z 521) was subjected to collision-induced dissociation using argon as the collision gas. The tandem mass spectrometry spectra revealed peaks at m/z 78.93, 96.99, and 115.12 corresponding to the fragments PO$_3^-$, PO$_{4}^{2-}$, and $C_6H_11O_2$, respectively.

**Crystallization and Data Collection**—Crystals in the space group P21 were obtained as described previously (17). Selenomethionine protein from construct 2 crystals was used for phasing (crystal 1). Crystals were produced by mixing 2 μl of protein (10 mg/ml) with 1 μl of mother liquid containing 20% polyethylene glycol 3350, 0.1 M Li$_2$SO$_4$. Crystal 3 was obtained using the same ratio of chloroform/methanol (1:1), placed in a disposable glass tube, dried under a stream of nitrogen vapor, resuspended in the assay buffer with detergent, and sonicated for 5 min in a water bath sonicator. After preincubating the assay mixture for 5 min at 25 °C, ATP was added and the reaction was stopped after 30 min by adding 750 μl of chloroform/methanol (1:2) containing 1% HCl. Extraction and separation of phospholipids were performed as follows. 1 ml of chloroform/methanol (1:1) and 500 μl of 1 M KCl plus 0.2 M H$_2$PO$_4$ were added, and the mixture was mixed thoroughly on a vortexer; samples were centrifuged at 2000 rpm for 5 min. The lower density phase (lipids) was transferred to a new glass reaction tube and dried under a stream of nitrogen gas. 50 μl of chloroform/methanol (2:1) were added to each tube, and the samples were applied to TLC silica plates (TLC plates 20 × 20 cm silica gel 60; Merck) that had been treated with 0.15 M ammonium sulfate and activated for 2 h and 30 min at 120 °C. The plates were developed in acetone/toluene/water (91:30:8), and the radioactive lipids were visualized by autoradiography. For direct quantification, the amount of phosphate incorporated into the lipid phase was determined by scintillation counting. For assays employing Mg$^{2+}$ or Ca$^{2+}$, both EGTA and 5 mM MgCl$_2$ were omitted. To remove contaminating metals before the calcium effect measurements, the purified protein sample was incubated in 1 mM EGTA solution for 30 min at 4 °C. To remove EGTA, the protein solution was passed sequentially over two Sephadex G-25 PD-10 columns (GE Healthcare) that had been equilibrated in 40 mM Bis-Tris, pH 7.5. The protein samples as well as the other assay buffers/solutions were deionized using a Chelex 100 column (Bio-Rad) (21). Radioactive [$\gamma^{33}$P]ATP (9000 TBq/mmol) was purchased from PerkinElmer. Lipids were purchased from Avanti Polar Lipids and Sigma. Salts (MgCl$_2$, CaCl$_2$, and NaCl) and detergents were purchased from Sigma. Reference lipids on the developed TLC plates were visualized with iodine vapor.

**Crystallization and Data Collection**—Crystals in the space group P21 were obtained as described previously (17). Selenomethionine protein from construct 2 crystals was used for phasing (crystal 1). Crystals were produced by mixing 2 μl of protein (10 mg/ml) with 1 μl of mother liquid containing 20% polyethylene glycol 3350, 0.1 M Li$_2$SO$_4$, 0.2 M Li$_2$SO$_4$. Crystal 3 was obtained using the same ratio of chloroform/methanol (1:1), placed in a disposable glass tube, dried under a stream of nitrogen vapor, resuspended in the assay buffer with detergent, and sonicated for 5 min in a water bath sonicator. After preincubating the assay mixture for 5 min at 25 °C, ATP was added and the reaction was stopped after 30 min by adding 750 μl of chloroform/methanol (1:2) containing 1% HCl. Extraction and separation of phospholipids were performed as follows. 1 ml of chloroform/methanol (1:1) and 500 μl of 1 M KCl plus 0.2 M H$_2$PO$_4$ were added, and the mixture was mixed thoroughly on a vortexer; samples were centrifuged at 2000 rpm for 5 min. The lower density phase (lipids) was transferred to a new glass reaction tube and dried under a stream of nitrogen gas. 50 μl of chloroform/methanol (2:1) were added to each tube, and the samples were applied to TLC silica plates (TLC plates 20 × 20 cm silica gel 60; Merck) that had been treated with 0.15 M ammonium sulfate and activated for 2 h and 30 min at 120 °C. The plates were developed in acetone/toluene/water (91:30:8), and the radioactive lipids were visualized by autoradiography. For direct quantification, the amount of phosphate incorporated into the lipid phase was determined by scintillation counting. For assays employing Mg$^{2+}$ or Ca$^{2+}$, both EGTA and 5 mM MgCl$_2$ were omitted. To remove contaminating metals before the calcium effect measurements, the purified protein sample was incubated in 1 mM EGTA solution for 30 min at 4 °C. To remove EGTA, the protein solution was passed sequentially over two Sephadex G-25 PD-10 columns (GE Healthcare) that had been equilibrated in 40 mM Bis-Tris, pH 7.5. The protein samples as well as the other assay buffers/solutions were deionized using a Chelex 100 column (Bio-Rad) (21). Radioactive [$\gamma^{33}$P]ATP (9000 TBq/mmol) was purchased from PerkinElmer. Lipids were purchased from Avanti Polar Lipids and Sigma. Salts (MgCl$_2$, CaCl$_2$, and NaCl) and detergents were purchased from Sigma. Reference lipids on the developed TLC plates were visualized with iodine vapor.
Crystal Structure and a Novel Function of YegS

TABLE 1
Data collection, phasing, and refinement statistics

Two crystals were used for YegS structure determination, Crystal 1 for phasing and Crystal 2 for refinement.

Data from crystals of construct 1 were processed using MOSFLM (26) and scaled using SCALA (27). Details of the data collection are shown in Table 1.

The low resolution model in the first data set was built into the initial 3.1 Å resolution electron density map using XtalView (28) and refined using REFMAC5 (29). The coordinates were transferred to the second data set that extends to 1.9 Å resolution by molecular replacement using AMoRe (30) and refined using REFMAC5 with subsequent runs of manual fitting using XtalView. Simulated annealing using CNS (31) was performed to help with the positioning of less ordered loops. When the data were very anisotropic a “smoothed” σ cutoff procedure was used for filtering.3 Data for refinement were selected by calculating an average σ on F for each reflection and all neighboring reflections. Thus, up to 27 reflections contributed to the average σ cut-off value for each reflection. Then, several “smoothed” σ cutoffs were tried to select data for refinement. The final selected smoothed σ cutoff was chosen because it produced a data set that refined well. This averaged σ cut-off was set at 4σ, resulting in a refinement data set 98% complete to 2.7 Å, 73% complete overall to 1.9 Å, and 37.8% complete in the final shell to 1.9 Å. We did in parallel try refinement using a softened or σ cutoff applied as described under “Experimental Procedures.”

RESULTS

YegS has been classified as a putative DGK based on the primary sequence analysis. To address the question of whether YegS is a bona fide lipid kinase, a functional activity assay was performed. TLC revealed no phosphatidic acid accumulation when using DAG as substrate (Fig. 1, lane 1) but detected a strong incorporation of γ-32P from ATP into a major band when using E. coli total lipid extracts or phosphatidylglycerol (PG) (Fig. 1A, lanes 2 and 3 and Fig. 1B). Analysis of the lipid product by mass spectrometry, after using a short chain analogue of PG (C 6:0) as acceptor substrate, confirmed that YegS acts directly on a PG-type head group, yielding the phosphatidylglycerol phosphate (PGP) product (see supplemental Fig. S1). From the substrate-activity relationships in Fig. 1B, an apparent Kₘ for PG of 0.70 mM, and a Vₘₐₓ of 29.7 nmol/min/mg, could be calculated. Activity was obtained from pH 5 to 9. The Kₘ value could be compared with a molar fraction of PG in the mixed PG/CHAPS aggregates of a few percent, substantially below the 10–15 mole percent PG in the E. coli inner membrane in vivo. To the best of our knowledge a phosphatidylglycerol phosphorylation reaction has not been reported earlier in E. coli or in any other organism.

The structure of YegS was initially solved to 2.9 Å resolution in a tetragonal crystal form (P4₁2₁2) by single wavelength anomalous dispersion techniques using a selenomethionine-substituted variant of YegS. An initial model was subsequently used to solve the structure of the native protein crystallized in a monoclinic (P2₁) crystal form that diffracted to 1.9 Å resolution and contained two protomers in the asymmetric unit. An anisotropic smoothed σ cutoff procedure was used (see “Experimental Procedures”) when this helped in promoting the convergence of the refinement. Although the average B-values remain

3 K. A. Johnson, unpublished information.
Crystal Structure and a Novel Function of YegS

YegS shows the closest structural similarity to the fold recently described in NAD kinases (34–36). Superposition of the structure of each of the N- and C-terminal domains of YegS with NAD kinase shows a root mean square deviation of 2.6 Å (109 residues to 1Z0S) and 2.7 Å (115 residues to 2AN1), respectively. The superposition of the entire proteins gives a significantly higher root mean square distance because the two domains display slightly different orientation to each other in the two proteins (Fig. 2B). The N-terminal domain of YegS also shows good structural superposition with the N-terminal domain of PFK (37). A signature GGDG motif of the PFK family is also present in YegS (GGDG in YegS), as well as in NAD kinase, and superimposes well in the structures. This motif has been implicated by structural and mutagenesis studies to be essential for the phosphoryl transfer reaction in the PFK family (38) and more recently in the NAD kinase structure from Archaeoglobus fulgidus (36). However, a detailed understanding of how this motif promotes the reaction has been elusive. A DALI (39) search using the N-terminal domain of YegS found a large number of significant hits to Rossman fold-containing proteins involved in nucleotide and dinucleotide interactions. Searching with the C-terminal domain found, in addition to NAD kinase, only the structure of Vp4, a sialic acid binding domain from rhesus rotavirus (40). However, Vp4 does not contain a Rossman fold nucleotide binding domain. Therefore, YegS and NAD kinases appear to represent a unique subgroup of the Rossman fold kinases and are structurally distinct from the PFK family. The similarity of the C-terminal domain to that of NAD kinases is especially apparent in the core β-sandwich of this domain (Fig. 2B). In contrast, large deviations in structure between YegS and NAD kinases are noticeable in the region corresponding to the NAD binding site. YegS has a loop protruding into the region where the NAD/NADP binding site is found in NAD kinase (residues 203–204). Furthermore, the secondary structural elements α6, α7, β7, and β8 of YegS do not have corresponding elements in NAD kinase, and this region is likely to constitute the substrate binding region of the lipid kinases.

The structure of YegS underlines the importance of four signature sequence motifs constituting the active site crevice, three in the N-terminal domain and one in the C-terminal domain highly conserved in both bacterial and eukaryotic lipid kinases (2) (Fig. 2C). The motifs in the N-terminal domain are NXKS (residues 12–15 and X represents any amino acid), GGDG (residues 65–69 also present in NAD kinases and PFK) and LPLGTXDXA (residues 91–100), which we label the T-motif. All three motifs are found to make interactions to the bound pyrophosphate in the ADP co-crystallized structure (Fig. 3). The importance of the GGDG motif for the mammalian lipid kinases has been demonstrated by several mutagenesis studies (2, 41–43). In the YegS structure, the GGDG motif is found at the beginning of the third α-helix, and the amino groups of the helix termini (Gly-66 and Gly-68) directly coor-

relatively high, all residues are well defined in protomer A, except for the loop containing residues 148–157, whereas in the B protomer residues 148–161 and residues 228–245 were poorly defined.

The structure of YegS reveals a two-domain protein with the active site crevice found between the two domains (Fig. 2A). The N-terminal domain is formed by residues 5–124 and 286–301 and comprises five α-helices (α1–α5) and six β-strands (β1–β5 and β19). This domain resembles the Rossmann fold nucleotide binding motif (33). The C-terminal domain (residues 125–285) contains 13 β-strands (β6–β18) and two α-helices (α6 and α7). Data were also collected from crystals co-crystallized with ADP and the substrate analogue glycerol-3-phosphate to 2.5 Å resolution. In the resulting density from these crystals, only the ligand density for the pyrophosphate moiety and some minor residual density is observed, and neither the ribose, adenosine ring of the ADP, nor the glycerol-3-phosphate is visible in the electron density. Although the density clearly reveals a pyrophosphate moiety, the high B-values in the refinement (~100) might indicate that it is not fully occupied. The pyrophosphate moiety is found in close proximity to the crevice between the two domains that constitutes the active site.

YegS shows the closest structural similarity to the fold recently described in NAD kinases (34–36). Superposition of the structure of each of the N- and C-terminal domains of YegS with NAD kinase shows a root mean square deviation of 2.6 Å (109 residues to 1Z0S) and 2.7 Å (115 residues to 2AN1), respectively. The superposition of the entire proteins gives a significantly higher root mean square distance because the two domains display slightly different orientation to each other in the two proteins (Fig. 2B). The N-terminal domain of YegS also shows good structural superposition with the N-terminal domain of PFK (37). A signature GGDG motif of the PFK family is also present in YegS (GGDG in YegS), as well as in NAD kinase, and superimposes well in the structures. This motif has been implicated by structural and mutagenesis studies to be essential for the phosphoryl transfer reaction in the PFK family (38) and more recently in the NAD kinase structure from Archaeoglobus fulgidus (36). However, a detailed understanding of how this motif promotes the reaction has been elusive. A DALI (39) search using the N-terminal domain of YegS found a large number of significant hits to Rossman fold-containing proteins involved in nucleotide and dinucleotide interactions. Searching with the C-terminal domain found, in addition to NAD kinase, only the structure of Vp4, a sialic acid binding domain from rhesus rotavirus (40). However, Vp4 does not contain a Rossman fold nucleotide binding domain. Therefore, YegS and NAD kinases appear to represent a unique subgroup of the Rossman fold kinases and are structurally distinct from the PFK family. The similarity of the C-terminal domain to that of NAD kinases is especially apparent in the core β-sandwich of this domain (Fig. 2B). In contrast, large deviations in structure between YegS and NAD kinases are noticeable in the region corresponding to the NAD binding site. YegS has a loop protruding into the region where the NAD/NADP binding site is found in NAD kinase (residues 203–204). Furthermore, the secondary structural elements α6, α7, β7, and β8 of YegS do not have corresponding elements in NAD kinase, and this region is likely to constitute the substrate binding region of the lipid kinases.

The structure of YegS underlines the importance of four signature sequence motifs constituting the active site crevice, three in the N-terminal domain and one in the C-terminal domain highly conserved in both bacterial and eukaryotic lipid kinases (2) (Fig. 2C). The motifs in the N-terminal domain are NXKS (residues 12–15 and X represents any amino acid), GGDG (residues 65–69 also present in NAD kinases and PFK) and LPLGTXDXA (residues 91–100), which we label the T-motif. All three motifs are found to make interactions to the bound pyrophosphate in the ADP co-crystallized structure (Fig. 3). The importance of the GGDG motif for the mammalian lipid kinases has been demonstrated by several mutagenesis studies (2, 41–43). In the YegS structure, the GGDG motif is found at the beginning of the third α-helix, and the amino groups of the helix termini (Gly-66 and Gly-68) directly coor-
dinate the pyrophosphate moiety. The aspartic acid (Asp-67) is not directly involved in phosphate binding but instead is found some 6 Å from the closest phosphate moiety. Instead, it makes hydrogen bonds to main chain amide groups of two residues of the T-motif (Asp-98 and Phe-99). This interaction anchors the end of a very flexible loop, which in *A. fulgidus* NAD kinase (1Z0S) contains a GXG motif conserved in the NAD kinases and involved in coordinating phosphate groups of a bound

**FIGURE 2.** The overall structure of YegS and comparison to NAD kinase. A, YegS structure. The pyrophosphate moiety (POP) of ADP molecule as well as key residues marking the four signature motifs of the lipid kinases are shown. Mg^{2+} marks the novel metal site. The N-terminal domain is the nucleotide binding core of the structure. The two helices shown in green constitute the likely substrate binding region of YegS. B, NAD kinase structure (Protein Data Bank entry 1Z0Z combined with the pyrophosphate moiety of 1Z0S, based on a superposition of the two structures) is shown in an equivalent orientation. The pyrophosphate is indicated. The NAD substrate/product binding site is at the cleft between the N- and C-terminal domains. Figures were generated using PyMOL (pymol.sourceforge.net). C, sequence alignment of YegS to its close relatives, together with the secondary structure elements. YPO2856 is assigned as putative diacylglycerol kinase catalytic domain containing protein from *Yersinia pestis* (other related prokaryotic sequences are not included). Human DGKα is a representative sequence from the diacylglycerol kinase family, MulK is the recently characterized multlipid kinase, CERK is the human ceramide kinase, and SK-1 is the human sphingosine kinase isofrom 1. For a more extensive sequence alignment of lipid kinases, see Fig. 3 in Ref. (2).
pyrophosphate moiety. The similar positioning of the pyrophosphate moieties in A. fulgidus NAD kinase, as in YegS, supports the assignment of the density in YegS as pyrophosphate.

In YegS this loop contains Thr-95, which makes a hydrogen bond to the pyrophosphate molecule. The threonine following the GGDG motif (Thr-69) also makes a hydrogen bond to the pyrophosphate molecule. A threonine in the corresponding position is highly conserved in the NAD kinase family (36). The conserved Asn-12 and Lys-14 of the NXKS motif are positioned at the end of strand B1, and the N terminus of α1 is in close proximity to the pyrophosphate binding site. Asn-12 directly coordinates the pyrophosphate, whereas Lys-14 is positioned so it could potentially provide a ligand to a phosphate moiety, but in the pyrophosphate-bound structure it is not within hydrogen bonding distance. In the pyrophosphate-bound NAD kinase structure from A. fulgidus (1Z0S), Lys-8 is in the same position as Lys-14 in YegS and is coordinating a pyrophosphate moiety. However, Lys-8 is not fully conserved.

The likely substrate for YegS, phosphatidylglycerol, would be located in the membrane. The protein should therefore have structural features allowing it to access membranes. However, YegS remains in the soluble fraction during purification, even after ultracentrifugation, suggesting that membrane interactions are likely to be transient. This is in analogy with some other lipid-interacting proteins such as cytosolic phospholipase D and phospholipase C but in contrast to, for example, cyclooxygenases and Coq7 where the protein binds strongly to the membrane with interfacial hydrophobic interactions.

Because of the strong similarities of the YegS and NAD kinase active sites, it is reasonable to presume conserved reaction geometries in the two enzymes. The position of the head group of the YegS substrate phosphatidylglycerol would then be restricted to the face of the active site crevice formed by the C-terminal domain, that is, the region where the substrate NAD binds in NAD kinase (Figs. 2 and 3). The helix α6, which is not present in NAD kinase, plays an important role in contributing residues to this potential substrate binding region of the YegS active site crevice and provides a fourth sequence motif at the active site. The residues Tyr-164, His-167, and Arg-171 of α6 face the active site (Fig. 3). Tyr-164 is highly conserved in prokaryotic and eukaryotic lipid kinases family members (see Fig. 3 in Ref. 2) and is a Tyr in all human enzymes except ceramide kinase (Fig. 2C). This tyrosine could potentially play an important role interacting with the hydrophobic moiety of lipid substrates (2). Arg-171 and His-167 could potentially play a role in interacting with the phosphate moiety of the phosphoglycerol head group of the substrate. These two residues are not conserved in the mammalian DAG kinases, consistent with the different head groups of their lipid substrates. However, Arg-171 is conserved in many of the prokaryotic family members, indicating a similar substrate specificity for these enzymes as with YegS. In the recently characterized M. tuberculosis DAG kinase Rv2252 the corresponding residue is not conserved, consistent with this enzyme acting on a different acceptor substrate (19).

To define regions that are appropriate for membrane association, we analyzed the electrostatic surface potential of YegS using APBS (44). However, no obvious regions of positive potential or exposed hydrophobic patches were identified on the surface of YegS. Interesting, the Membrane Protein Explorer (MPEx) server (45) predicted the region 157–175, including α6 containing the conserved Tyr-164, as a likely membrane interaction region. The disordered region 147–158 is located close to α6 but is not predicted as a potential membrane-linking segment.

Surprisingly, the structure of YegS reveals a metal binding site in the C-terminal domain. This metal binding site is located between the loop connecting B12 and B13, and B6 that is connecting the two domains. The metal is coordinated by the side chain of Asp-218, together with the two carbonyl oxygens of Leu-215 and Leu-220 and three ordered water molecules (Fig. 4B). One of the water molecules has a second sphere coordination by Asp-125. The electron density and octahedral coordination supports Mg$^{2+}$ as the bound metal, consistent with 200 mM MgCl$_2$ in the crystallization solution. As discussed above, mammalian DAGKs contain accessory domains involved in the regulation of the enzymatic activity. The role of calcium in regulating these enzymes has been suggested to be effected via the high affinity Ca$^{2+}$ binding sites in the tandem EF-hand motifs present in the DAGKs. Unlike the mammalian counterparts, YegS has no accessory domains to regulate the activity; however, it does have a novel metal binding site that is both unpredicted and structurally unique, i.e. not found in any other known protein structure. Interestingly, both the direct coordinating residue, Asp-218, and the second sphere coordinating residue, Asp-125, are conserved in many lipid kinases including human lipid kinases (see alignment in Ref. 2 and Fig. 2C). This indicates that the metal site will most likely be present in many members of the lipid kinase family. The human sphingosine kinase 1 does not contain Asp-218, implying that it may lack this metal binding site. The metal binding site is found in the region of the structure that connects the N- and C-terminal domains. Two aspartate residues that lie between the metal...
binding site and the active site (Fig. 3A), Asp-98 and Asp-269, are also highly conserved in the lipid kinase family. Both of these aspartate residues are at the interface between the N- and C-terminal domains where they stabilize buried loops on the C-terminal domain by coordinating backbone amino groups (residues 202, 205 and 136, 138 respectively). The four conserved aspartates discussed above, Asp-98, -125, -218, and -269, were mutated in a study of Ca\(^{2+}\)/H\(_{11001}\) activation of porcine DGK\(_{H11001}\). Mutation of Asp-650, corresponding to the primary metal ligand Asp-218 in YegS, abolished kinase activity. Mutation of Asp-497, corresponding to the second sphere metal ligand Asp-125 in YegS, reduced kinase activity to 0.9% of wild type. Mutations of the two domain interface aspartates, Asp-465 (Asp-98 in YegS) and Asp-697 (Asp-269 in YegS), abolished all but a significant residual activity, 0.1 and 4%, respectively. The residual activities of the aspartic acid residue Asp-497 (Asp-125 in YegS) and Asp-697 (Asp-269 in YegS) mutations were used to study the effect of these mutations on the Ca\(^{2+}\) regulation, showing that these mutations dramatically attenuated the activation by calcium. It was surprising that these mutations in the catalytic domain affected the calcium regulation, when such regulation had previously been ascribed to localization events mediated by the accessory calcium-binding, EF-hand-containing domain (13).

To define the role of the metal binding site in YegS, we investigated the dependence of the kinase activity on Ca\(^{2+}\) and Mg\(^{2+}\) concentration. The enzyme showed a complicated behavior toward these two metal ions (Figure 5). YegS showed an absolute dependence on the presence of at least one metal. The metal-free YegS lacked enzymatic activity. However, high concentrations of Ca\(^{2+}\) (1 mM) maintain the enzyme activity at a constant level independent of the Mg\(^{2+}\) concentrations. The activity levels of YegS were higher at 5 mM Mg\(^{2+}\), and this activity was slightly reinforced in the presence of 1 mM Ca\(^{2+}\), whereas the activity was diminished upon the addition of 1 mM Ca\(^{2+}\).

The profile of the metal effects on YegS is similar to the human broad specificity DAG kinase MulK, in which the YegS metal binding site is conserved (Fig. 2C) but which does not have EF-hand accessory domains. DGK\(_{a}\)s show a strong activating effect by Ca\(^{2+}\) at 1 mM and 10 mM Mg\(^{2+}\) (2, 46). The interpretation of the Ca\(^{2+}\) data is complicated by the fact that metals, primarily Mg\(^{2+}\), are likely to participate in ATP binding and potentially also in binding the protein to membrane phospholipids. However, the similar Ca\(^{2+}\) activation behavior of MulK and YegS, both lacking accessory domains, implies a similar mechanism of metal activation.

DISCUSSION

The structure of YegS reveals a number of important features that are likely to be conserved in lipid kinases and provides a
structural basis for existing biochemical data for this enzyme family. The lipid kinases have previously been suggested to form a structural family together with PFK and NAD kinase based on the presence of the conserved GGDG motif (1, 47). From the YegS structure it is clear that the lipid kinases and NAD kinase form a subfamily distinct from PFK and other structurally characterized non-protein kinases. A common feature of the lipid kinase and NAD kinase families is that they bind substrates with large hydrophobic moieties, which might explain the common origin of their C-terminal substrate binding domains. The presence of three conserved motifs involved in phosphate interactions in the active site of YegS and NAD kinases implies a very similar phosphorly transfer chemistry in the two enzymes. A detailed understanding of the mechanism of phosphorlyl transfer in NAD kinase and PFK has remained elusive because, as yet, no structure of an informative ternary complex has been determined. An intriguing ping pong-like substrate binding scheme was recently proposed for NAD kinase, based on a structure where ATP was observed to bind in the pockt harboring the NAD nicotine moiety (36). It remains to be shown that this observed ATP binding regime is, in fact, physiologically relevant, and a direct phosphorly transfer from ATP bound in the region as observed in our structure remains the most likely reaction scenario for this kinase family.

The conserved GGDG motif has been shown by mutagenesis studies to be a key component of the catalytic machinery of NAD (35), PFK (38), and the lipid kinases (2, 41–43) and is the most likely site for catalysis of phosphorly transfer. In all available structures in complexes with phosphate donors, the main chain amide groups of this motif coordinate the pyrophosphate structures in complexes with phosphate donors, the main most likely site for catalysis of phosphorly transfer. In all available studies to be a key component in the catalytic machinery of the most likely reaction scenario for this kinase family.

The novel metal binding site in YegS, which is conserved in most lipid kinases, appears to confer a metal dependence to the kinase activity of YegS that is similar to that of mammalian MulK. The previous mutational studies of conserved aspartates in DGKs indicate that this site is functionally important and suggest a role for the domain interface in regulating the activity. The Ca\(^{2+}\) activation mechanism in the mammalian lipid kinase, which contains accessory domains, is likely to be more complex. However, a scenario is emerging where the activity could be controlled both by EF-hand domains and the novel metal site of the catalytic domain, identified in the present study.

We have demonstrated that YegS phosphorylates the common membrane lipid PG to PGP \textit{in vitro}. PGP is normally the intermediate and minor precursor to PG in bacteria and synthesized by an essential enzyme, phosphatidylglycerolphosphate synthase (50), thereby funneling the PG toward the production of cardiolipin (CL). However the reverse reaction has not been described earlier. It is known that the synthesis rate of the PG/CL pathway can increase many-fold upon certain stimuli (51), but PGP does not accumulate, even after overexpression of phosphatidylglycerolphosphate synthase. Interestingly, the \textit{E. coli} Co-Response Data Base (csdb.mpimp-golm.mpg.de/index.html) (52) shows a significant correlation between expression of the YegS gene and a handful of stress response genes (\textit{e.g.} \textit{iddC}, \textit{ycaC}, \textit{ynbC}, \textit{ynCG}, \textit{csbD}, \textit{hchA}, and \textit{gadX}). Several of these are involved in compensatory mechanisms for acid stress (53). Acid stress is encountered by bacteria entering the stomach-intestinal systems in animals, which may explain the occurrence of YegS homologues in several pathogenic bacteria. A possible role for the YegS reaction in bacteria is to promote addition of negative phosphate moieties on PG, which could compensate for charge loss on lipid head groups upon internal acidification of, for example, cardiolipin groups. Likewise, it was suggested that the related \textit{M. tuberculosis} Rv2252 DAG kinase may have a potential role during pathogenesis (16).

The metal dependence of YegS is complex (Fig. 5). The most likely mechanistic interpretation of the Mg\(^{2+}\) effects is binding of the metal in the active site, together with the phosphate donor. An alternative additional possible function for Mg\(^{2+}\) is to mediate interactions with the lipid bilayer. The requirement of calcium for significant YegS activity is interesting and intriguing. With the observed activation profile it is less likely that calcium is directly competing with Mg\(^{2+}\) binding, except maybe at high Mg\(^{2+}\) concentrations (5 mM). Instead, the molecular mechanism behind the calcium effect appears to be unique and indicating different binding sites for the two metals. The observation of an attenuation of calcium activation upon mutation of a residue in DGK\(\alpha\), corresponding to the second sphere metal ligand Asp-125 in YegS (2), indicates that the novel metal site may indeed, at least in some of the lipid kinases, respond to calcium. In \textit{E. coli} and other bacteria, calcium seems to have several signaling effects (54), but it is unknown whether variations in the Ca\(^{2+}\) concentration have a role in controlling the \textit{in vivo} activity of any lipid kinase. However, further studies are required to establish the detailed role of calcium and magnesium for individual lipid kinases, as well as the role of the
Crystal Structure and a Novel Function of YegS

novel metal site in the modulation of the activity of lipid kinases.

The structure determination of a prototypic lipid kinase family member, together with the identification of a novel metal site, conserved in most lipid kinases, now allows for focused efforts directed toward elucidating the structure-function relationship of enzymes of this important family protein.

Acknowledgments—We thank staff at beamline BL-1 of Bessy, Berlin, and beamline ID14-4 and ID14-3 at European Synchrotron Radiation Facility, Grenoble, for technical support during data collection. We also thank Martin Andresson and Benita Engvall for help with the original cloning of YegS.

REFERENCES

Crystal Structure of YegS, a Homologue to the Mammalian Diacylglycerol Kinases, Reveals a Novel Regulatory Metal Binding Site
H. M. Amin Bakali, Maria Dolores Herman, Kenneth A. Johnson, Amélie A. Kelly, Åke Wieslander, B. Martin Hallberg and Pär Nordlund


Access the most updated version of this article at doi: 10.1074/jbc.M604852200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2007/03/19/M604852200.DC1

This article cites 53 references, 19 of which can be accessed free at
http://www.jbc.org/content/282/27/19644.full.html#ref-list-1