Structure, Vol. 11, 703–713, June, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S0969-2126(03)00094-7

The Crystal Structure of Choline Kinase Reveals a Eukaryotic Protein Kinase Fold

Daniel Peisach, Patricia Gee, Claudia Kent, and Zhaohui Xu*
Department of Biological Chemistry
University of Michigan Medical School
1301 East Catherine Road
Ann Arbor, Michigan 48109

Summary

Choline kinase catalyzes the ATP-dependent phosphorylation of choline, the first committed step in the CDP-choline pathway for the biosynthesis of phosphatidylcholine. The 2.0 Å crystal structure of a choline kinase from *C. elegans* (CKA-2) reveals that the enzyme is a homodimeric protein with each monomer organized into a two-domain fold. The structure is remarkably similar to those of protein kinases and aminoglycoside phosphotransferases, despite no significant similarity in amino acid sequence. Comparisons to the structures of other kinases suggest that ATP binds to CKA-2 in a pocket formed by highly conserved and catalytically important residues. In addition, a choline binding site is proposed to be near the ATP binding pocket and formed by several structurally flexible loops.

Introduction

Phosphatidylcholine is a major structural component of eukaryotic cellular membranes and serum lipoproteins and serves as a precursor for the production of lipid second messengers (Exton, 2000). The predominant pathway for the biosynthesis of phosphatidylcholine in animals and plants is the three-step CDP-choline pathway, whereby choline is first converted into phosphocholine, which then reacts with CTP to form CDP-choline. Finally, the phosphocholine moiety is transferred to diacylglycerol to produce phosphatidylcholine. Choline kinase catalyzes the first step of the pathway, the ATP-dependent phosphorylation of choline, and has been reported to regulate the CDP-choline pathway (Kent, 1990). In mammals, phosphocholine has been reported to be mitogenic (Cuadrado et al., 1993), and inhibitors of choline kinase are being investigated as therapeutic agents in cancer research (Hernandez-Alcoceba et al., 1997). In addition to its importance in the biosynthesis of phosphatidylcholine, choline kinase functions in the biosynthesis of cell wall-associated phosphocholine in bacteria (Bean and Tomasz, 1977; Poxton and Leak, 1977; Tuomanen, 1999). For example, the necessity of phosphocholine to *S. pneumoniae* cell wall structure (Tuomanen, 1999) renders its choline kinase as a possible antibacterial target.

A partially purified preparation of choline kinase from brewer's yeast was first studied fifty years ago (Wittenberg and Kornberg, 1953). The enzyme from several mammalian sources (Ishidate et al., 1984; Porter and Kent, 1990; Uchida and Yamashita, 1990) as well as *Saccharomyces cerevisiae* (Kim et al., 1998) has been subsequently purified to homogeneity. Most choline kinases utilize both choline and ethanolamine as substrates to accept phosphotransfer from ATP, although K_m values for choline are lower than those for ethanolamine. Enzymes that are specific for ethanolamine have also been reported (Lykidis et al., 2001; Pavlidis et al., 1994; Uchida, 1997). Genes and cDNAs for bona fide choline/ethanolamine kinases have been isolated from mammals (Aoyama et al., 1998, 2000; Uchida and Yamashita, 1992), plants (Al-Malki et al., 2000; Monks et al., 1996), and yeast (Hosaka et al., 1989), and recent genome sequencing efforts have identified many more putative choline and/or ethanolamine kinases. Alignment of the protein sequences of choline kinase from several species reveals several highly conserved regions, including two clusters of conserved residues (Figure 1). The first, also known as the Brenner's motif, has been identified in many enzymes catalyzing phosphotransfer reactions, including protein kinases (Aoyama et al., 2000; Brenner, 1987). The second conserved cluster appears to be specific for members of the choline/ethanolamine kinase family (the choline kinase motif), raising the possibility that it might be important for binding choline or ethanolamine.

Multiple choline kinase isozymes are found in many organisms, as indicated by both biochemical (Ishidate et al., 1982, 1985; Porter and Kent, 1990) and genetic evidence (Aoyama et al., 2000; Monks et al., 1996; Uchida, 1994). It is possible that induction of expression of different isozymes is a mechanism for regulating choline kinase activity, but roles for the various isoforms have not been well defined. Analysis of the *C. elegans* genome has identified seven choline kinase-like genes, which fall into three groups based on similarity in sequence to mammalian choline kinases; group A is most similar while groups B and C are less so (Gee and Kent, 2003). Four of the seven genes (two from the A group and two from the B group) have recently been cloned and expressed in insect cell lines. All of them have choline kinase activity; those that most closely resemble the mammalian choline kinases are the most active. An example of this is isoform A-2 (CKA-2). It exists as a dimer composed of two identical subunits of 429 residues each. Biochemical studies have shown that this enzyme has K_m values of 1.6 mM and 2.4 mM for choline and ATP, respectively, and a k_{cat} of 74 s^{-1}. The enzyme requires magnesium for activity, has a very alkaline pH optimum, and is much more active with choline as a substrate than with ethanolamine. We have begun to use this enzyme to study the functions of the various choline kinase-like genes in *C. elegans*, as well as to provide a model for structure-function studies of the broader choline kinase family.

In the present study, we determined the three-dimensional structure of CKA-2 in order to understand the

*Correspondence: zhaohui@umich.edu

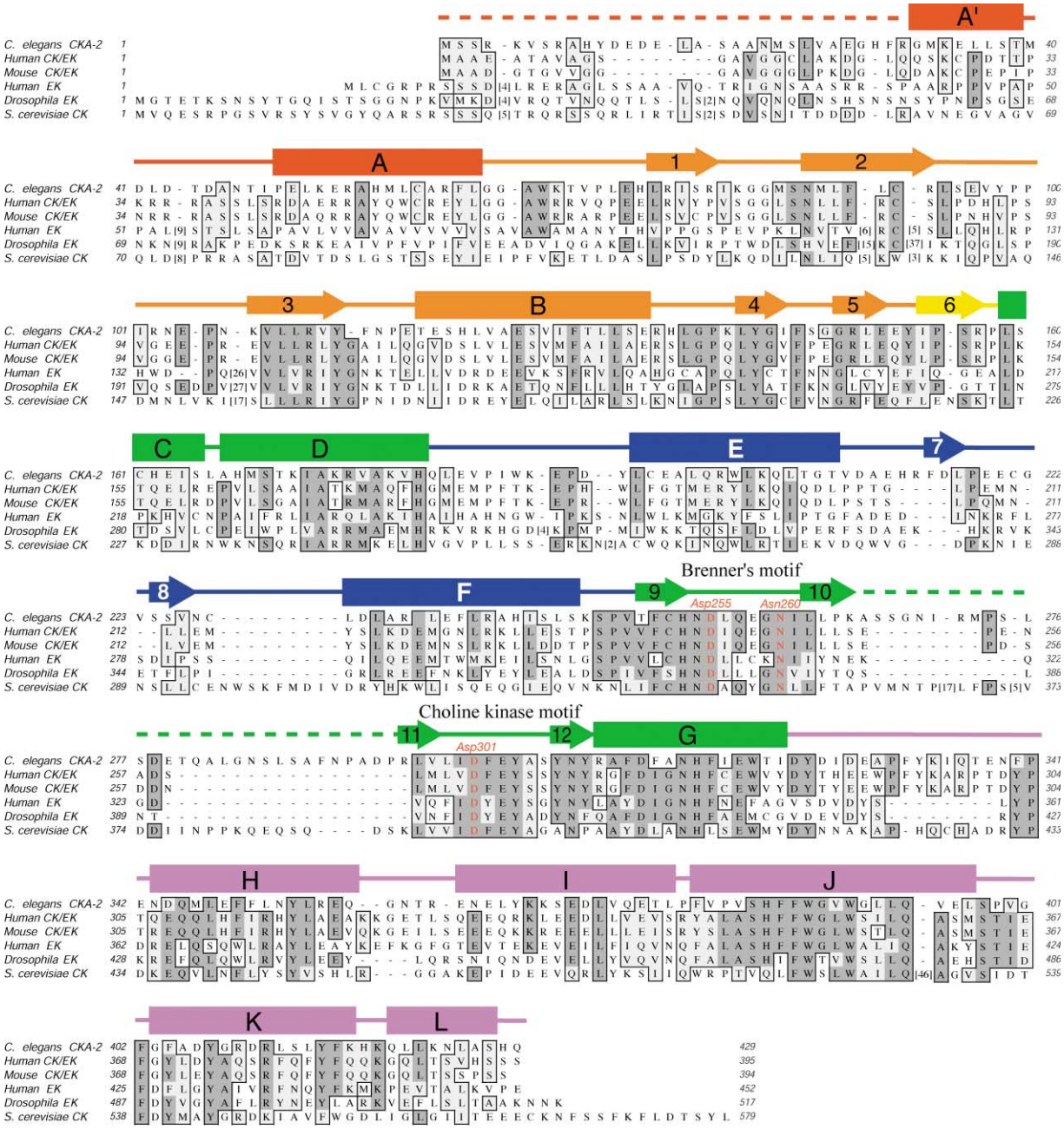

Figure 1. Sequence Alignment of Choline/Ethanolamine Kinases

The sequences of *C. elegans* choline kinase CKA-2, human choline/ethanolamine kinase (CK/EK), mouse choline/ethanolamine kinase, human ethanolamine kinase, *Drosophila* ethanolamine kinase, and *S. cerevisiae* choline kinase were aligned with the program ClustalW (Higgins et al., 1994). Invariant residues are darkly shaded and outlined. Similar, but not identical, residues are lightly shaded and outlined. Secondary structure elements are indicated above the sequence: α helices, rectangles; β strands, arrows; other elements, solid lines; structurally unobserved residues at the N terminus and midsequence, dashed lines. These elements are colored on the basis of their locations in the sequence (see text for detail). For purpose of clarity, sequences in non-CKA-2 proteins where no homology is identified with CKA-2 have been replaced by bracketed numbers indicating the number of residues omitted from the display. Three highly conserved residues from the proposed ATP binding site are indicated with red letters.

molecular mechanisms of its catalysis and substrate selectivity for choline over ethanolamine and to establish a foundation for structure-function relationships. We report here the general features of this choline kinase and detailed structural information on its active and substrate binding sites. It is noteworthy that CKA-2 has remarkable structural similarity to the eukaryotic protein kinase family and the bacterial aminoglycoside phosphotransferase family, despite limited sequence similarity. Our comparison suggests that choline kinase could

use a similar mechanism to those of other phosphotransferases in catalysis. This work represents the first structural analysis of a choline/ethanolamine kinase.

Results and Discussion

Structure Determination

The CKA-2 protein used in the structure determination was overexpressed and purified from *E. coli*. The bacterially expressed protein had a nearly identical catalytic profile to the one expressed in eukaryotic (insect) cells and was crystallized in two different, yet related, crystal forms. Crystals grown from the selenomethionine-substituted protein belong to the orthorhombic space group $P2_12_12$, and those from the native protein belong to the monoclinic space group $P2_1$. The two have nearly identical crystal packing, with the crystallographic two-fold c axis in the $P2_12_12$ crystals becoming a noncrystallographic two-fold axis in the $P2_1$ crystals. The structure of CKA-2 was first determined by the multiple-wavelength anomalous diffraction (MAD) method with selenium atoms in selenomethionine as the anomalous scattering atoms and later refined against a 2.0 Å resolution native data set. The refined model consists of a choline kinase dimer including residues 32–265 and residues 296–426 for subunit A, residues 40–265 and residues 296–426 for subunit B, one calcium ion for each subunit, and a total of 342 water molecules, with an R factor of 20.9% and an R_{free} of 23.1%. Of all the nonglycine residues, 92.9% are located in the most favored regions on the Ramachandran plot, and the rest are all located in additional allowed regions. Crystallographic statistics are summarized in Table 1.

Overall Fold

CKA-2 from *C. elegans* is a homodimeric enzyme with a two-fold symmetry axis. Each monomer is composed of two domains, a small 153-residue N-terminal domain and a large 271-residue C-terminal domain, connected by a five-residue linker (Figure 2). The overall architecture of the protein is similar to that seen for members of the eukaryotic protein kinase (ePK) family and the bacterial aminoglycoside phosphotransferase (bAP) family, including the catalytic domains of casein kinase, cyclic AMP-dependent protein kinase (PKA), cyclin-dependent kinase, src tyrosine-protein kinase, insulin-like grown factor 1 receptor kinase, and 3′, 5″-aminoglycoside phosphotransferase type IIIa [APH(3′)-IIIa] (Favelyukis et al., 2001; Hon et al., 1997; Knighton et al., 1991; Lawrie et al., 1997; Longenecker et al., 1996).

The N terminus of the protein begins with 31 unobserved and presumably disordered residues, followed by a small subdomain that contains two helices (A′ and A; Figures 1 and 2A, red) that lie across a well-conserved structural motif (Figures 1 and 2A, orange). This motif is composed of a five-stranded highly twisted antiparallel β sheet made up of strands 1, 2, 3, 5, and 4, in that order. Helix B is inserted between strands 3 and 4 and lies beneath strands 4 and 5. This helix provides a large interface that stabilizes choline kinase as a dimer (see next section). The structural similarity among CK, ePK, and bAP in this region is striking. For example, despite the low sequence identity between CKA-2 and PKA, the rms deviation for the structurally conserved main chain atoms in the N-terminal domain is only 1.7 Å.

The N-terminal domain is linked to the C-terminal domain through a short strand (strand 6) bordered by two 90° turns (Figures 1 and 2A, yellow). The C-terminal domain can be subdivided into three regions: a central core subdomain (Figures 1 and 2A, green), an insertion subdomain (Figures 1 and 2A, blue), and a C-terminal subdomain (Figures 1 and 2A, magenta). The central core consists of three helices (C, D, and G) and a long hairpin-shaped loop that incorporates two short stretches of antiparallel β strands (9, 10, 11 and 12). Strands 10 and 11, together with strand 6, form a small antiparallel β sheet on the right side of the molecule (Figure 2A). This long hairpin loop contains many of the conserved, functionally important residues. For example, strands 9–10 house the Brenner's motif, and strands 11–12 house the choline kinase motif (Figure 2B). As will be discussed in following sections, these sequences are most likely involved in binding substrates and performing catalysis. In addition, strands 10 and 11 are connected by a long stretch of sequence (residues 265–296) that appears to extend away from the core structure. Since there is no electron density observed for this sequence, it is assumed to be disordered. The structurally conserved main chain atoms in this subdomain display a high degree of conformational similarity to several segments invariantly present in ePK and bAP. For example, the superimposition of these segments between CKA-2 and PKA gives an rms deviation for main chain atoms of 2.0 Å.

The insertion subdomain is located between helix D and strand 9 and is composed of two helices (E and F) connected by a β hairpin motif (strands 7 and 8). This insertion wraps around the left side of the structure (Figure 2A). Helix E connects to helix D via a long S-shaped loop that is part of the dimer interface.

The C-terminal subdomain has five helices (H–L) that form the bottom section of the structure. At the back of the molecule, helix G (from the conserved core subdomain) connects to helix H through a long hairpin-like loop (Figure 2A). In a following section, we propose that this loop could control the access to the active and substrate binding sites. Helix H runs beneath helix D and links to helices I and J through a reverse turn. Helices I and J can be viewed as one very long helix that runs from one side of the molecule to the other. Two prolines in the middle of the sequence introduce a near-30° kink, thus breaking the helix in two. Following another four-residue reverse turn are helices K and L, which extend to the C terminus.

Structural comparison between CKA-2, PKA, and APH(3′)-IIIa shows that all three structures contain the conserved N-terminal structural motif (strands 1–5, including helix B) and the C-terminal central core subdomain (helices C, D, and G and strands 9–12). While PKA appears to have the minimal amount of insertion and decoration on this structural core, both APH(3′)-IIIa and CKA-2 have evolved to have more structural elements attached (Figure 3) (Hon et al., 1997). For example, PKA

Table 1. Crystallographic Data Statistics

Data Collection Statistics						
Crystal	Space Group		Cell Dimensions			
Native	$P2_1$		a = 55.4 Å, b = 102.1 Å, c = 85.4 Å, β = 90.66°			
SeMet	$P2_12_12$		a = 85.7 Å, b = 102.4 Å, c = 55.4 Å			
Data Set	d_{min} (Å)	Number of Measurements	Number of Unique Reflections	Completeness (%)[a]	I/σ[a]	R_{sym} (%)[a,b]
Native	2.0	195,184	60,790	99.1 (99.8)	14 (3.2)	6.4 (30.2)
SeMet λ1 (0.9792 Å)	2.2	293,359	24,791	99.9 (100.0)	31 (8.7)	9.2 (22.9)
SeMet λ2 (0.9793 Å)	2.3	257.998	21,840	99.9 (100.0)	31 (8.2)	8.6 (23.5)
SeMet λ3 (0.9565 Å)	2.5	197,049	17,090	99.9 (100.0)	27 (6.7)	9.9 (27.8)

MAD Phasing Statistics	Peak		Edge		Remote	
	Isomorphous	Anamolous	Isomorphous	Anamolous	Isomorphous	Anamolous
Phasing power[c]						
Centric	—	—	0.37	—	0.39	—
Acentric	—	1.54	0.45	0.97	0.46	0.72
FOM[d]						
Centric	0.192					
Acentric	0.482					
FOM after DM[e]	0.924					

Refinement Statistics (Native Data Set)	
Number of reflections (working/test)	54,579/4,367
Number of nonhydrogen atoms	6,155
Number of water molecules	342
Resolution (Å)	37.5–2.00
R_{cryst}/R_{free} (%)[f]	20.9/23.1
Bond length deviation (Å)	0.006
Bond angle deviation (°)	1.2
Average B factor (Å^2)	25.3

[a] Values in parentheses are for the highest resolution bin.
[b] $R_{sym} = \Sigma_h\Sigma_i|I_i(h) - <I(h)>|/\Sigma h<I(h)>$, where $I_i(h)$ is the i^{th} measurement and $<I(h)>$ is the weighted mean of all measurements of I(h).
[c] Phasing power = $<|F_H(calc)|>/<E>$, where E is the phase-integrated lack of closure.
[d] Figure of merit.
[e] Figure of merit after solvent flipping in the program DM.
[f] $R = \Sigma(|F_{obs}| - K|F_{calc}|)/\Sigma|F_{obs}|$. R_{free} is the R value obtained for a test set of reflections that consisted of a randomly selected 8% subset of the diffraction data used during refinement of σ_A value calculations.

has a short connection between helix D and the long β hairpin loop (starting with strand 9), but APH(3′)-IIIa and CKA-2 add two helices and an 18-residue turn, with CKA-2 forming a small β sheet in this turn. In addition, the C-terminal subdomain after the conserved helix G is very different in all three structures. APH(3′)-IIIa has only two short helices. PKA, on the other hand, has three helices connected by fairly floppy loops and finishes by wrapping its last 40 residues around the molecule and back to the N-terminal domain. Finally, CKA-2 has five helices connected with loops of various lengths. Together with the C-terminal central core subdomain, these helices create a sizable pocket that could be the choline binding site of the enzyme. Clearly, there is much more variability in the C-terminal domain than in the N-terminal domain, but this is most likely due to the variety of substrates on which these proteins operate.

Structure of the Dimer and Dimer Interface

Mammalian choline kinases are either dimers (Uchida and Yamashita, 1990) or tetramers (Porter and Kent, 1990) in solution. Gel filtration and dynamic light-scattering experiments show that CKA-2 has a molecular weight of about 100,000 Da, suggesting a dimer structure (data not shown). The asymmetric unit of the native crystal contains two monomers related by a noncrystallographic symmetry (NCS) two-fold axis (179.5°). Inspection of crystal packing suggests that these two molecules represent the biological dimer in solution (Figure 4A). Each monomer in this pair buries 8.5% of its total solvent-accessible surface (2676 Å^2, as calculated by the program CNS) in intermolecular contacts, an area that is in the range expected for intersubunit contacts in oligomeric proteins (as little as 5% for larger proteins and up to 20% for smaller ones) (Janin and Chothia, 1990). All other molecular pairs involved in crystalline contacts bury less than 4.8% of solvent-accessible surface. A similar dimer interface is also observed in crystals of the selenomethionine-substituted protein (space group $P2_12_12$) as a crystallographic contact along the two-fold c axis.

The dimer interface is 55% nonpolar, which is on the low end, but close enough to the average value of 67% for the subunit interfaces of oligomeric proteins. Polar interactions including salt bridges and hydrogen bonds

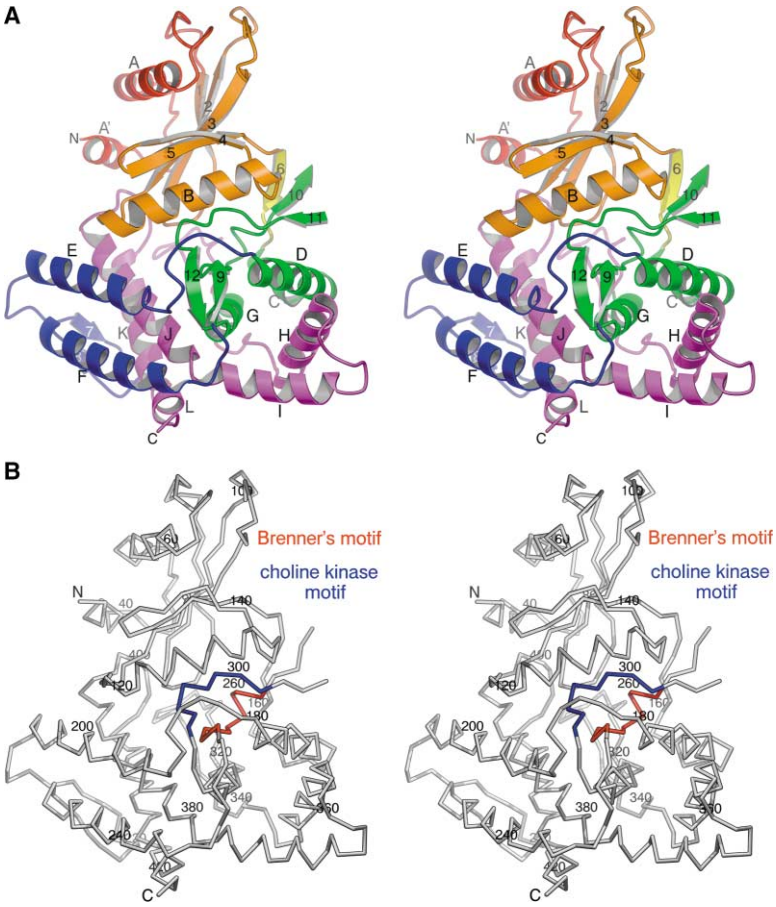

Figure 2. The Structure of the Choline Kinase CKA-2 Monomer

(A) A stereo ribbon drawing of the CKA-2 monomer. The secondary structure elements are labeled and colored as in Figure 1. α helices are lettered and drawn as coils, β strands are numbered and drawn as arrows, and other elements are drawn as tubes.

(B) A stereo Cα trace of the CKA-2 monomer with every 20th amino acid numbered. The Brenner's motif, red; the choline kinase motif, blue.

also contribute to the interface. In the following description, the two subunits are, when necessary, labeled A and B in order to distinguish residues belonging to one or the other. Four regions are particularly involved in the dimerization: the loop following helix A, strand 4, helix B, and the S-shaped loop that connects helices D and E. The most significant interface is composed of two symmetry-related helices B forming a right-handed coiled-coil structure (Figure 4B). The interaction between the two helices is notably hydrophobic, arising from Leu122, Val123, the β carbon of Ser126, Val127, the γ carbon of Thr130, and Leu131 from both mono-

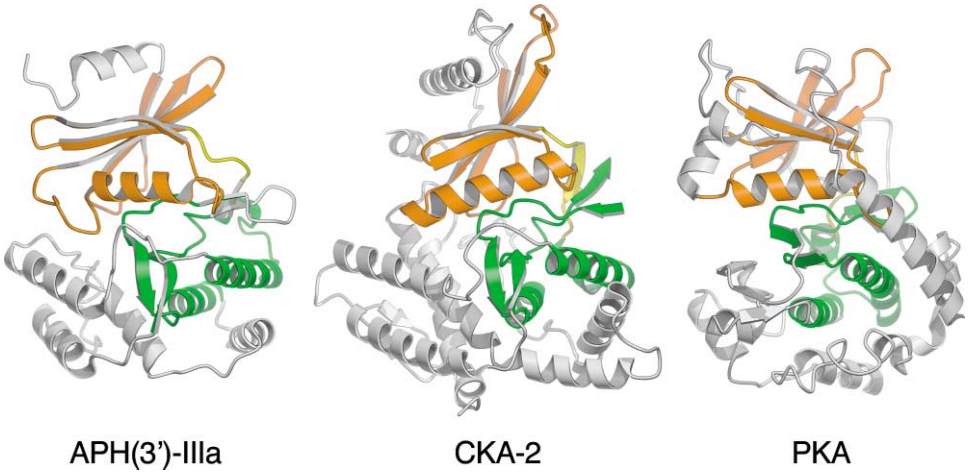

Figure 3. Structure Comparison of CKA-2 with Aminoglycoside 3′-Phosphotransferase [APH(3′)-IIIa] and the Catalytic Domain of cAMP-Dependent Protein Kinase (PKA)

Ribbon diagrams of APH(3′)-IIIa (Protein Data Bank accession code 1J7U), CKA-2, and PKA (Protein Data Bank accession code 1JLU). The structurally conserved cores of the N-terminal domain, C-terminal domain, and linker region are colored orange, green, and yellow, respectively.

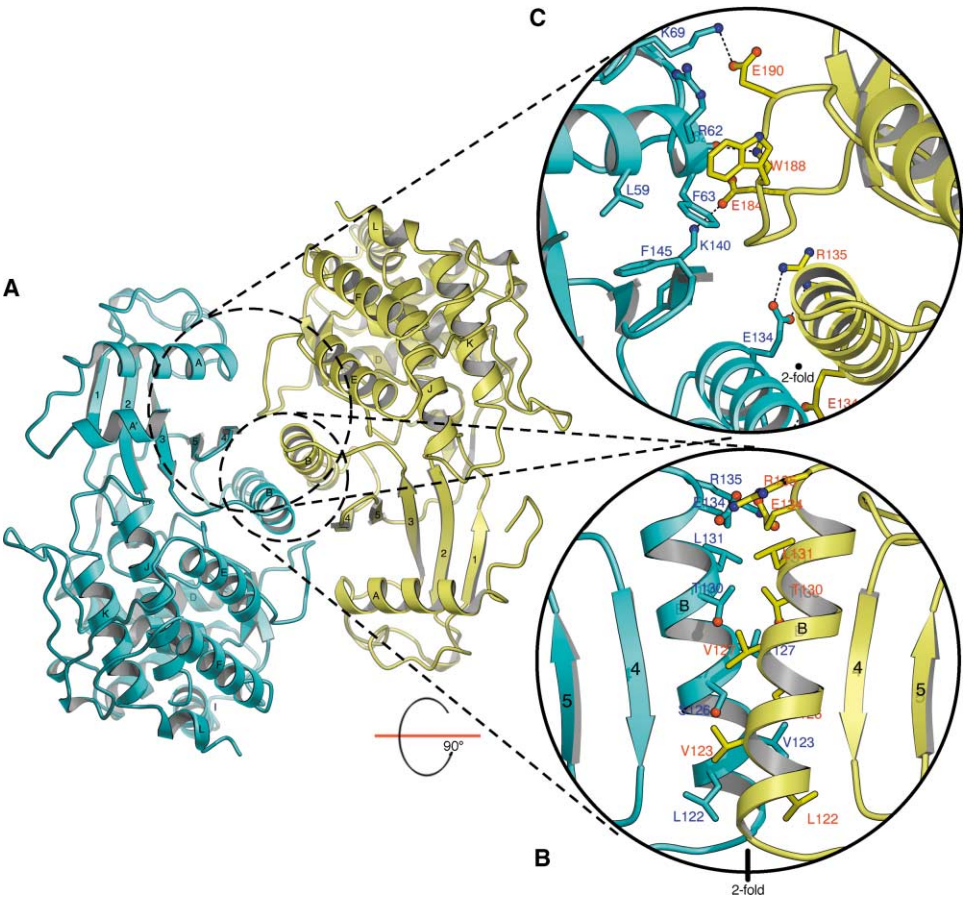

Figure 4. The Structure of the Choline Kinase CKA-2 Dimer

(A) A ribbon drawing of the CKA-2 dimer looking straight down the two-fold axis of symmetry. The two subunits are colored cyan and yellow. Rotating the CKA-2 molecule in Figure 2 approximately 90° around a vertical axis (bringing the left half out of the page and the right half into the page) produces the cyan subunit.

(B) An enlarged view of the residues involved in interactions between the two symmetry-related helices B. Note the predominant hydrophobic nature of the interaction.

(C) An enlarged view of the residues involved in the dimer interface between the S-shaped loop of the yellow subunit and the loop following helix A and strand 4 of the cyan subunit. Only residues that are involved in direct salt bridges, hydrogen bonds, or van der Waals interactions are shown. Salt bridges and hydrogen bonds, dashed lines; labels of residues from the yellow subunit, red; labels of residues from the cyan subunit, blue. The view in (B) is orthogonal to those in (A) and (C).

mers. No water molecules are found at the helix-helix interface. At the C-terminal ends of these two helices are bidentate salt bridges between GluA134 and ArgB135 as well as between GluB134 and ArgA135. An additional dimer interface is formed between the S-shaped loop of one subunit and the loop following helix A and strand 4 of the second and symmetrical subunit. Only a small number of nonpolar residues are involved in this interface (Figure 4C). More precisely, Trp188 from one monomer faces a hydrophobic patch on the other formed by the Cβ-Cγ bond of Arg62, Phe63, Leu59, and Phe145. The rest are a combination of polar interactions, including salt bridges between GluA184 and LysB140 and between GluA190 and LysB69, and a hydrogen bond between the main chain amide of TrpA188 and the main chain carbonyl of ArgB62 (Figure 4C). Water-mediated hydrogen bonds are also found between GluA190 and the main chain carbonyl of AlaB61, between GluA190 and ArgB62, between LysA189 and the main chain carbonyls of TyrB142 and PheB63, and between the main chain carbonyl of ProA186 and the main chain amide of IleB144. Because the structure of the CKA-2 dimer has a two-fold axis of symmetry, all interactions described also exist in a symmetrical fashion.

Proposed ATP Binding Site

The structure of CKA-2 does not have any bound substrates. High degrees of structural similarity between CK, ePK, and bAP, however, allow us to speculate where the substrate binding site might be located. On the basis of a structural alignment identified by the program DALI (Holm and Sander, 1993), we propose that ATP would be bound in a similar pocket to that seen for ePK and bAP, which is defined by helix B, the loop between strands 1 and 2, the β sheet made up of strands 6, 10, and 11, and the loop connecting strands 11 and 12. This pocket has been used to bind ATP in all known structures of ePK and bAP. For clarity of discussion, we

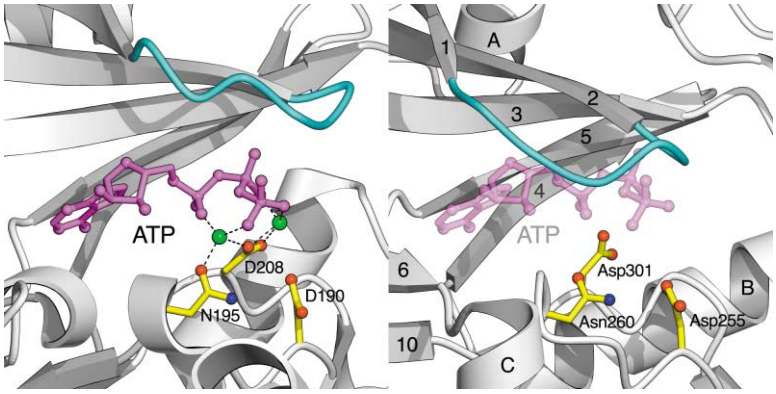

Figure 5. The Proposed ATP Binding Site in CKA-2

A close-up view of the ATP binding sites in APH(3′)-IIIa (left) and CKA-2 (right). Proteins are shown as gray ribbon drawings, except for the following: the ATP binding loop is cyan; the side chains of three highly conserved and catalytically important residues D190 (Asp255), N195 (Asn260), and D208 (Asp301) are yellow ball and stick models; ATP molecules are magenta ball and stick models; and the Mg^{2+} ions are green spheres. Please note that the ATP on the right panel was not seen in the structure but has been drawn here to help identify the highly conserved nature of the region and to illustrate how the cyan loop lies in the ATP binding site.

will use one-letter amino acid codes followed by number for either PKA or APH(3′)-IIIa residues and three-letter amino acid codes for CKA-2 residues. The conserved phosphotransferase Brenner's motif [(FL)(CS)HNDhX_3N] is located on the connecting loop between strands 9 and 10 (Figures 1 and 2B). In APH(3′)-IIIa, residues F186 (Phe251) and H188 (His253) from this motif play structural roles in forming the ATP binding site. The phenylalanine is a part of a conserved hydrophobic cluster, which, in CKA-2, is made up of two phenylalanines and three leucines. The histidine forms hydrogen bonds with the main chain amide of residue D190 (Asp255) and the main chain carbonyl of residue I207 (Ile300). These two sets of interactions appear to help orient the side chain of the catalytically important residues D190 (Asp255) and D208 (Asp301). D208 (Asp301), which is located on the choline kinase motif (see below), directly coordinates two magnesium ions necessary for catalysis. D190 (Asp255) is in the Brenner's motif and probably orients the second (choline in CKA-2) substrate. Finally, N195 (Asn260), which is in the Brenner's motif, also coordinates one of the magnesium ions. The similarity of spatial arrangement of these residues in CKA-2 to those in APH(3′)-IIIa (Figure 5) suggests that CKA-2 might bind ATP and catalyze phosphotransfer reactions utilizing a similar mechanism.

A second conserved motif, (ILV)X_2ID(FWY)E(YF)X_3NX_3(FYW)DX_6E, also known as the choline kinase motif because it appears specifically in choline and ethanolamine kinases, is located on strands 11 and 12 and helix G. A related sequence, IDX_{11}D, is seen in bAP. The first aspartate in this sequence in APH(3′)-IIIa is the aforementioned D208 (Asp301), which coordinates to the catalytic magnesium ions. The second aspartate (Asp313) is located on helix G, and its side chain is involved in hydrogen bonding to the main chain amide of residues His253 and Asn254, both of which are well conserved. His253 has been mentioned above to stabilize the main chain geometry of Asp255 and Ile300. Similarly, the side chain of Asn254 interacts via hydrogen bonding to the side chain of Asn308, the conserved asparagine residue in the choline kinase motif. These interactions make up a network of hydrogen bonds involving highly conserved residues that seems to stabilize catalytically important residues. In addition to the aforementioned interactions, the hydroxyl group of Tyr353 hydrogen bonds to the imidazole group of His181, which, in turn, interacts with the carboxylate group of Asp313. These interactions are notably all located just beneath the dimer interface. Although we do not fully understand their functions yet, we speculate that they could serve to stabilize the dimer interface.

In the apostructure of APH(3′)-IIIa, the loop between strands 1 and 2 (ATP binding loop, the structural equivalent of the P loop in protein kinases [Kinoshita et al., 1999]) lies in the empty ATP binding site (Burk et al., 2001). In our CKA-2 structure, the corresponding loop also lies in the pocket where ATP is suggested to bind (Figure 5). On the basis of the structural similarity, we propose that this loop will have to move out of the way for ATP to bind and for catalysis to occur. In known ePK structures, substrate binding is associated with large domain swings, but, in the APH(3′)-IIIa structure, little else except for the loop that lies over the ATP binding site changes upon nucleotide binding. It was suggested that, unlike ePK, whose structure is usually dynamic and whose activity is highly regulated, APH(3′)-IIIa is structurally rigid and probably constitutively active (Burk et al., 2001). Thus, little domain movement is associated with nucleotide binding. Since choline kinase is structurally more similar to APH(3′)-IIIa than to PKA, it is likely that there will be no major domain movement accompanying ATP binding.

Proposed Choline Binding Site

Studies on reaction mechanisms of ePK and bAP have pointed to an ordered, sequential mechanism (McKay and Wright, 1995; Whitehouse et al., 1983). First, ATP binds to the enzyme, then the second substrate binds to the enzyme, and then the phosphotransfer reaction occurs. The product then quickly diffuses out, and the rate-limiting release of ADP occurs. The ternary complex of PKA with ATP and an inhibitor peptide complex shows that both the terminal phosphate and the acceptor hydroxyl group bind to D166, the aspartate that has been proposed to function as the catalytic base in the reaction (Knighton et al., 1991).

Since Asp301 is the structurally analogous residue to D166 in our alignments, it is therefore expected that there might exist a pocket near Asp301 that could serve

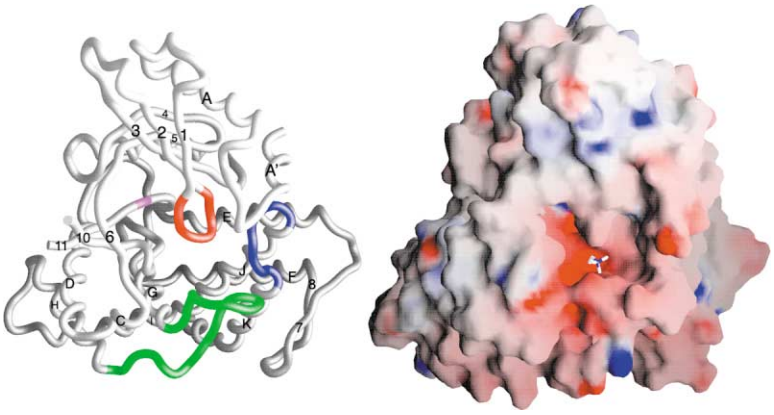

Figure 6. The Proposed Choline Binding Pocket in CKA-2

A Cα trace (left) and a solvent-accessible surface (right) of CKA-2 drawn in the same orientation, which is approximately from the backside of Figure 2. On the left panel, the loops are colored as follows: the ATP binding loop (residues 81–88), red; residue Asp301 (from the choline kinase motif), magenta; the choline pocket loop 1 (residues 322–343), green; choline pocket loop 2 (residues 396–405), blue. On the right panel, the surface is colored on the basis of the electrostatic potential of the molecule (ranging from −21 to +21 kT). To illustrate the size of the pocket opening, we have modeled in a choline molecule.

as the choline binding site. Examination of the CKA-2 surface suggests that such a pocket is located at the domain interface of the molecule (Figure 6). The pocket is lined on the bottom by the loop between helices G and H (choline pocket loop 1), on the right by the reverse turn between helices I and J (choline pocket loop 2), and on the top by the loop between strands 1 and 2 (ATP binding loop). The left side of the pocket is connected to the ATP binding site. The pocket is large enough to fit a choline molecule, and its volume would increase in the presence of ATP when the loop between strands 1 and 2 is expected to move up to accommodate the nucleotide.

We propose that this pocket is the choline binding site on the basis of the following observations. First, it is located in an area where choline kinase is uniquely different from ePK and bAP. Second, the structural elements lining the pocket have the highest average temperature factors within the structure, suggesting their flexibilities, which are necessary for choline binding and release. Third, the pocket also has an overall negative electrostatic potential that can accommodate the positive charge on the quaternary amine of choline when it binds (Figure 6). Finally, the pocket is away from the dimer interface (Figure 4). Although choline kinase is a dimer in solution, there is little cooperativity in catalysis between the two subunits. The distance between the proposed choline binding site and the dimer interface supports the catalytic independence of the two subunits. The precise details of the binding sites can only be answered when a structure of a holoenzyme becomes available.

Calcium Binding

There is one strong solvent peak (6 σ) in the electron density map for each monomer that cannot be accounted for by a water molecule. It forms tight interactions (<2.5 Å) with the side chains of Glu258 and Glu320 and the main chain carbonyls of Glu320 and Ile323 (Figure 7). Because the protein was crystallized in the presence of 200 mM calcium chloride, we interpret this putative solvent peak as arising from a calcium ion. Since calcium ions typically form octahedral coordination, a water molecule not observed at the current resolution likely provides the missing sixth ligand.

Calcium inhibits the activity of choline kinase (P.G. and C.K., unpublished data), but, curiously, the calcium binding site lies a full 10 Å away from the proposed Mg^{2+}-ATP binding site. It is thus unlikely that calcium competes directly with magnesium for the catalytic cation binding site. On the other hand, Glu258 is part of the Brenner's motif and is located between highly conserved residues Asp255 and Asn260 (Figure 7). It is conceivable that the binding of Glu258 to the calcium ion would add structural rigidity to the Benner's motif, making the phosphotransfer reaction less efficient. Currently we are attempting to crystallize the protein with either or both of its substrates bound in the absence of calcium.

Biological Implications

Choline kinase catalyzes the first committed step in the CDP-choline pathway for the biosynthesis of phosphatidylcholine. Although choline kinase has been studied for more than fifty years, there has been no structural information available, which has hindered the development of an understanding of the mechanism for this enzyme. The first structure determination of a choline kinase presented here provides new insights into the mechanism of this enzyme family. The realization of the structural relationship of choline kinase to protein kinases and aminoglycoside phosphotransferases provides a foundation from which we can derive a better understanding of the choline kinase mechanism. Together with the structure determination for CDP:phosphocholine cytidylyltransferase (Weber et al., 1999), the choline kinase structure should lead to a deeper understanding of the phosphatidylcholine biosynthesis pathway.

CKA-2 has a two-domain protein kinase fold with the active site located between the two domains. Residues that are highly conserved in the choline kinase family have been implicated by this structure study as either being located in the active site, and likely to be involved in substrate binding/catalysis, or stabilizing the conformation of catalytically important residues. Although the level of sequence identity varies among choline kinase family members, the high degree of conservation in the Brenner's motif and the choline kinase motif suggests that all choline kinases have a similar structure. The structure of CKA-2, therefore, could serve as a model for studies of other members of this large kinase family.

Given the importance of choline kinase in membrane biosynthesis and lipid second-messenger signaling, lit-

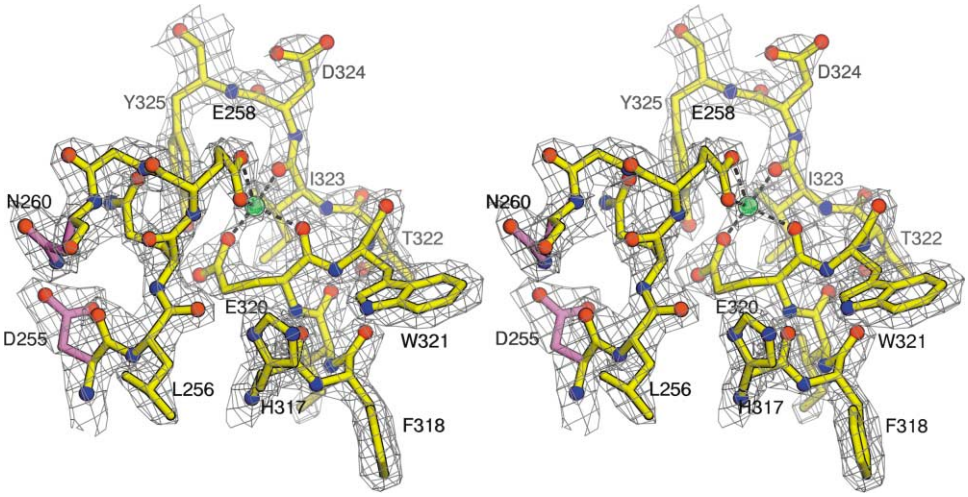

Figure 7. Experimental Electron Density Map of the Calcium Binding Site in CKA-2

A stereo view of an $m|F_o|\exp(i\phi^{mod})$ electron density map contoured at 1.25 σ and superimposed upon the refined model at the calcium binding site. m, figure of merit; $|F_o|$, observed amplitudes; ϕ^{mod}, MAD phases after density modification. The calcium ion is represented as the large green sphere coordinated (dashed lines) by the side chains of Glu258 and Glu320 and the main chain carbonyl of I323. Residues Asp255 and Asn260, which are proposed to be involved in Mg-ATP binding, are highlighted in magenta.

tle is known about the regulation of this enzyme. Although some regulation probably occurs at the level of gene expression (Aoyama et al., 2000), there is also likely to be regulation of activity of the enzyme itself. For example, yeast choline kinase is subject to phosphorylation and activation by protein kinase A (Yu et al., 2002). While we do not know whether CKA-2 is subject to similar regulation, the current structure will provide a model for studying the relationship of phosphorylation sites to the active site in yeast choline kinase to understand the mechanism of this type of regulation.

Choline kinase may also serve as a model for other choline-utilizing proteins. The cellular concentration of choline is balanced by uptake through membrane-associated choline transporters and consumption through choline kinase and, in cholinergic neurons, choline acetyltransferase. Certain cholinergic antagonists, such as hemicholinium-3, which inhibit choline uptake, also inhibit choline kinase (Ansell and Spanner, 1974). The availability of structural information on choline kinase provides a fresh opportunity to examine the effects of these drugs at a molecular level.

Experimental Procedures

Protein Expression and Purification

For the production of soluble *C. elegans* CKA-2 in *E. coli*, the coding region was inserted into a *pET43a*-derived plasmid (Novagen) to express as an N-terminal NusA fusion protein (Davis et al., 1999). A polyhistidine tag followed by a TEV protease recognition site was constructed between NusA and the choline kinase-coding region to facilitate protein purification. *E. coli* BL21(DE3) cells harboring the recombinant plasmid were grown in LB medium, and protein expression was induced with 0.4 mM IPTG at 20°C for 16 hr. The selenomethionine form of the protein was expressed in a previously described M9 minimal medium supplemented with 75 mg · l^{-1} selenomethionine.

All subsequent procedures were carried out at 4°C. Approximately 10 g of cell paste was resuspended in 100 ml of Buffer A (25 mM Tris-HCl [pH 8.0], 300 mM KCl, 5 mM β-mercaptoethanol, and 10 mg · ml^{-1} PMSF) with two Complete EDTA-free Protease Inhibitor Tablets from Roche. The cells were lysed by sonication, and the lysate was cleared by centrifugation. The supernatant was batch bound for one hour to 5 ml of Ni^{2+}-NTA resin (Qiagen) preequilibrated with Buffer A. The resin mixture was then poured into a column and washed with Buffer A and then with Buffer A containing 20 mM imidazole. The protein was eluted with Buffer A containing a 20–250 mM imidazole gradient. Fractions containing choline kinase were pooled and buffer-exchanged against Buffer A. To remove the N-terminal NusA and polyhistidine tags from the protein, we added His-tagged TEV protease at a 1:100 (protease to protein) mass ratio, and digestion was allowed to proceed for 16 hr. The digestion mixture was then loaded onto a second Ni^{2+}-NTA column equilibrated with Buffer A, and tag-free choline kinase was collected in the flow-through fraction. After buffer-exchanging the protein into Buffer C (25 mM Tris-HCl [pH 7.5], 1 mM EDTA, and 1 mM DTT), the protein was further purified by loading onto a 20 ml Source-Q anion exchange column (Pharmacia) preequilibrated with Buffer C and eluted with a 0–500 mM KCl gradient. The typical yield for the above scheme was 3 mg of pure choline kinase per liter of starting bacterial culture.

Crystallization and Data Collection

CKA-2 crystals were grown by the sitting drop vapor diffusion method. Initial drops containing 5 μl of 4 mg · ml^{-1} choline kinase in crystallization buffer (20 mM bicine [pH 8.0], 1 mM EGTA, and 2 mM β-mercaptoethanol), 0.5 μl of 2 M nondetergent sulfobetaine 201, and 2 μl of well solution (8%–10% PEG 2000, 200 mM calcium acetate, and 100 mM bicine [pH 8.0]) were allowed to equilibrate against 1 ml of well solution at 23°C. Crystals appeared overnight and grew to optimal size ($0.2 \times 0.3 \times 0.05$ mm^3) in 3 to 4 days. Before the crystals were flash-cooled, they were slowly transferred in three steps into a cryoprotecting solution containing 30% PEG 400, 20% ethylene glycol, 50 mM calcium acetate, and 100 mM bicine [pH 8.0]. All data were collected with crystals flashed-cooled in a cold nitrogen stream (110 K). Three-wavelength MAD data were collected at the SBC beamline 19-ID at the Advanced Photon Source and processed with the HKL2000 package (Otwinowski and Minor, 1997). Native data were collected at the DND beamline 5-ID at the Advanced Photon Source and processed with DENZO and SCALEPACK (Otwinowski and Minor, 1997) (Table 1).

Structure Determination and Refinement

Experimental phases were obtained by the MAD method (Hendrickson and Ogata, 1997) (Table 1). Six of the ten selenium sites expected

in the choline kinase monomer were found by the RANTAN heavy-atom search method as implemented in the CCP4 package (three of the remaining selenium sites were in the disordered N-terminal portion of the selenomethionine-substituted protein structure) (CCP4, 1994). Subsequently, heavy-atom parameters were refined, and MAD phases were calculated and improved by solvent flipping to 2.2 Å resolution with SHARP (de la Fortelle and Bricogne, 1998). The high quality of the experimental map (Figure 7) and the known selenium sites allowed unambiguous tracing of the protein main chain and most of the side chains with O (Jones et al., 1991). All refinement procedures were carried out in CNS, with progress measured by crossvalidation (R_{free} calculation) with an 8% randomly selected test set (Brünger et al., 1998). Initial refinement consisted of several iterations of conjugated gradient minimization, grouped B factor refinement, torsion angle dynamics simulated annealing by the maximum-likelihood target function with experimental phases as a prior phase distribution (MLHL), and model rebuilding in O. These were performed against the MAD data set collected at the peak wavelength. Once the structure was partially refined, it was used as a molecular replacement search model to obtain the native structure in the $P2_1$ space group. The molecular replacement was done with CNS.

All subsequent refinement was performed against a 2.0 Å native data set. Again, all progress was measured by crossvalidation with a new 8% randomly selected test set. Two monomers were placed into the monoclinic cell and oriented by rigid body refinement. Initial refinement was performed with data to 2.5 Å and strict noncrystallographic symmetry (NCS) imposed between the two monomers. Later rounds of refinement were performed on an NCS-restrained dimmer, with slowly decreasing weights for the main chain and side chain atoms and slowly increasing high-resolution data. This included rounds of selecting chemically reasonable water molecules in phase-combined σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ maps, conjugate gradient minimization, grouped atomic B factor refinement, and model rebuilding with experimental, σ_A-weighted (MLF target), and phase-combined σ_A-weighted (MLHL target) $2F_o - F_c$ and $F_o - F_c$ maps (Read, 1986). In the final rounds, all of the data from 100–2.0 Å were used with bulk solvent correction, no NCS, and individual atomic B factors refined. The final model includes residues 32–265 and residues 296–426 for subunit A, residues 40–265 and residues 296–426 for subunit B, one calcium ion for each subunit, and a total of 342 water molecules. Residues not included in the model are those that are not visible and assumed to be disordered.

Figure Preparation

Molscript (Kraulis, 1991), POVSCRIPT (http://people.brandeis.edu/~fenn/povscript), and POV-ray (http://www.povray.org) were used to produce Figures 2, 3, 4, 5, and 7. Grasp (Nicholls et al., 1991) was used to produce Figure 6.

Acknowledgments

We thank Jeanne Stuckey for maintaining the X-ray facility at the University of Michigan Medical School; Rongguang Zhang, Stephen Ginell, and Andrzej Joachimiak for access and help at the APS SBC beamline; Zdzislaw Wawrzak, Steven Weigand, and John Quintana for access and help at the APS DND beamline; Rowena Matthews for critically reading the manuscript; and Chong Yuan and the Xu lab for helpful discussions. This work was supported in part by NIH grants to Z.X. (R01-GM60997) and C.K. (R01-GM50610). Z.X. is a Pew Scholar in Biomedical Sciences.

Received: January 13, 2003
Revised: March 3, 2003
Accepted: March 27, 2003
Published: June 3, 2003

References

Al-Malki, A., Morby, A., and Harwood, J.L. (2000). Pea choline kinase: purification, properties and isolation of a cDNA. Biochem. Soc. Trans. *28*, 721–723.

Ansell, G.B., and Spanner, S.G. (1974). The inhibition of brain choline kinase by hemicholinium-3. J. Neurochem. *22*, 1153–1155.

Aoyama, C., Nakashima, K., Matsui, M., and Ishidate, K. (1998). Complementary DNA sequence for a 42 kDa rat kidney choline/ethanolamine kinase. Biochim. Biophys. Acta *1390*, 1–7.

Aoyama, C., Yamazaki, N., Terada, H., and Ishidate, K. (2000). Structure and characterization of the genes for murine choline/ethanolamine kinase isozymes alpha and beta. J. Lipid Res. *41*, 452–464.

Bean, B., and Tomasz, A. (1977). Choline metabolism in pneumococci. J. Bacteriol. *130*, 571–574.

Brenner, S. (1987). Phosphotransferase sequence homology. Nature *329*, 21.

Brünger, A.T., Adams, P.D., Clore, G.M., Delano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Burk, D.L., Hon, W.C., Leung, A.K., and Berghuis, A.M. (2001). Structural analyses of nucleotide binding to an aminoglycoside phosphotransferase. Biochemistry *40*, 8756–8764.

CCP4 (Collaborative Computational Project 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763.

Cuadrado, A., Carnero, A., Dolfi, F., Jimenez, B., and Lacal, J.C. (1993). Phosphorylcholine: a novel second messenger essential for mitogenic activity of growth factors. Oncogene *8*, 2959–2968.

Davis, G.D., Elisee, C., Newham, D.M., and Harrison, R.G. (1999). New fusion protein systems designed to give soluble expression in Escherichia coli. Biotechnol. Bioeng. *65*, 382–388.

Exton, J.H. (2000). Phospholipase D. Ann. NY Acad. Sci. *905*, 61–68.

Favelyukis, S., Till, J.H., Hubbard, S.R., and Miller, W.T. (2001). Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. Nat. Struct. Biol. *8*, 1058–1063.

de la Fortelle, E., and Bricogne, G. (1998). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472–494.

Gee, P., and Kent, C. (2003). Multiple isoforms of choline kinase from Caenorhabditis elegans: cloning, expression, purification, and characterization. Biochim. Biophys. Acta *1648*, 33–42.

Hendrickson, W.A., and Ogata, C.M. (1997). Phase determination from multiwavelength anomalous diffraction measurements. Methods Enzymol. *276*, 494–523.

Hernandez-Alcoceba, R., Saniger, L., Campos, J., Nunez, M.C., Khaless, F., Gallo, M.A., Espinosa, A., and Lacal, J.C. (1997). Choline kinase inhibitors as a novel approach for antiproliferative drug design. Oncogene *15*, 2289–2301.

Higgins, D., Thompson, J., Gibson, T., Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673–4680.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

Hon, W.C., McKay, G.A., Thompson, P.R., Sweet, R.M., Yang, D.S., Wright, G.D., and Berghuis, A.M. (1997). Structure of an enzyme required for aminoglycoside antibiotic resistance reveals homology to eukaryotic protein kinases. Cell *89*, 887–895.

Hosaka, K., Kodaki, T., and Yamashita, S. (1989). Cloning and characterization of the yeast CKI gene encoding choline kinase and its expression in Escherichia coli. J. Biol. Chem. *264*, 2053–2059.

Ishidate, K., Kihara, M., Tadokoro, K., and Nakazawa, Y. (1982). Induction of choline kinase by polycyclic aromatic hydrocarbons in rat liver. I. A comparison of choline kinases from normal and 3-methylcholanthrene-induced rat liver cytosol. Biochim. Biophys. Acta *713*, 94–102.

Ishidate, K., Nakagomi, K., and Nakazawa, Y. (1984). Complete purification of choline kinase from rat kidney and preparation of rabbit

antibody against rat kidney choline kinase. J. Biol. Chem. *259*, 14706–14710.

Ishidate, K., Iida, K., Tadokoro, K., and Nakazawa, Y. (1985). Evidence for the existence of multiple forms of choline (ethanolamine) kinase in rat tissues. Biochim. Biophys. Acta *833*, 1–8.

Janin, J., and Chothia, C. (1990). The structure of protein-protein recognition sites. J. Biol. Chem. *265*, 16027–16030.

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. A *47*, 110–119.

Kent, C. (1990). Regulation of phosphatidylcholine biosynthesis. Prog. Lipid Res. *29*, 87–105.

Kim, K.H., Voelker, D.R., Flocco, M.T., and Carman, G.M. (1998). Expression, purification, and characterization of choline kinase, product of the CKI gene from Saccharomyces cerevisiae. J. Biol. Chem. *273*, 6844–6852.

Kinoshita, K., Sadanami, K., Kidera, A., and Go, N. (1999). Structural motif of phosphate-binding site common to various protein superfamilies: all-against-all structural comparison of protein-mononucleotide complexes. Protein Eng. *12*, 11–14.

Knighton, D.R., Zheng, J.H., Ten Eyck, L.F., Ashford, V.A., Xuong, N.H., Taylor, S.S., and Sowadski, J.M. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science *253*, 407–414.

Kraulis, P.J. (1991). MOLSCRIPT: a program to reproduce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Lawrie, A.M., Noble, M.E., Tunnah, P., Brown, N.R., Johnson, L.N., and Endicott, J.A. (1997). Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. Nat. Struct. Biol. *4*, 796–801.

Longenecker, K.L., Roach, P.J., and Hurley, T.D. (1996). Three-dimensional structure of mammalian casein kinase I: molecular basis for phosphate recognition. J. Mol. Biol. *257*, 618–631.

Lykidis, A., Wang, J., Karim, M.A., and Jackowski, S. (2001). Overexpression of a mammalian ethanolamine-specific kinase accelerates the CDP-ethanolamine pathway. J. Biol. Chem. *276*, 2174–2179.

McKay, G.A., and Wright, G.D. (1995). Kinetic mechanism of aminoglycoside phosphotransferase type IIIa. Evidence for a Theorell-Chance mechanism. J. Biol. Chem. *270*, 24686–24692.

Monks, D.E., Goode, J.H., and Dewey, R.E. (1996). Characterization of soybean choline kinase cDNAs and their expression in yeast and Escherichia coli. Plant Physiol. *110*, 1197–1205.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins Struct. Funct. Genet. *11*, 281.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Pavlidis, P., Ramaswami, M., and Tanouye, M.A. (1994). The Drosophila easily shocked gene: a mutation in a phospholipid synthetic pathway causes seizure, neuronal failure, and paralysis. Cell *79*, 23–33.

Porter, T.J., and Kent, C. (1990). Purification and characterization of choline/ethanolamine kinase from rat liver. J. Biol. Chem. *265*, 414–422.

Poxton, I.R., and Leak, D.J. (1977). The biosynthesis of a choline nucleotide by a cell-free extract from Streptococcus pneumoniae. J. Gen. Microbiol. *100*, 23–29.

Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr. *A42*, 140–149.

Tuomanen, E. (1999). Molecular and cellular biology of pneumococcal infection. Curr. Opin. Microbiol. *2*, 35–39.

Uchida, T. (1994). Regulation of choline kinase R: analyses of alternatively spliced choline kinases and the promoter region. J. Biochem. *116*, 508–518.

Uchida, T. (1997). A novel high-molecular mass mammalian ethanolamine kinase. Biochim. Biophys. Acta *1349*, 13–24.

Uchida, T., and Yamashita, S. (1990). Purification and properties of choline kinase from rat brain. Biochim. Biophys. Acta *1043*, 281–288.

Uchida, T., and Yamashita, S. (1992). Molecular cloning, characterization, and expression in Escherichia coli of a cDNA encoding mammalian choline kinase. J. Biol. Chem. *267*, 10156–10162.

Weber, C.H., Park, Y.S., Sanker, S., Kent, C., and Ludwig, M.L. (1999). A prototypical cytidylyltransferase: CTP:glycerol-3-phosphate cytidylyltransferase from bacillus subtilis. Struct. Fold. Des. *7*, 1113–1124.

Whitehouse, S., Feramisco, J.R., Casnellie, J.E., Krebs, E.G., and Walsh, D.A. (1983). Studies on the kinetic mechanism of the catalytic subunit of the cAMP-dependent protein kinase. J. Biol. Chem. *258*, 3693–3701.

Wittenberg, J., and Kornberg, A. (1953). Choline phosphokinase. J. Biol. Chem. *202*, 431–444.

Yu, Y., Sreenivas, A., Ostrander, D.B., and Carman, G.M. (2002). Phosphorylation of Saccharomyces cerevisiae choline kinase on Ser30 and Ser85 by protein kinase A regulates phosphatidylcholine synthesis by the CDP-choline pathway. J. Biol. Chem. *277*, 34978–34986.

Accession Numbers

Atomic coordinates and structure factors have been deposited into the Protein Data Bank under accession code 1NW1.