Structural Basis for the Catalysis and Substrate Specificity of Homoserine Kinase†

S. Sri Krishna,§ Tianjun Zhou,§ Matthew Daugherty,‡ Andrei Osterman,‡ and Hong Zhang*,§

Department of Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390-9038, and Integrated Genomics, Inc., 2201 West Campbell Park Drive, Chicago, Illinois 60612

Received April 25, 2001; Revised Manuscript Received July 6, 2001

ABSTRACT: Homoserine kinase (HSK), the fourth enzyme in the aspartate pathway of amino acid biosynthesis, catalyzes the phosphorylation of L-homoserine (Hse) to L-homoserine phosphate, an intermediate in the production of L-threonine, L-isoleucine, and in higher plants, L-methionine. The high-resolution structures of *Methanococcus jannaschii* HSK ternary complexes with its amino acid substrate and ATP analogues have been determined by X-ray crystallography. These structures reveal the structural determinants of the tight and highly specific binding of Hse, which is coupled with local conformational changes that enforce the sequestration of the substrate. The δ-hydroxyl group of bound Hse is only 3.4 Å away from the γ-phosphate of the bound nucleotide, poised for the in-line attack at the γ-phosphorus. The bound nucleotides are flexible at the triphosphate tail. Nevertheless, a Mg^{2+} was located in one of the complexes that binds between the β- and γ-phosphates of the nucleotide with good ligand geometry and is coordinated by the side chain of Glu130. No strong nucleophile (base) can be located near the phosphoryl acceptor hydroxyl group. Therefore, we propose that the catalytic mechanism of HSK does not involve a catalytic base for activating the phosphoryl acceptor hydroxyl but instead is mediated via a transition state stabilization mechanism.

Homoserine kinase (HSK, EC 2.7.1.39)[1] catalyzes the following reaction: L-homoserine + ATP → *O*-phospho-L-homoserine + ADP. It occupies a key position in the bacterial threonine biosynthetic pathway as it performs the first reaction after the branching-point to methionine. The kinetic properties of homoserine kinases from several species, such as *Escherichia coli* (*1*, *2*), yeast (*3*), *Arabidopsis thaliana* (*4*), wheat germ (*5*), and *Methanococcus jannaschii* (*6*), have been characterized. Despite the high degree of similarities in sequences, significant variations in kinetic properties are observed among HSKs from different species. The enzymes from various species are shown to bind substrates mostly in a random order, with a slight preference of binding to ATP first (*2*, *6*). Wormser and Pardee first showed that in *E. coli* one of the ways threonine could regulate its own synthesis was by inhibiting the HSK activity (*7*). High concentrations of Hse also have an inhibitory effect on the *E. coli* HSK activity (*2*). However, in HSKs from plants and archaea, high concentrations of Hse do not alter the enzyme activity (*4*−*6*). It was also found that HSK from plants is not inhibited by threonine (*4*, *5*), which is probably a requirement for being a precursor of multiple amino acids biosynthesis pathways.

To identify functional groups that are involved in substrate recognition and catalysis, extensive biochemical and mutagenesis studies have been carried out on the *E. coli* HSK. Chemical modifications with group specific reagents have indicated the involvement of a histydinyl and a lysyl residue at or near the enzyme's active site (*8*). Mutation of His139 to Leu resulted in a diminished kinase activity but elevated ATPase activity, suggesting that His139 is close to the active site (*9*). Mutagenesis and kinetics measurements of enzyme activity with α-modified homoserine analogues further suggested that Arg234 interacts with the carboxylate group of the bound amino acid (*9*).

HSK belongs to a large superfamily of enzymes, the GHMP kinases (*10*) (Figure 1). Other members in this superfamily include galactokinase, mevalonate kinase, phosphomevalonate kinase (*11*), mevalonate diphosphate decarboxylase (*12*), 4-diphosphocytidyl-2C-methyl-D-erythritol kinase (*13*), and archaeal shikimate kinase (*6*). We have recently solved the crystal structures of homoserine kinase from *M. jannaschii* and its complex with ADP (*14*), which are the first three-dimensional structures of the GHMP superfamily published. These structures reveal a unique kinase fold and a novel nucleotide-binding mode that are conserved among members of the GHMP kinase superfamily. The bound ADP adopts a rare *syn* conformation, and its location in the structure reveals the substrate binding and active site. The most conserved sequence motif among GHMP kinases, with a consensus PX_3GSSAA, forms a novel phosphate-binding loop and interacts extensively with the α- and β-phosphates of the nucleotide. The cavity adjacent to the ADP is well suited for binding of homoserine.

† The atomic coordinates and the structure factors of the homoserine kinase complexes have been deposited in the Protein Data Bank (PDB; http://www.rcsb.org/pdb/), entry codes 1h72(Hse+AMPPNP), 1h73 (Thr+AMPPNP), and 1h74 (Ile+ATP-γ-S).

* Corresponding author. E-mail: zhang@chop.swmed.edu. Phone: 214-648-9299. Fax: 214-648-9099.

§ University of Texas Southwestern Medical Center.

‡ Integrated Genomics, Inc.

[1] Abbreviations: HSK, homoserine kinase; MK, mevalonate kinase; Hse, homoserine; HEPES, *N*-(2-hydroxyethyl)-piperazine-*N*-2-ethanesulfonic; AMPPNP, 5′-adenylylimidotriphosphate; ATP-γ-S, adenosine-5′-diphosphate monothiophosphate; *M. jannaschii*, *Methanococcus jannaschii*; NMP kinase, nucleoside monophosphate kinase; c.c., correlation coefficient; rmsd, root-mean-square deviation.

10.1021/bi010851z CCC: $20.00
Published on Web 08/16/2001

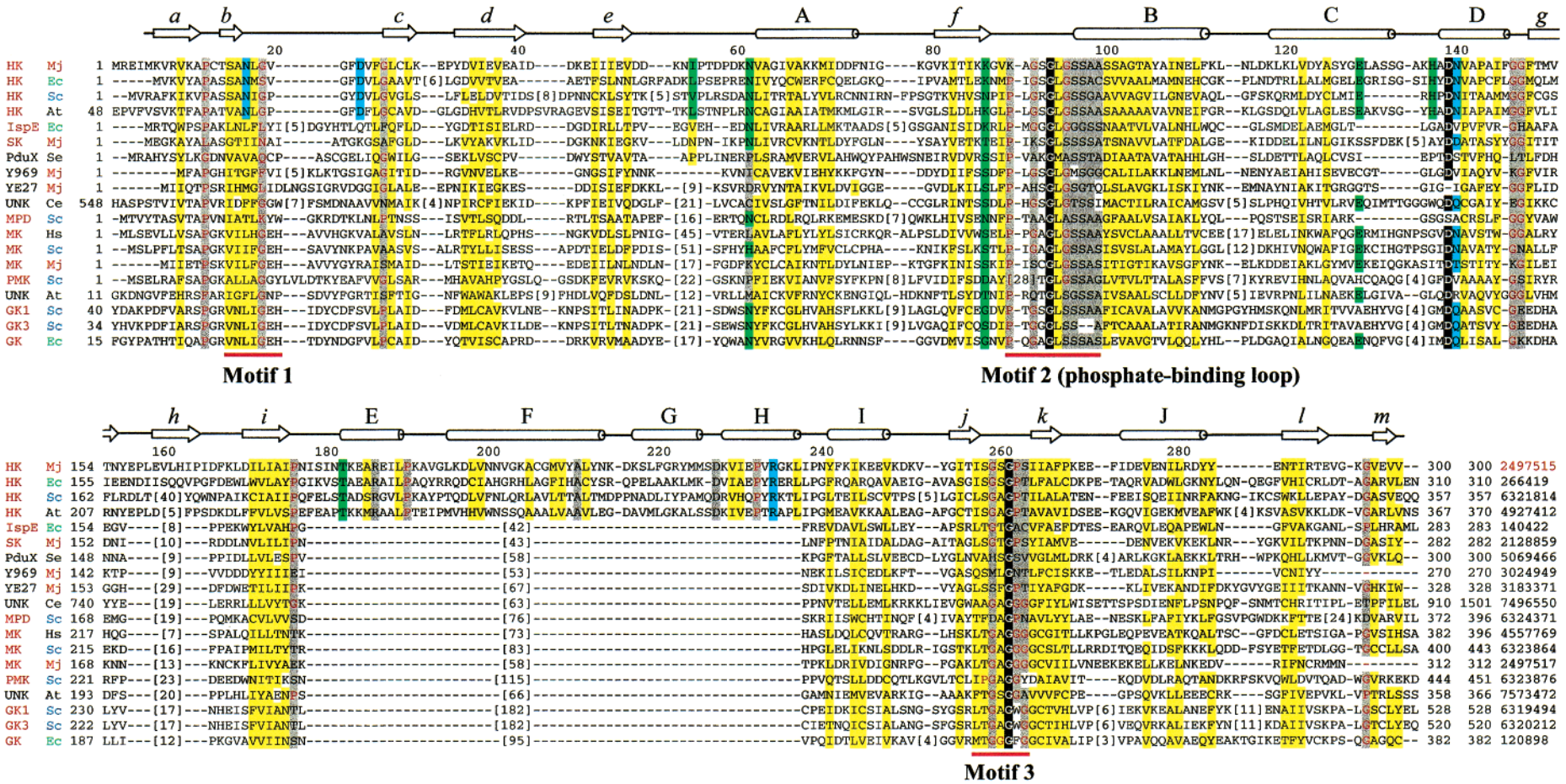

FIGURE 1: Multiple alignment of the GHMP superfamily. The three highly conserved motifs as defined by Bork et al. (1993) (*10*) are underlined and marked as Motif 1−3. The first column lists the protein (gene) names, and the second column lists abbreviated species names. Protein name abbreviations are HK, homoserine kinase; IspE, 4-diphosphocytidyl-2C-methyl-D-erythritol kinase (*13*); SK, shikimate kinase (*6*); UNK, unknown; MPD, mevalonate pyrophosphate decarboxylase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; GK, galactokinase. Species name abbreviations are Mj, *Methanococcus jannaschii*; Ec, *Escherichia coli*; Sc, *Saccharomyces cerevisiae*; At, *Arabidopsis thaliana*; Se, *Samonella enterica*; Ce, *Caenorhabditis elegans*; Hs, *Homo sapiens*. To the right, the numbers indicate the positions of the first residue of the aligned region in the respective protein sequence. Numbers in parentheses indicate how many residues in an insertion are not shown. The numbers at the end of each sequence indicate the last residue number in the aligned region followed by the number of total residues in the protein. The NCBI gene identification number (gi) is at end of the alignment. The secondary structure elements in the HSK crystal structure are indicated above the sequences with arrows representing β strands and cylinders denoting α helices. The invariant residues are highlighted in black and are shown in white letters. The mostly conserved residues are highlighted in gray. Uncharged residues in mainly hydrophobic sites are highlighted in yellow. Residues that interact with MgATP are highlighted in green; residues that interact with homoserine are highlighted in light blue.

Table 1: Crystal Data and Refinement Statistics

complexes	HK+Hse+AMPPNP	HK+Thr+AMPPNP	HK+Ile+ATP-γ-S
space group	$P4_32_12$	$P4_32_12$	$P2_1$
cell dimensions	$a = b = 87.99$ Å, $c = 99.30$ Å	$a = b = 88.64$ Å, $c = 99.55$ Å	$a = 62.42$ Å, $b = 129.30$ Å, $c = 109.93$ Å, $\beta = 105.71°$
X-ray source	APS SBC-19BM	Rigaku RU-H3R	APS SBC-19BM
resolution (Å)	50−1.8	25−2.0	50−1.9
total measurements	206 493	235 303	533 801
unique reflections	36 728	27 305	131 766
completeness (last shell)(%)	99.8 (99.9)	99.8 (98.7)	99.3 (94.9)
I/σ(last shell)	21.1 (2.0)	45.8 (6.8)	26.9 (1.3)
R_{sym}[a]	0.089	0.043	0.072
refinement statistics			
R_{work} (R_{free})[b]	0.177 (0.207)	0.178 (0.222)	0.210 (0.238)
no. of protein atoms	2269	2269	9076
no. of water molecules	308	249	442
no. of Mg^{2+}	0	0	1
no. of substrate atoms	47	38	117
rmsd bond	0.016 Å	0.014 Å	0.014 Å
rmsd bond angle	1.75°	1.87°	1.62°
average B (Å^2)			
protein	24.99	26.20	50.65
solvent	41.86	37.91	51.05
nucleotide:	47.38	51.04	69.43
base	21.56	28.73	55.31
ribose	44.04	47.94	68.10
phosphates	71.39	71.68	84.74
amino acid substrate	29.60	39.12	69.45

[a] $R_{sym} = \sum_{hkl}[(\sum_j|I_j - \langle I\rangle|)/\sum_j|I_j]$. Because of a problem in the data processing software dealing with the unique detector configuration at APS 19BM beamline, the data collected at this beamline had a systematically high R_{sym}. However, this does not appear to affect refinement significantly. [b] $R_{work} = \sum_{hkl}|F_o - F_c|/\sum_{hkl}|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. R_{free} is the R factor calculated for a randomly selected 5% of the reflections that were omitted from the refinement.

To further understand the active site and catalytic mechanism of HSK, we have determined the crystal structures of *M. jannaschii* HSK complexed with ATP analogue and substrate homoserine as well as with inhibitors threonine and isoleucine. These structures reveal highly specific interactions between the enzyme and homoserine, coupled with local conformational changes that effectively sequester the amino acid substrate. The bound ATP analogues display a high degree of flexibility at the triphosphate tails. In the HSK ternary complex with Hse and ATP analogue AMPPNP, the δ-hydroxyl of Hse is only 3.4 Å away from the phosphorus of the γ-phosphate. Our structures suggest a mechanism in which the Hse hydroxyl group directly attacks the γ-phosphate of ATP in the absence of a general base provided by a protein residue.

EXPERIMENTAL PROCEDURES

Crystallization and Data Collection. The *M. jannaschii* HSK was expressed and purified as described before (*6*, *14*). Cocrystals of HSK−substrate complexes were grown by the hanging drop vapor-diffusion method in conditions similar to that of the uncomplexed proteins (*14*). Prior to crystallization, 15 mg/mL HSK, in 50 mM Hepes, pH 7.2, 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT, was incubated with various substrates at a final substrate concentration of 5 mM. 2 μL HSK−substrate complex solution was then mixed with an equal volume of the reservoir solution which contained 0.1 M Tris, pH 8.0−8.5, 0.2 M sodium acetate, and 20−25% PEG4000, and equilibrated against the reservoir. Diffraction data for cocrystals of HSK+Thr+AMPPNP were collected on a RAXIS−IV image plate detector equipped with Osmic mirrors at a temperature of 100 K. Data for the cocrystals of HSK+Ile+ATP-γ-S and HSK+Hse+AMPPNP were collected at the Advanced Photon Source, Argonne National Laboratory, beamline 19BM. The Ile+ATP-γ-S complex crystals are isomorphous to the native crystal and belong to the space group $P2_1$ with cell dimensions of $a = 61.99$ Å, $b = 129.72$ Å, $c = 109.42$ Å, and $\beta = 105.57°$. The Thr+AMPPNP and Hse+AMPPNP complexes were crystallized in the space group $P4_32_12$, with cell dimensions of $a = b = 87.31$ Å and $c = 99.57$ Å. While there are four independent HSK monomers in the asymmetric unit in the monoclinic crystals, there is only one HSK monomer in the asymmetric unit of the tetragonal crystals. All diffraction data were processed with DENZO and SCALEPACK using the HKL2000 package (*15*). The data statistics for these crystals are listed in Table 1.

Structure Determination and Refinement. Since HSK+Ile+ATP-γ-S cocrystals are isomorphous to the uncomplexed HSK crystal, the refined uncomplexed HSK structure was used as the starting model for the refinement of HK+Ile+ATP-γ-S structure using the CNS package (*16*). The substrates were built into the difference densities at a later stage of the refinement before solvent molecules were included.

Initial phases for the HSK+Thr+AMPPNP complex were solved by the molecular replacement method using the program AMoRe (*17*) of the CCP4 package (*18*). A monomer of the uncomplexed HSK structure was used as the search model. The rotation function search gave a peak of 8.2σ with a correlation coefficient (c.c.) of 0.175 for data between 20 and 4 Å. Translation searches carried out in both $P4_12_12$ and $P4_32_12$ space groups indicated that the correct space group is $P4_32_12$. The highest peak in the translation function was 15.7σ with a c.c. of 0.577 and R_{factor} of 0.42 for data from 20 to 3.5 Å. Rigid body refinement in CNS (*16*) yielded an

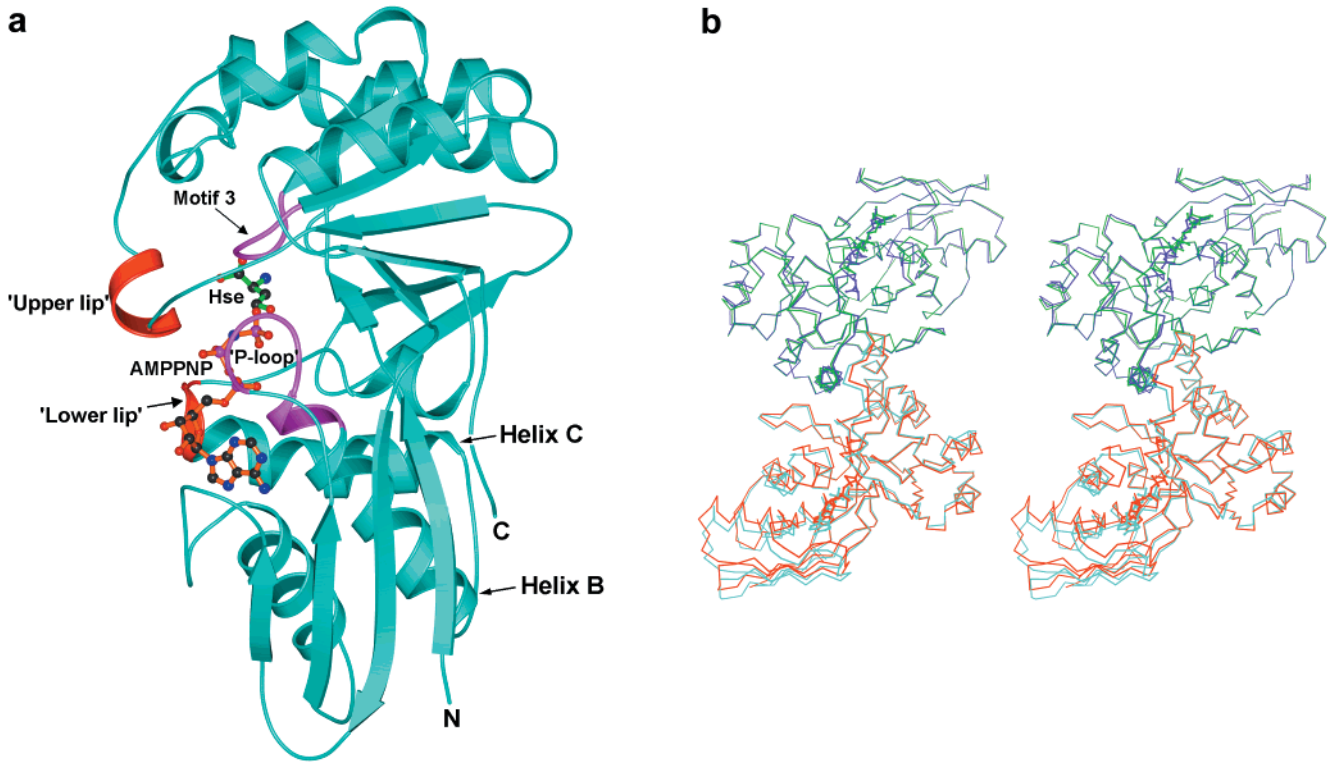

FIGURE 2: (a) The structure of HK complexed with Hse and AMPPNP. The protein is represented as a ribbon and the substrates are represented as ball-and-stick. The phosphate binding loop ("P-loop") and the motif 3 are shown in purple and the flexible "lips" in red. (b) Stereoview of the C_α superposition of the physiological dimer of HK from the $P2_1$ space group HK+ADP complex and the $P4_32_12$ space group HK+Hse+AMPPNP complex. The two subunits of the enzyme belonging to the $P2_1$ space group are colored green and cyan, respectively. The molecule in the $P4_32_12$ cell is colored purple and its symmetry mate is in red.

R-factor of 0.347 for data from 20 to 2.0 Å. Subsequent refinement and model building was carried out using the packages CNS and O (*15*, *19*). The refined HSK+Thr+AMPPNP structure model, excluding substrates and solvents, was then used as the initial model for the refinement of the HSK+Hse+AMPPNP complex structure. The refinement statistics of all three complexes are listed in Table 1.

RESULTS

The Overall Structures of HSK−Substrate Complexes. The monomer structures in the different HSK−substrate complexes are very similar to that of the uncomplexed enzyme (Figure 2a). The root-mean-square deviations (rmsd) between C_α positions among various monomers range from 0.4 to 0.7 Å. The tetragonal structures have lower overall B-factors (~25 Å^2) as compared to the monoclinic structures (~50 Å^2). The intermolecular packing in the tetragonal space group is very different from that in the monoclinic lattice, but the interface of the dimerization is mostly preserved. This packing stabilizes all regions of the molecule. In contrast, in the $P2_1$ structure, there are four crystallographically independent monomers in the asymmetric unit, and the N-terminal domain of monomer B is largely disordered due to the lack of packing interactions. This is likely to be the main cause for the high overall B-factors of the $P2_1$ crystals. As a result, the two tetragonal structures were refined to a better R_{work} and R_{free} values while maintaining comparable geometric restraints as the monoclinic structure (Table 1).

Although the monomer structure is very similar among various HSK−substrate complexes in the two different lattices, there is a substantial difference in the relative positioning of the two subunits in the dimer (Figure 2b). When one subunit is superimposed by least-squares fitting, the second subunit shows a slight rotation (~4°) around the dyad. As a result, the farthest region from the dyad in the second subunit shifts about 4 Å between the two lattices. We also located a Tris molecule in the structure of HSK+Hse+AMPPNP.

The Homoserine Binding Site. The homoserine density is well defined in the HSK+Hse+AMPPNP complex (Figure 3a). The average B for Hse is 29.6 Å^2, comparable with that of the surrounding protein side chain atoms. The bound homoserine is located in a cavity adjacent to the ATP binding site and forms extensive and highly specific interactions with the protein side chains (Figure 3b). The α-carboxylate of the bound Hse forms a salt bridge with the side chain of Arg235 in a very favorable hydrogen bonding orientation. The amino group of the Hse interacts with the side chains of Asp140 and Asp23. Unexpectedly, Asp140 (aligned with Asp204 of human mevalonate kinase (MK)) does not make contact with the δ-OH of Hse. The OD group of Asp140 is more than 5 Å away from the Hse hydroxyl.

The complex structure also revealed that Asn17, whose main chain conformation is in the disallowed region in the Ramanchandran plot (*20*) ($\phi = 79.2°$, $\varphi = -60.3°$), is in a perfect position to form hydrogen bonds with both carboxylate and amino group of Hse (Figure 3b). The strained conformation of Asn17 is unchanged in both uncomplexed and complexed structures. Another specific interaction between the enzyme and Hse is a hydrogen bond formed

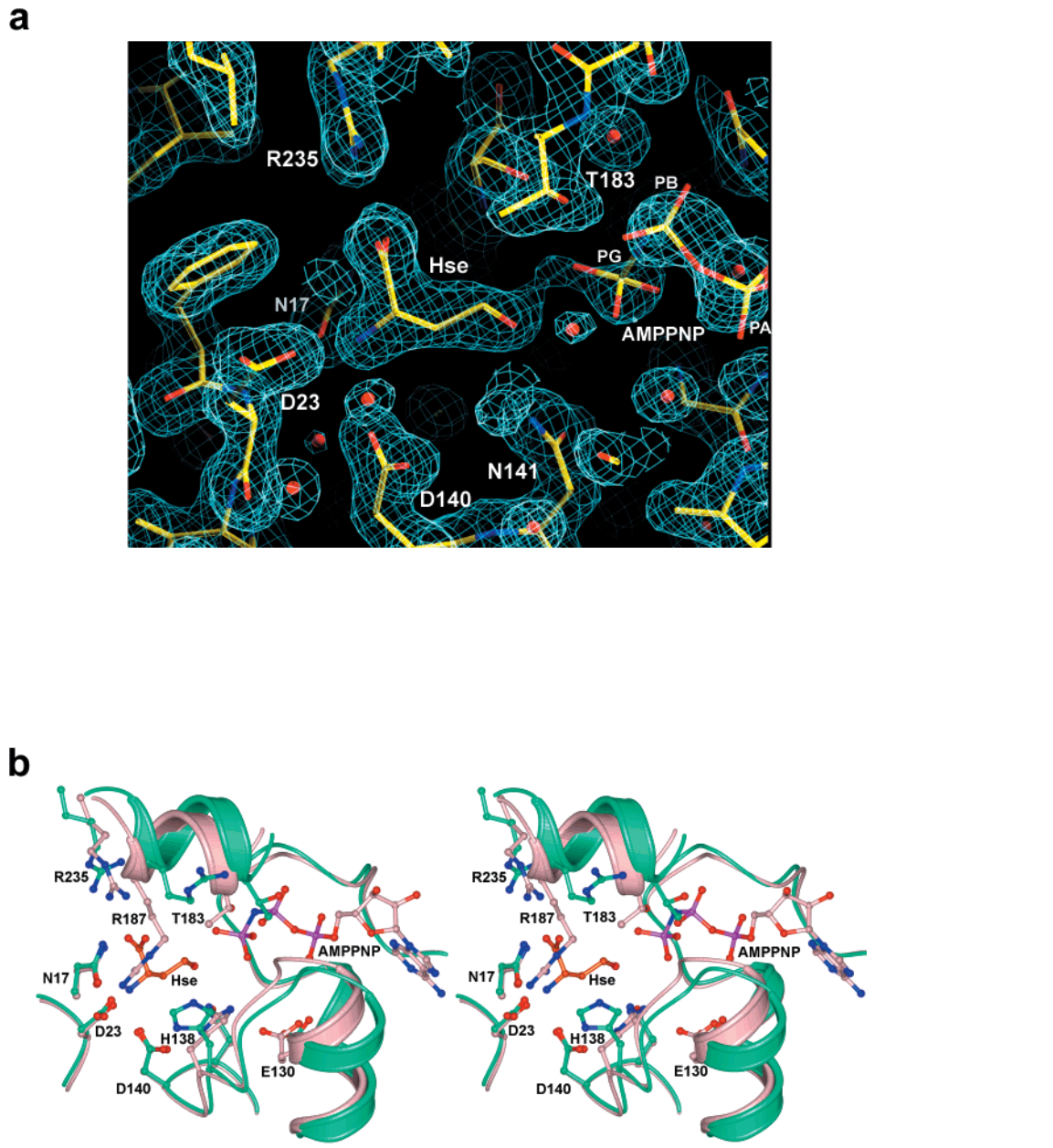

FIGURE 3: (a) Electron density around Hse binding site in HK+Hse+AMPPNP structure. The density was contoured at 1.2σ using the program O (*23*). The Hse is located near the ATP binding site and interacts extensively with the surrounding side chains. (b) Stereoview of the superposition of the apo-HK and HK+Hse+AMPPNP complex at the active site, showing the conformational changes induced by substrate binding and the interactions of bound Hse with protein residues. The apo-HK (green) was aligned on the Hse+AMPPNP (cream) using the program ALIGN (*23*).

between the Hse δ-OH group and the Asn141 side chain amide (distance 2.78 Å). This interaction between HSK and Hse side chain provides a discriminating factor that distinguishes Hse from other amino acids and appears to be necessary to anchor the bound amino acid. In the HSK+Ile+ATP-γ-S complex structure, lack of this interaction results in the disorder of bound Ile as reflected in the high B factors (69.4 Å^2, Table 1) and poor density. Steric hindrance and unfavorable hydrophilic environment of the binding pocket may also contribute to the disorder of the Ile. In the HSK+Thr+AMPPNP complex, the side chain hydroxyl OG of the bound Thr interacts with Asn141 weakly (distance ~ 3.5 Å), and the bound Thr is ordered but with somewhat higher B factors (average 39.1 Å^2) than the surrounding protein atoms (26.2 Å^2). The Asn residue at this position is also conserved in the mevalonate kinase family. In galactokinases, the equivalent residue is a Gln (Figure 1). This conservative replacement suggests that an Asn or a Gln residue at this position may play the same functional role of correctly orienting the phosphoryl acceptor hydroxyl groups in GHMP kinases.

In addition to these specific interactions, drastic local conformational changes occur upon homoserine binding: the "upper lip" helix (residues 181−189) near the active site closes down toward the substrate, and the side chain of Arg187 reorients by about 180°. As a result, the tip of Arg187 side chain moves by ~10 Å, now forming a salt bridge with Asp23, thus completely shielding the Hse from exposure to the solvent (Figure 3b). Upon Hse binding, the ion pair formed between Hse carboxylate and Arg235 becomes completely buried (Figure 3b). The two distinct conformations of Arg187 side chain appear strongly cor-

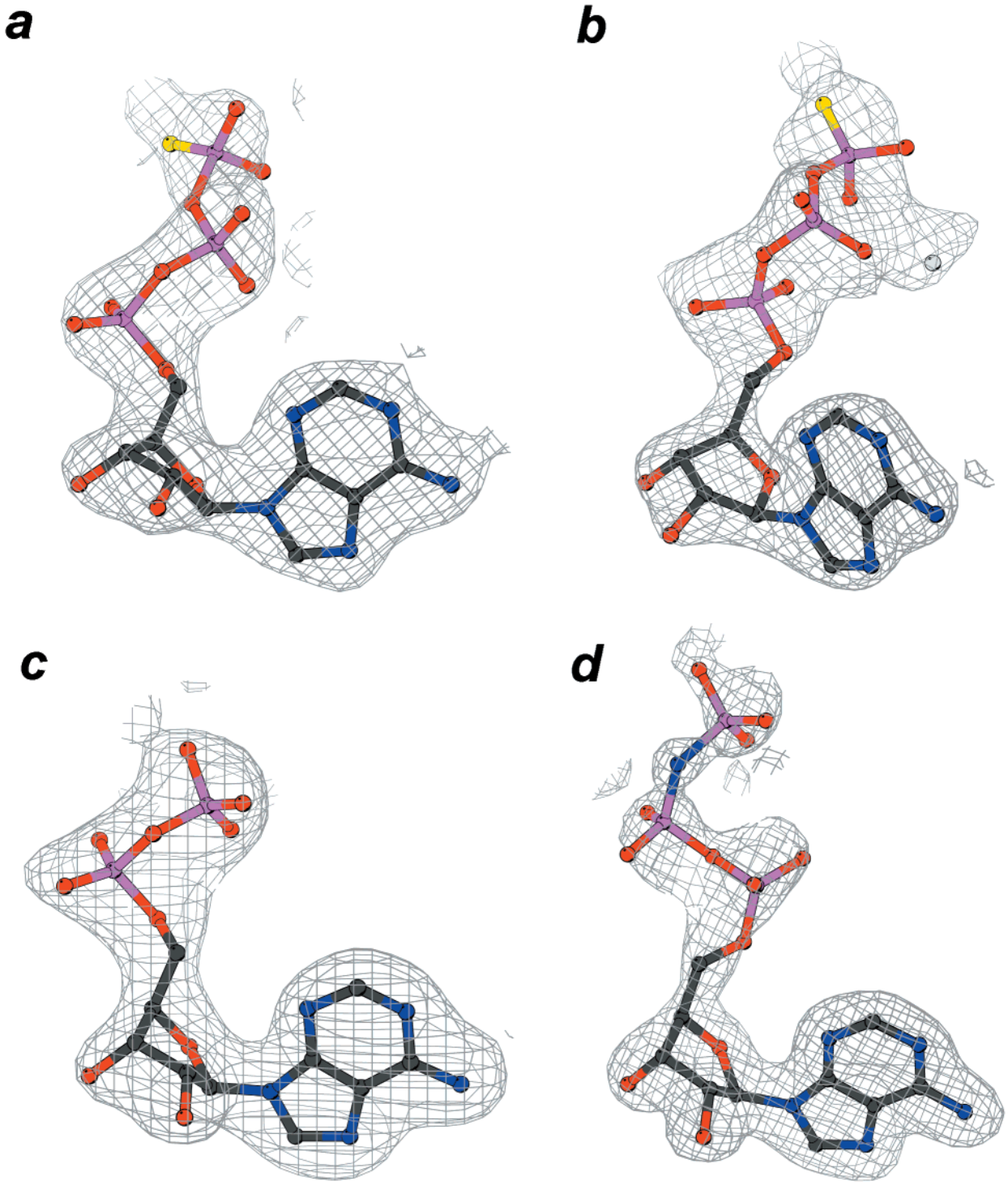

FIGURE 4: Representative electron densities of the bound nucleotide substrates. (a) ATP-γ-S in chain D of the HK+Ile+ATP-γ-S complex, (b) ATP-γ-S and Mg^{2+} in chain B, and (c) ADP in chain C; (d) AMPPNP in the HK+Hse+AMPPNP complex. All the densities were contoured at 1σ using the program Bobscript (*24*).

related with the binding of Hse or other amino acid substrates. In the apo-enzyme, it adopts the "open" conformation. In the Thr+AMPPNP and Hse+AMPPNP complex, where the Hse or Thr appear to have a full occupancy, Arg187 adopts the "closed" conformation. In the Ile+ATP-γ-S complex, Ile has only a partial occupancy, and the Arg187 side chain is largely disordered indicating alternative conformations.

The ATP and Mg^{2+} Binding Site. Unexpectedly, the triphosphate tails of the bound ATP analogues in different HSK−substrate complexes are rather flexible. In the cocrystal of HSK+Ile+ATP-γ-S, there are four independent HK monomers in the asymmetric unit. Monomers A and B make up one physiologically relevant HSK dimer, while monomers C and D compose another. In monomers B and D, the densities for the complete ATP-γ-S are clearly visible (Figure 4a,b), although the B factors of the nucleotide are much higher at the ribose and the triphosphate region as compared to that of the base. The conformations of the ATP-γ-S in monomers B and D are very similar. In chain B, a Mg^{2+} ion

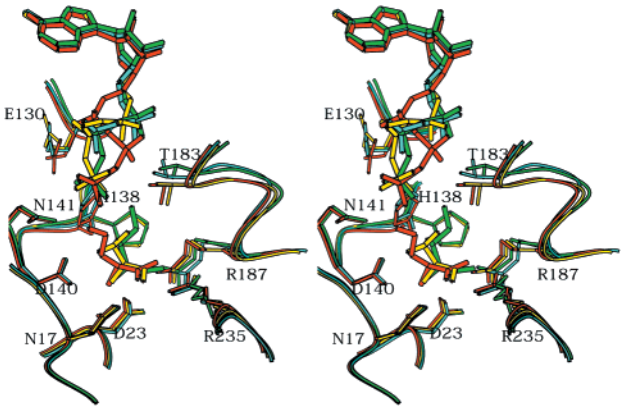

FIGURE 5: Stereoview of the superposition of various HK−substrate complexes around substrate binding sites, illustrating the conformational flexibility of the bound substrates and the interactions of the bound substrates with the protein side chains. Chains C and D of the HK+Ile+ATP-γ-S complex are colored cyan and green, respectively; HK+Hse+AMPPNP is colored red and HK+Thr+AMPPNP is in yellow.

can be located coordinating the β- and γ-phosphate (Figure 4b). Although the B-factor of this Mg^{2+} ion is high (87.7 Å^2), it is comparable to that of the phosphates in the nucleotide (Table 1) and has good ligand geometry. Glu130 is found to coordinate this Mg^{2+}. However, no Mg^{2+} was found in monomer D at this position (Figure 4a). Surprisingly, in monomers A and C of the same crystal, the γ-phosphate of the nucleotide is largely disordered. We were only able to model an ADP molecule in the density (Figure 4c). The ADP moiety of the bound ATP-γ-S in all four monomers in the monoclinic Ile+ATP-γ-S complex adopts essentially the same conformation as the ADP in the HSK+ADP complex structure and occupies the same binding site (*14*).

The flexibility of the nucleotide conformation is also profound in the HSK+AMPPNP complexed with either Hse or Thr. Since AMPPNP is more resistant to hydrolysis than ATP-γ-S, we assume that a majority of the AMPPNP molecules were not hydrolyzed. However, the density for the γ-phosphate is much weaker as compared to the other parts of the nucleotide (Figure 4d). The B-factors increase from mid-20s of the base, to mid-40s of the ribose, and up to 60s of the α-phosphate and 70s of the β-and γ-phosphates (Table 1), indicating increased flexibility along the nucleotide. Additionally, the AMPPNP molecule appears to adopt at least two alternative conformations, one is similar to that of ATP-γ-S in the HSK+Ile+ATP-γ-S complex, the other deviates at the α- and β-phosphate positions (Figure 5). The poor density and high B-factor of the amide link between β- and γ-phosphates of AMPPNP indicates that this part of the molecule is particularly flexible.

Inspection of the substrate binding cavities of HSK complexed with its substrates reveals that Hse molecule completely fills its binding pocket. However, the nucleotide-binding pocket appears to be larger than the ATP molecule, and thus allowing the latter to adopt multiple conformations and to be flexible. It appears that the nucleotide is firmly anchored by the specific interactions between the base and protein atoms, as well as by the electrostatic field of the anion hole created by the phosphate-binding loop (the PX3GSSAA motif) and the positive dipole of helix B (Figure 2a). These interactions provide a specific and tight binding and at the same time allowing a certain degree of flexibility.

Conformational Changes around the Active Site. Comparison of the structures of the HSK−substrate ternary complexes with the uncomplexed HSK reveals substantial conformational changes around the active site (Figure 3b). As mentioned before, the flexible "upper lip", formed by a short helix composed of residues 181−189, closes down toward both the nucleotide and amino acid substrates (Figures 2a, 3b). At the same time, the side chain of Arg187 undergoes a large conformational change to form a salt bridge with Asp23, thus enforcing the binding of Hse. Coincidently, the Thr183 side chain also moves by ∼3.5 Å and now directly interacts with the β-phosphate. The conformation of this region is better ordered in these complexes than in the apo-enzyme. The details of the interactions between bound substrates and protein are shown in Figure 6.

In the monoclinic HSK+Ile+ATP-γ-S complex structure, the "lower lip" near the active site (residues 131−137) is largely disordered in chains B and D but is better ordered in chains A and C and in the tetragonal structures. The side chain of His138, which has been implicated in the active site by the mutagenesis studies of the *E. coli* enzyme, is flexible and adopts at least two conformations: a "close-in" position as in chains A and C and an "open-up" position as in chains B and D (Figure 5). In the close-in position, His138 interacts with the α- and β-phosphates through a solvent molecule (data not shown).

DISCUSSION

Hse Substrate Specificity and Identification of Active Site Residues. So far only limited biochemical and mutagenesis data are available on the potential functional groups of GHMP kinases. They were mostly done on the *E. coli* HSK and human mevalonate kinase (MK) (*8*, *9*, *21*, *22*). Multiple alignment of all members in the GHMP kinase superfamily helps to interpret these data (Figure 1). Highly conserved residues in all members of GHMP kinase superfamily are likely to participate in the common steps of catalysis, while residues that are conserved within one family but are different in others are candidates for specific interactions with the substrates. Studies of the human MK suggest that Glu193 is directly involved in MgATP binding, while Asp204 is critical for the catalysis but not directly involved in MgATP binding (*21*). It was thus proposed that the Asp204 in human MK might serve as the catalytic base that facilitates deprotonation of the phosphoryl receptor hydroxyl on mevalonate (*21*). Glu193 and Asp204 in human MK (aligned with Glu130 and Asp140 in *M. jannaschii* HSK, respectively) are the two most conserved functional residues throughout the superfamily, except for the phosphate-binding motifs. The high-resolution crystal structures of HSK−substrate ternary complexes presented in this work provide direct visualization of how these groups interact with the substrates. Glu130 was found to interact with the Mg^{2+} that binds between the β- and the γ-phosphates of the nucleotide, thereby confirming its functional assignment in the previous studies (*21*). However, Asp140 is found to form a salt bridge with the positively charged amino group of the bound amino acid rather than interacting with the δ-OH of Hse. The distance between the oxygen of the Asp140 side chain and δ-OH of

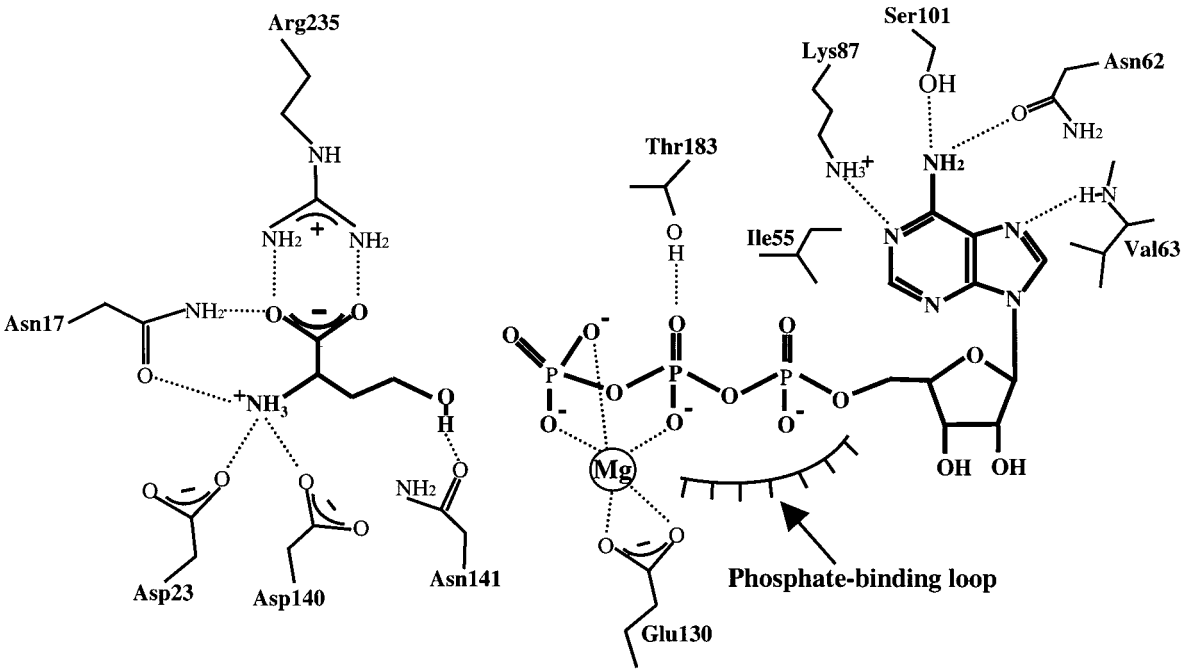

FIGURE 6: Summary of the interactions between HSK and the substrates. Dashed lines indicate the hydrogen bonds.

Hse is greater than 5 Å, well beyond the range of hydrogen bond interaction. Instead, Asn141 is seen to interact with the δ-OH of Hse in the HSK+Hse+AMPPNP complex. The positioning of Asn141 side chain is optimal for interacting with the δ-hydroxyl of a homoserine but not with any other amino acid. The conservation pattern at this position suggests that the same Asn in mevalonate kinase and the similar Gln in galactokinase may play a similar role of coordinating phosphoryl acceptor hydroxyl for the in-line attack at the γ-phosphate of ATP. It needs to be pointed out that HSK and MK are very divergent in sequence (~20% identity), and there are likely structural variations at the active site. It is possible that in MK, Asp204 may interact with mevalonate hydroxyl as a catalytic base. Structures of the MK−substrate complex will further clarify this issue.

Mutagenesis and biochemical studies on the *E. coli* HSK indicate that Arg234 (equivalent to Arg235 of *M. jannaschii* HSK), which is specifically conserved only in HSK, interacts with the carboxylate of Hse (*9*). The structures of the complexes confirmed this assessment of the role of Arg235. Several additional residues are revealed in the current study that specifically interact with the bound Hse. Asn17, with its main chain adopting a conformation that is in the disallowed region of the Ramachandran Plot (*20*), interacts with both the amino and carboxylate of the Hse backbone. Asp23 and Asp140 were found to interact with the amino group of Hse. Arg187 enforces the sequestration of Hse by a change in side chain conformation and forms a salt bridge with Asp23.

The most unexpected observations are the high flexibilities of the nucleotide phosphates, and as a result, the disorder of the Mg^{2+} ion. The structures of HSK complexed with different amino acids, ATP analogues, and crystallized in different space groups provided a population of conformations of the enzyme active sites and substrates/inhibitors (Figure 5). His138 is found to be in the proximity of the active site and adopts multiple conformations. In its "close-in" conformation, His138 side chain interacts with the nucleotide phosphates through a solvent molecule as observed in monomers A and C of the HSK+Ile+ATP-γ-S complex. Mutation of H139 to Leu in the *E. coli* enzyme resulted in a diminished kinase activity but a drastically enhanced ATPase activity. Whether the same effects would be observed for His138 of *M. jannaschii* HSK remains to be determined. It should be noted that the His138 is conserved only in HSKs but not in other GHMP enzymes, and the sequences between residues 131 and 137 are very diverse and are of different lengths even among HSK family members. Therefore, the effects of mutation in this region may have different consequences in HSKs from different organisms.

Proposed Catalytic Mechanism. So far, two major mechanisms have been proposed for kinase-catalyzed reactions. In a large group of nucleotide triphosphate phosphotransferases, which include phosphofructokinase (*25*) and ribokinase (*26*), an aspartic acid residue acts as a base to encourage deprotonation of the pertinent sugar hydroxyl group and so as to activate it for nucleophilic attack on the γ-phosphate of ATP (*27*, *28*). An alternate mechanism has been proposed for several structurally related phosphotransferases, such as NMP kinases, G proteins, and fructose 6-P,2-kinase (summarized in ref *29*), in which no catalytic base is involved and the transition state is stabilized by an anion hole created by the combination of main chain amide nitrogen's and conserved positively charged side chains. The mechanism of HSK appears to fall in the second category, even though the general fold of HSK is different as compared to the other kinases.

Combining the observations of various HSK−substrate complexes and previous biochemical and mutagenesis studies, we propose the following mechanisms for the HSK catalyzed reaction. The specific binding of Hse and ATP with the enzyme brings the two substrates close to each other. Mg^{2+} coordinated by Glu130 activates the γ-phosphate of ATP. The δ-OH group of Hse directly attacks the phosphor of the γ-phosphate. In the absence of a catalytic base, the

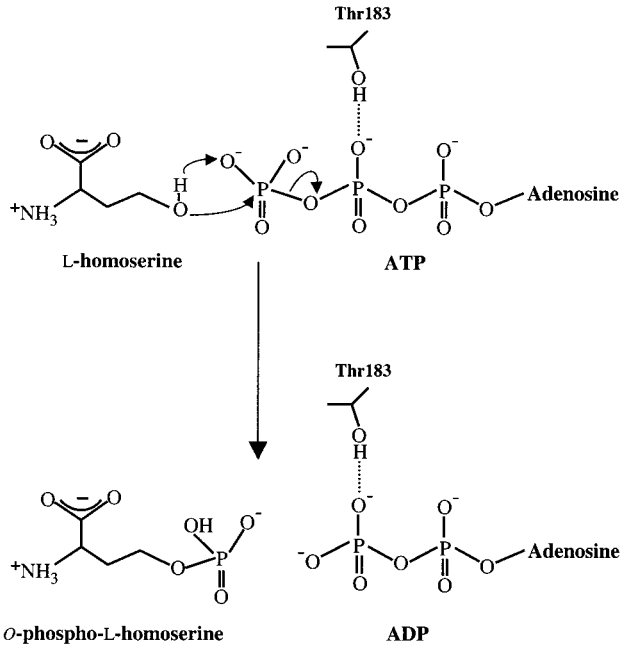

FIGURE 7: Proposed catalytic mechanism for HSK.

deprotonation of the δ-OH group of Hse may be achieved by the interaction with the γ-phosphate. Since the pK_a of protonating a -PO_3^{2-} to -HPO_3^- is around 7, the substrate binding induced protein conformational changes at the active site may promote such a change, stabilizing the transition state of the reaction and thus facilitating the phosphoryl transfer. Additionally, Thr183 was found to interact with the β-phosphate upon substrate binding which may also help to stabilize the transition state of the reaction. This mechanism is illustrated in Figure 7.

It has been shown by various studies that the two substrates, Hse and ATP, bind to the enzyme mostly in a random order, with slight preference of ATP binding before Hse (*2*, *6*). This is probably because binding of Hse strongly induces a closing-down of Arg187, which in turn induces the closing-down of the flexible "upper lip" near the active site. These conformational changes may partially block the entrance to the nucleotide binding site. However, due to the intrinsic flexibility of the protein structure, the ATP molecule is still able to bind to the enzyme in the presence of Hse with only a slightly increased binding constant. On the other hand, binding of ATP causes few conformational changes and does not affect Hse binding.

The high-resolution three-dimensional structures of HSK−substrate complexes reveal that the interactions between the enzyme and substrates are highly specific. It would be very interesting to know the structural basis of the different substrate specificity in other members of the GHMP kinase superfamily. Because of the highly divergent sequences, detailed catalytic mechanism may also vary among different GHMP members. Therefore, it is highly desirable to have the three-dimensional structures of other GHMP kinases and their complexes with substrates to understand the mechanism that are shared among different members of the superfamily and to reveal the differences in the structure, mechanism, and substrate specificities. This knowledge will be valuable for understanding the evolution in structure and function of this interesting protein superfamily.

ACKNOWLEDGMENT

We thank Mischa Machius and Diana Tomchick for their help with synchrotron data collection, Xinlin Du and Chad Brautigam for enlightening discussions, and Nick Grishin for critical reading of the manuscript.

REFERENCES

1. Theze, J., Kleidman, L., and St. Girons, I. (1974) *J. Bacteriol. 118*, 577−581.
2. Shames, S. L., and Wedler, F. C. (1984) *Arch. Biochem. Biophys. 235*, 359−370.
3. Mannhaupt, G., Pohlenz, H. D., Seefluth, A. K., Pilz, U., and Feldmann, H. (1990) *Eur. J. Biochem. 191*, 115−22.
4. Lee, M., and Leustek, T. (1999) *Arch. Biochem. Biophys. 372*, 135−42.
5. Riesmeier, J., Klonus, A., and Pohlenz, H. (1993) *Phytochemistry 32*, 581−584.
6. Daugherty, M., Vonstein, V., Overbeek, R., and Osterman, A. (2001) *J. Bacteriol. 183*, 292−300.
7. Wormser, E. H., and Pardee, A. B. (1958) *Arch. Biochem. Biophys. 78*, 416−432.
8. Huo, X., and Viola, R. E. (1996) *Arch. Biochem. Biophys. 330*, 373−379.
9. Huo, X., and Viola, R. E. (1996) *Biochemistry 35*, 16180−16185.
10. Bork, P., Sander, C., and Valencia, A. (1993) *Protein Sci. 2*, 31−40.
11. Tsay, Y. H., and Robinson, G. W. (1991) *Mol. Cell. Biol. 11*, 620−631.
12. Cordier, H., Lacombe, C., Karst, F., and Berges, T. (1999) *Curr. Microbiol. 38*, 290−294.
13. Lüttgen, H., Rohdich, F., Herz, S., Wungsintaweekul, J., Hecht, S., Schuhr, C. A., Fellermeier, M., Sagner, S., Zenk, M. H., Bacher, A., and Eisenreich, W. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 1062−7.
14. Zhou, T., Daugherty, M., Grishin, N. V., Osterman, A. L., and Zhang, H. (2000) *Structure Fold. Des. 8*, 1247−1257.
15. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
16. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr. 54*, 905−921.
17. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
18. Collaborative Computational Project Number 4. (1994) *Acta Crystallogr. D 50*, 760−763.
19. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
20. Ramachandran, G. N., and Sasisekharan, V. (1968) *Adv. Protein Chem. 23*, 283−438.
21. Potter, D., and Miziorko, H. M. (1997) *J. Biol. Chem. 272*, 25449−25454.
22. Potter, D., Wojnar, J. M., Narasimhan, C., and Miziorko, H. M. (1997) *J. Biol. Chem. 272*, 5741−5746.
23. Cohen, G. E. (1997) *J. Appl. Crystallogr. 30*, 1160−1161.
24. Esnouf, R. (1997) *J. Mol. Graph. Model. 15*, 133−138.
25. Hellinga, H. W., and Evans, P. R. (1987) *Nature 327*, 437−439.
26. Sigrell, J. A., Cameron, A. D., Jones, T. A., and Mowbray, S. L. (1998) *Structure 6*, 183−193
27. Knowles, J. R. (1980) *Annu. Rev. Biochem. 49*, 877−919.
28. Yoshida, M., and Amano, T. (1996) *FEBS Lett. 359*, 1−5.
29. Hasemann, C. A., Istvan, E. S., Uyeda, K., and Deisenhofer, J. (1996) *Structure 4*, 1017−1029.

BI010851Z