THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 32, Issue of August 6, pp. 33882–33889, 2004

Printed in U.S.A.

Substrate-induced Conformational Changes in Human UMP/CMP Kinase*[S]

Received for publication, February 24, 2004, and in revised form, May 14, 2004
Published, JBC Papers in Press, May 26, 2004, DOI 10.1074/jbc.M401989200

Dario Segura-Peña‡, Nikolina Sekulic‡, Stephan Ort§, Manfred Konrad§, and Arnon Lavie‡¶

From the ‡University of Illinois at Chicago, Department of Biochemistry and Molecular Genetics, Chicago, Illinois 60607 and the §Max Planck Institute for Biophysical Chemistry, Department of Molecular Genetics, Am Fassberg 11, 37077 Göttingen, Germany

Human UMP/CMP kinase plays a crucial role in supplying precursors for nucleic acid synthesis by catalyzing the conversion of UMP, CMP, and dCMP into their diphosphate form. In addition, this kinase is an essential component of the activation cascade of medicinally relevant nucleoside analog prodrugs such as AraC, gemcitabine, and ddC. During the catalytic cycle the enzyme undergoes large conformational changes from open in the absence of substrates to closed in the presence of both phosphoryl donor and phosphoryl acceptor. Here we report the crystal structure of the substrate-free, open form of human UMP/CMP kinase. Comparison of the open structure with the closed state previously reported for the similar *Dictyostelium discoideum* UMP/CMP kinase reveals the conformational changes that occur upon substrate binding. We observe a classic example of induced fit where substrate-induced conformational changes in hinge residues result in rigid body movements of functional domains to form the catalytically competent state. In addition, a homology model of the human enzyme in the closed state based on the structure of *D. discoideum* UMP/CMP kinase aids to rationalize the substrate specificity of the human enzyme.

UMP/CMP kinase plays a critical role in the pathway that supplies nucleotide precursors for DNA and RNA synthesis. The physiological substrates UMP, CMP, and dCMP are reversibly converted by UMP/CMP kinase into their diphosphate form according to the following reaction scheme.

$$\underset{\text{(d)CMP}}{\text{UMP}} + \text{ATP-Mg} \rightleftharpoons \underset{\text{(d)CDP}}{\text{UDP}} + \text{ADP-Mg}$$

REACTION 1

Subsequent phosphorylation by nucleoside diphosphate kinase results in the triphosphorylated form of these nucleotides, which are the substrates for DNA and RNA polymerases. Additionally, UMP/CMP kinase participates in the activation of clinically relevant nucleoside analog prodrugs that are used in cancer and viral chemotherapy. Therefore, understanding the determinants of substrate specificity, the precise roles played by its catalytic residues, and the conformational changes that take place during the catalytic cycle is of paramount importance for the development of more effective anti-cancer and anti-viral agents.

Despite the physiological and medicinal importance of this enzyme, the human form was cloned only in 1999 (1). This work was followed by biochemical characterization of the enzyme, which revealed its ability to phosphorylate synthetic nucleosides (2–4). Important examples are the monophosphate form of the cytidine analogs 1-β-D-arabinofuranosylcytosine (AraC)[1] and 2′,2′-difluorodeoxycytidine (gemcitabine), a mainstay of leukemia and lymphoma therapy. The first step in the conversion of these cytidine analogs to their pharmacologically active triphosphorylated form is catalyzed by deoxycytidine kinase. We have recently published the structure of human deoxycytidine kinase in complex with AraC, gemcitabine, and its physiological substrate deoxycytidine (5). The second step of AraC and gemcitabine activation, from the monophosphate to the diphosphate form, is catalyzed by UMP/CMP kinase. The third phosphate group is added by the non-base-specific nucleoside diphosphate kinase. In addition, UMP/CMP kinase has the remarkable ability to phosphorylate L-nucleotides (3). Of special importance is 3TC (β-L-2′,3′-dideoxy-3′-thiacytidine) used in the treatment of human immunodeficiency virus infection. However, prior to this work, our structural understanding of this important enzyme was limited to the yeast uridylate kinase (6) (50% sequence identity to the human enzyme) and the UMP/CMP kinase from the slime mold *Dictyostelium discoideum* (52% identity) (7). Here we present the structure of the human UMP/CMP kinase in its apo conformation.

UMP/CMP kinase belongs to the nucleoside monophosphate (NMP) kinase family that includes adenylate kinase, guanylate kinase, and thymidylate kinase. All of these globular proteins contain a typical α/β fold that consists of a β-sheet core surrounded by α-helices, with the P-loop motif at the N terminus that binds the phosphoryl donor (in general ATP). Previous structural studies on these enzymes revealed considerable conformational changes that occur in response to substrate binding (8). These conformational changes can be described as the rigid body movement of regions referred to as the NMP-binding and LID regions relative to the CORE domain. The substrate-free enzyme is characterized by a more "open" state where the LID, which typically contains catalytic residues, is most distant from the

* This work is supported by National Institutes of Health Grant AI46943 (to D. S.-P., N. S., and A. L.) and by funds from the Deutsche Forschungsgemeinschaft and the Max-Planck-Gesellschaft (to S. O. and M. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains a supplemental figure and a movie.

The atomic coordinates and structure factors (code 1TEV) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ To whom correspondence should be addressed: University of Illinois at Chicago, Dept. of Biochemistry and Molecular Genetics, 900 South Ashland Ave., Chicago, IL 60607. Tel.: 312-355-5029; Fax: 312-355-4535; E-mail: lavie@uic.edu.

[1] The abbreviations used are: AraC, 1-β-D-arabinofuranosylcytosine; rmsd, root mean square deviation; NMP, nucleoside monophosphate.

NMP-binding region. Upon binding of one the substrates, an intermediate "partially closed" state is formed. Only in the presence of both nucleotides do these enzymes adopt the "closed" (and active) conformation that is characterized by the LID region interacting with the substrate at the phosphoryl donor site.

There is no single NMP kinase of the same organism for which all of the above states have been observed. Because of pronounced enzyme flexibility, it has been difficult to obtain crystal structures of NMP kinases in the absence of ligands that act to stabilize a conformation compatible with crystal growth. In the case of UMP/CMP kinase, all of the structures solved of the yeast and mold enzymes contain substrates or inhibitor occupying both donor and acceptor sites (6, 7, 9–11). Surprisingly, we were able to grow crystals of human UMP/CMP kinase that diffracted to high resolution in the absence of substrates. Comparison of this structure with the homologous *D. discoideum* enzyme reveals the considerable conformational changes that occur upon substrate binding. Here we analyze these conformational changes. In addition, we have built a homology model of the closed form of human UMP/CMP kinase and use this model to analyze the substrate specificity of the human enzyme.

EXPERIMENTAL PROCEDURES

Cloning and Expression of UMP/CMP Kinase—Using a commercially available brain tissue cDNA library (Invitrogen) as template, we retrieved by standard PCR procedures the 588-bp DNA fragment corresponding to the coding region of the so-called "short" form of human UMP/CMP kinase. This 196-residue protein was reported to represent the physiological form of the enzyme in human cells (3). Primer oligonucleotides were designed based on the published DNA sequence (1) introducing suitable endonuclease restriction sites at either end to facilitate direct cloning into the pGEX-RB bacterial expression vector (12): sense primer, 5′-GGAATTCCAT**ATG**AAGCCGCTGGTCGTGTTC-3′ (NdeI restriction site underlined; ATG codon of the first Met residue in bold letters); antisense primer, 5′-CGGGATCCTTAGCCTTCCTTGTCAAAAATCTG-3′ (BamHI site underlined). Clone integrity was verified by automatic DNA sequencing. Expression of the glutathione *S*-transferase fusion protein in the *Escherichia coli* BL21(DE3) strain was induced with 0.1 mM isopropyl-thio-β-D-galactopyranoside, and the cells were grown ~14 h after induction at 22 °C in 2YT medium.

Protein Purification and Nucleotide Kinase Activity Measurements—A purification protocol was developed that included a glutathione-Sepharose column (Amersham Biosciences) and elution of the fusion protein from the column with 10 mM glutathione in 200 mM KCl and 10 mM Tris/HCl, pH 8.0. The glutathione *S*-transferase tag was cleaved with thrombin (SERVA) while dialyzing against 200 mM of KCl, 25 mM Tris/HCl, pH 8.0, 8 mM β-mercaptoethanol. After cleavage and dialysis, the glutathione *S*-transferase was removed by loading for a second time on the glutathione-Sepharose column. The UMP/CMP kinase obtained from the previous step was further purified using gel filtration (S75; Amersham Biosciences), in which it eluted as a single peak of 22 kDa, consistent with a monomer being the physiological form of this enzyme. The protein was stored frozen at −80 °C until used for crystallization experiments.

Enzymatic activity in the forward reaction (ATP and UMP as substrates) was determined at 25 °C in the spectrophotometer using an NADH-dependent assay as described for TMP kinase (12). To probe enzyme activity, characteristic steady state kinetic parameters were determined, keeping the ATP concentration constant at 2 mM and varying UMP between 10 μM and 1 mM, in the presence of 5 mM dithiothreitol. No substrate inhibition of the enzyme (varied from 1 to 5 μg/ml of reaction volume) was observed in this range of UMP concentrations. The observed maximum activity for UMP phosphorylation was 180 μmol/mg of enzyme/min, and the K_m value was 30 μM. The bisubstrate analog P^1-(5′-adenosyl)-P^5-(5′-uridyl)pentaphosphate (UP_5A) (Jena Biosciences) was found to be a very strong inhibitor. At the constant ATP concentration of 2 mM, the apparent UP_5A inhibition constant, K_i, was estimated to lie in the 1–5 nM range.[2] This value is very close to that reported for the *D. discoideum* enzyme (dissociation constant of 3 nM) (13).

[2] D. Segura-Peña, N. Sekulic, S. Ort, M. Konrad, and A. Lavie, manuscript in preparation.

Crystallization, X-ray Data Collection, and Structure Solution—Prior to crystallization, the protein was dialyzed against a solution containing 25 mM KCl, 10 mM Hepes, pH 7.5, and 5 mM dithiothreitol. The crystals were grown at room temperature using hanging drops containing equal volumes of protein and reservoir solutions. The protein solution contained 17 mg/ml of UMP/CMP kinase, 5 mM $MgCl_2$, and either 4 mM AppNHp and 2 mM UMP or 1 mM UP_5A, whereas the reservoir contained 25% polyethylene glycol 4000, 0.2 M $MgSO_4$, and 0.1 M sodium acetate, pH 4.6. Hexagonal crystals appeared after 1–2 days. The crystals were frozen by direct immersion in liquid nitrogen after a short soak in the mother liquor that was complemented with 10% xylitol and 10% of sucrose as cryoprotectants. X-ray data were collected at the Advanced Photon Source (Argonne National Laboratories) using the SER-CAT (Southeast Regional Collaborative Access Team) beamline. Data of high completeness and a signal-to-noise above 2 σ were collected to 2.1 Å and processed with XDS (14). For molecular replacement, we used the model of the *D. discoideum* UMP/CMP kinase in complex with UP_5A (7). A model of human UMP/CMP kinase in the closed conformation was generated using the protein structure homology modeling server Swiss Model (15).

Structure Comparison and Conformational Change Analysis—The structure comparison techniques used here have been previously described by Gerstein and Chothia (16) and Lesk and Chothia (17) and include the *sieve fit*, *fit and refit*, and *fit all* procedures. The conformational change analysis was done using the Cα atoms. We used the program O for graphical analysis (18) and the program LSQMAN (x-ray.bmc.uu.se/usf) for calculating distances between equivalent Cα atoms in different structures and for rmsd analysis. Based on this information, we identified hinge residues. The CCP4 supported program POLYPOSE (19) was used to obtain the rotation matrix and rotation values required for superimposing secondary structural elements between the open and closed conformational states.

RESULTS

Structure Solution—Human UMP/CMP kinase crystallized in the hexagonal space group $P6_522$ with a single molecule in the asymmetric unit. Crystals were grown in the presence of substrate or the bisubstrate inhibitor UP_5A. Therefore, we expected that the structural model of the closed conformation of the *D. discoideum* enzyme complexed with UP_5A (7) would be appropriate for molecular replacement. Using Molrep (20) we obtained a solution, but the crystallographic *R* factor was unexpectedly high (56%). This was consistent with poor electron density for considerable sections of the model. Nevertheless, because of the rather high resolution limit of the data, we decided to attempt automatic model building using the warpNtrace mode of Arp (21). This worked astonishingly well yielding a model missing only 12 of the 196 amino acids with a low crystallographic *R* factor.

Upon inspection of the model built by Arp, we were much surprised to observe a very different conformation of the enzyme compared with our search model. The Arp model revealed UMP/CMP kinase in an open conformation (Fig. 1). No electron density for the substrates could be observed. A strong peak of electron density close to the P-loop was interpreted as a sulfate ion; this is analogous to the sulfate ion observed in the apoadenylate kinase structure (22).

Quality of the Structure—Cycles of manual rebuilding reiterated with refinement using Refmac (23) and addition of water molecules using Arp resulted in the final model with an R_{work} value of 22% and an R_{free} value of 27%. In addition to 111 water molecules, all but the two most N-terminal residues, for which no electron density was observed, were modeled. A single sulfate ion is modeled bound to the P-loop main chain nitrogen atoms. All main chain dihedral angles are found in the allowed regions of the Ramachandran plot; the average *B* factor is 44 $Å^2$. The residues in the LID domain forming a loop connecting the two helices of this region have high temperature factors and poor side chain electron density. This is not surprising because of the lack of substrates in this structure. A summary of model statistics is given in Table I.

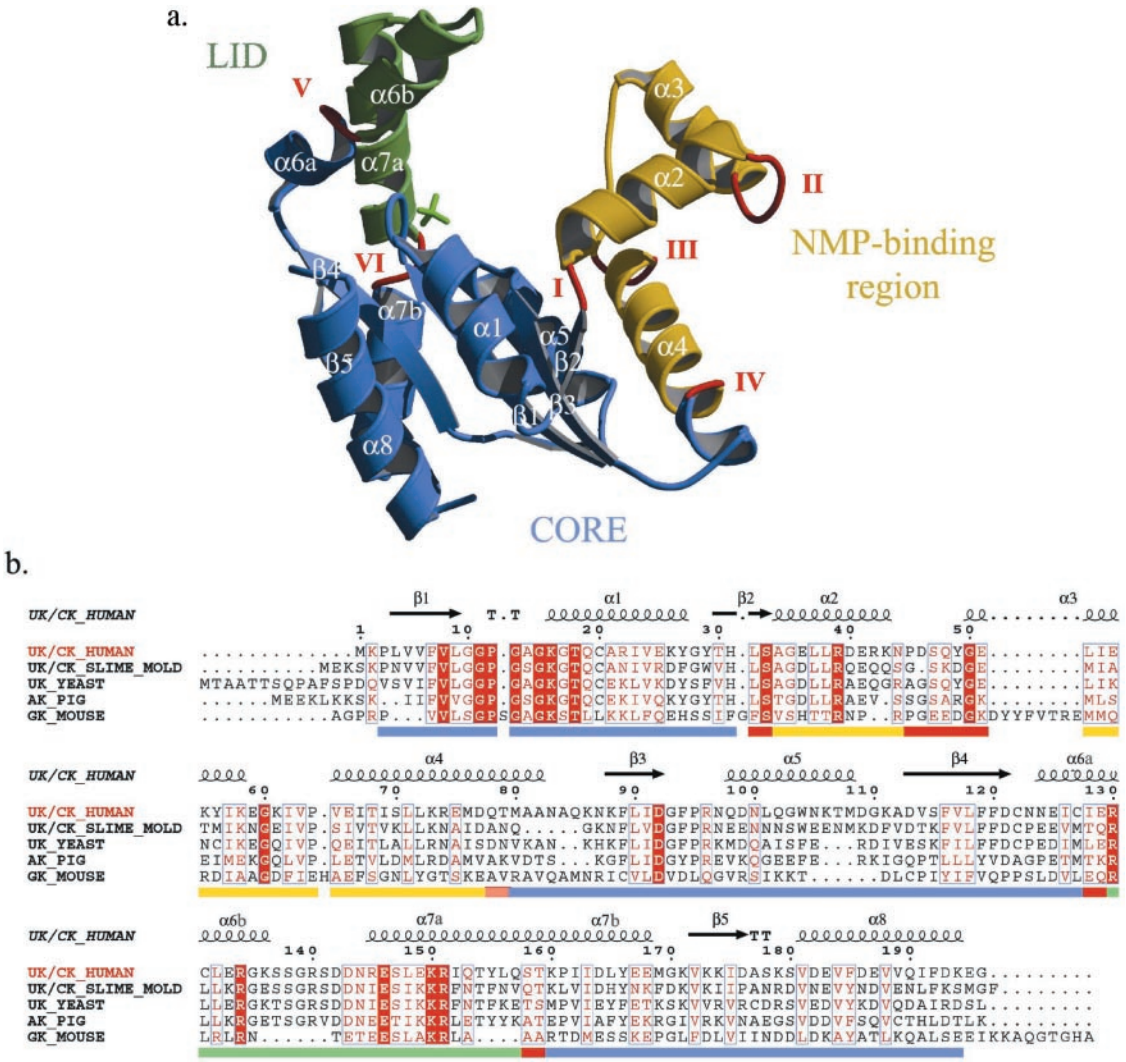

FIG. 1. **Ribbon diagram and sequence alignment of human UMP/CMP kinase.** *a*, ribbon diagram of human UMP/CMP kinase in the open conformation. Upon substrate binding the relative positions of the LID region (*olive*), NMP-binding region (*gold*), and CORE region (*blue*) change to form a more compact closed state. The change in conformation is mainly done by rotation of rigid bodies around hinges, labeled *I–VI*. The sulfate ion observed bound to the P-loop is shown in *green*. *b*, sequence alignment of UMP/CMP kinases with adenylate kinase and guanylate kinase. The secondary structure elements shown above the alignment are of UMP/CMP kinase. The coloring scheme of the *bar* below the alignment corresponds to that used in the ribbon diagram for the individual regions and hinges. All of the structure figures were generated using MOLSCRIPT (37) and RASTER3D (38).

Proline 95 assumes the *cis* configuration. This proline is located in a conserved loop sequence motif (hhh^{92}DGh^{95}P^{96}R, where h stands for bulky hydrophobic) among UMP/CMP kinases and adenylate kinases. In all known structures of adenylate and UMP/CMP kinases, this residue was observed in the *cis* configuration. The *cis* proline configuration confers a fold that enables Asp^{92} and Arg^{96} to participate in magnesium coordination and NMP binding, respectively.

Overall Structure—In an analogous fashion to other NMP kinases (24), the structure of UMP/CMP kinase has been divided into three regions based on functional and structural considerations. These are the CORE, NMP-binding. and LID regions (Fig. 1). The CORE is a rigid part of the protein molecule composed of five parallel β-strands (β1–β5) and helices α1, α5, α6a, α7b, and α8. The P-loop motif responsible for binding the phosphate groups of the phosphoryl donor nucleotide, which connects β1 to α1, was assigned as a part of the CORE region. In the here reported apo-structure, we observed a sulfate molecule bound near the P-loop at the position that corresponds to the β-phosphate of ATP. This highly conserved loop undergoes a slight conformational change upon substrate binding (maximum shift of ~1.9 Å in position of the Cα atom of Gly^{11}). A change of P-loop conformation in response to the nature of the nucleotide bound was also observed in thymidylate kinase (25, 26), suggesting that P-loop flexibility is a general feature in nucleoside monophosphate kinases.

The NMP-binding domain is the part of the molecule that binds the nucleotide monophosphate (in this case CMP or UMP) and consists of the three helices α2, α3, and α4 joined by two short loops. This designation of the NMP-binding region is different to that previously made based solely on the structure of the closed conformation of the enzyme, where helix 4 was not included (7). Our structural comparison of the open and closed conformation leads to the conclusion that this helix should be included in the definition of the NMP-binding region because of its concerted movement with helix 2 (see below).

The LID region is composed of helix 6b followed by a loop that connects to helix 7a (Fig. 1). Structural studies on several NMP kinases including UMP/CMP kinase (10) have revealed that the LID undergoes closure upon binding of substrates, bringing several catalytic residues in close proximity with the phosphate groups of both substrates. Although others have restricted the LID definition to the loop that connects helix 6b and helix 7a, we include helices 6b and 7a in the LID definition

TABLE I
Data collection and refinement statistics

Data collection	
X-ray source	APS(SER-CAT)λ=1.000Å
Unit cell (Å)	$a = b = 62.8$, $c = 226.6$
Space group	$P6_522$
Molecules/asymmetric unit	1
Resolution limit (Å)	15–2.1
Measured reflections	198,383
Unique reflections	15,467
Completeness (%; overall/last shell)	94.5/85.1
$I/\sigma I$ (overall/last shell)	15.3/6.2
R_{sym}(overall/last shell)[a]	11.6/39.3
Refinement	
Resolution limit (Å)	15–2.1
Reflections (working/free)	11,497/1,442
R_{cryst} (%)[b]	21.6
R_{free} (%)	27.4
Non-hydrogen atoms	1659
rmsd from ideal geometry (Å)	
Bond length	0.015
Angle distances	1.444
Coordinate error (Å) (Luzzati)	0.3
Ramachandran plot statistics (%)	
Residues in most favored regions	94.2
Residues in allowed regions	5.8
Residues in generously allowed regions	0
Residues in disallowed regions	0

[a] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$.
[b] $R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$. 10% randomly omitted reflections were used for R_{free}.

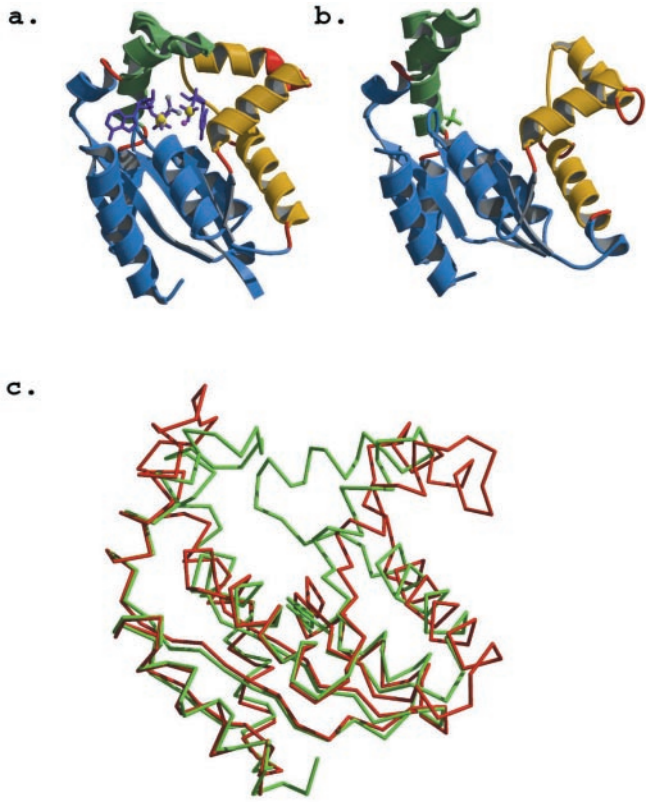

FIG. 2. **Comparing the open and closed conformations of UMP/CMP kinase.** *a*, ribbon diagram of the UMP/CMP kinase from *D. discoideum* in complex with CMP and ADP. *b*, ribbon diagram of the apo human UMP/CMP kinase. In both cases, the CORE is colored in *blue*, the NMP-binding region is in *gold*, and the LID is in *green*. Residues that belong to hinges are shown in *red*. *c*, overlay of apo human UMP/CMP kinase (*red*) and the nucleotide-bound *D. discoideum* UMP/CMP kinase (*green*). The superposition matrix used to achieve this overlay included only atoms from the CORE region. Note the good overlay of the CORE region but the significant differences in the LID and NMP-binding regions.

for human UMP/CMP kinase because both helices change conformation in response to substrate binding. We observed an analogous behavior in the LID region of mouse guanylate kinase (27).

Substrate-induced Conformational Change—To identify and analyze the conformational changes that occur upon substrate binding in UMP/CMP kinase, we compared the structure of the human UMP/CMP kinase in the absence of substrate (reported here) with the structure of the *D. discoideum* enzyme in complex with ADP, aluminum fluoride, and CMP (10) (Fig. 2). To rigorously characterize the movements taking place when the enzyme goes from the open to the closed conformation, three fitting procedures previously reported by Gerstein and Chothia (16) were used. These are *sieve fit*, *fit and refit*, and *fit all* (17).

The conformational change that occurs upon substrate binding is apparent when comparing the structures of the open and closed states. Overlaying the structures using all atoms results in the large rmsd of 4.13 Å. In contrast, the rmsd of the 104 Cα atoms of the CORE regions is only 0.97 Å. This significant difference reveals the presence of mobile regions in the protein molecule.

The overlay based on the CORE region reveals significant conformational changes in the LID and NMP-binding regions. In the absence of substrate, the enzyme is in the open conformation characterized by the LID and the NMP-binding regions being far away from each other. After substrate binding, the enzyme adopts a more compact conformation that brings the LID and the NMP-binding region in close proximity (Fig. 2). For example, the distance between the Cα atoms of Gly^{60}, located in helix 3 of the NMP-binding region, and Arg^{151}, located in helix 7a of the LID region, shortens from 21 Å in the open conformation to 4.5 Å in the closed state. Comparing the positions of Cα atoms between the two conformations identifies the residues that move in relationship to the static core (Fig. 3).

The most dramatic movement occurs in the NMP-binding domain, with a maximum Cα displacement between the open and the closed state of ~18 Å for the Cα atom of Glu^{59}. The LID also undergoes significant movement between the two states. Here the maximum displacement is ~10 Å for the Cα atom of Arg^{140}. Interestingly, this residue is one of the LID arginines that in the closed conformation participates in catalysis by direct interaction with the transferred phosphoryl group of ATP (10). The shift of the LID and NMP-binding regions from the open to the closed form fulfills the criteria of the induced fit mechanism that calls for catalytic residues to adopt a productive conformation only after substrate binding.

Conformational Changes of the NMP-binding Region—The three α-helices that make the NMP-binding region cannot be treated as a single rigid body because helix 3 rotates relative to helices 2 and 4. Thus, the closure of the NMP-binding region is due to motion around four separate hinges that we refer to as hinges I, II, III, and IV (Figs. 1 and 3). The complicated motion around the four hinges can be dissected into a two-step process where hinges I and IV, which connect the NMP-binding region and the CORE, operate on the whole domain, and then hinges II and III operate on helix 3.

Hinge I is composed of residues 32 and 33 and is located in β strand 2 at the boundary of the CORE and NMP-binding regions. The movement of hinge I produces a rotation of 11.5° in helix 2 toward the closed conformation. Analysis of hinge IV is complicated by the fact that the human UMP/CMP kinase contains a five-residue insert after helix 4 in comparison with the *D. discoideum* enzyme. Therefore, it is not possible to precisely identify the residues that constitute hinge IV. The

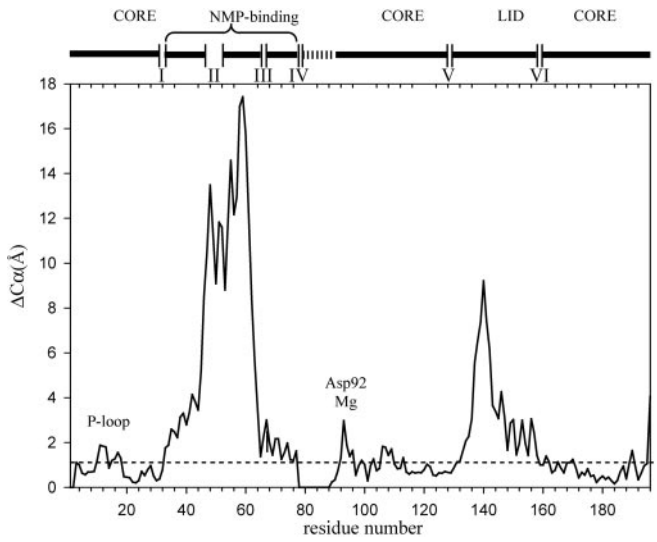

FIG. 3. **Cα-Cα distance plot.** The distance between the position of Cα atoms in the open and closed state is depicted as a function of residue number. The alignment of the enzyme in the two states was done based on Cα atoms of the CORE region. The different regions in the enzyme and the position of hinges are shown above the plot. We set a threshold of ~1 Å between the major moving and the more static residues (*dashed line*) because this is close to the rmsd value for atoms in the CORE region. It is clear from this plot that the most pronounced conformational change upon substrate binding occurs to the NMP-binding region and to a lesser but still significant extent to the LID region. Importantly, residues that belong to the CORE region but that interact with substrates, such as the P-loop and the magnesium binding carboxylic acid Asp^{92}, also exhibit observable movement. This implies that in addition to the major conformational changes observed for the NMP-binding and LID regions, substrate binding causes conformational changes of smaller magnitude in residues of the CORE region that directly interact with the substrate.

effect of hinge IV is to rotate helix 4 by 10.9°. This value is very similar to that observed for hinge I upon helix 2. Therefore, we conclude that helices 2 and 4 move in a concerted manner. This conclusion is supported by the fact that superposition of helices 2 and 4 together in the two different structures gives a rmsd of only 0.32 Å, whereas the overlap of each of the helices individually with its counterpart in the other structure produces a relatively similar rmsd. of 0.15 and 0.22 Å for helix 2 and helix 4, respectively. On the other hand, attempting to use the whole NMP-binding region for the overlay results in a rmsd of 3.5 Å. Based on this analysis, we were able to model a hypothetical intermediate structure between the open and the closed state (Fig. 4). This intermediate structure represents the effects of hinges I and IV upon closure. When helices 2 and 4 are treated as a rigid body, a rotation of 11.4° is required to bring the two helices from the open to the closed state.

Hinges II and III can be seen as hinges within the NMP-binding region. These internal hinges, which act in unison, result in a dramatic conformational change in the position of helix 3 upon closure. Hinge III is composed of residues 65 and 66 that connect helix 3 with helix 4. Hinge II is the most complex in UMP/CMP kinase. To define the residues of the loop that form part of the hinge II, we applied the *fit all* technique (16). The result of this procedure is presented in the form of a contour plot (supplemental figure) where regions with a very steep slope are indicative of residues with significant changes in the ψ and ϕ torsion angles. According to the contour plot, residues 45–51 form hinge II. Interestingly, residues 47 and 48 seem not to contribute to the conformational change, implying that hinge II is composed of two parts. The overlay of the loop residues from residues 45 to 51 in the open state with the equivalent residues in the closed conformation produces a rmsd of 2.95 Å for a total of only seven amino acids. This poor fitting is indicative of the significant change in the residue torsion angles of this region of the enzyme. The combined changes in the torsion angles of the seven residues in hinge II result in the dramatic movement of helix 3 characterized by a rotation of 57°. A rotation value of 64° for helix 3 was obtained by using the program DynDom (28).

It is worth noting that there is good electron density for hinge II, indicating a well structured loop. A total of six intrahinge hydrogen bonds and two interhinge hydrogen bonds (*i.e.* residues of hinge II making hydrogen bonds with other parts of the molecule) act to stabilize this loop. The homology model of the human UMP/CMP kinase in the closed conformation shows that even though the hydrogen bonds pairs change after closure, the total number of hydrogen bond interactions remains the same.

In summary to this section, we have divided the motion of the NMP-binding region into two steps. The first one is the concerted motion around hinges I and IV that moves helices 2 and 4 as a rigid body. The second motion is around hinges II and III that operates on helix 3, resulting in a dramatic rotation of ~57°. The combined effect of hinges I and IV and hinges II and III on helix 3 is a rotation of 68°. The implication of this value being equal to the sum of the individual rotational angles of the two set of hinges (11° from hinges I and IV and 57° from hinges II and III) is that the two sets of hinges move helix 3 in the same direction, in an additive manner.

Conformational Changes of the LID Region—The LID region of most NMP kinases (mammalian thymidylate kinases being a notable exception (29)) contains several arginines that are crucial for catalysis. Only in the substrate-bound state do these arginines adopt a conformation that allows a direct interaction with the phosphates of ATP. The LID of human UMP/CMP kinase is composed of helix 6b, a loop, and helix 7a. Two hinges determine the movement in the LID. The first hinge separates helix 6b from the CORE and is located at residues 128 and 129. The second hinge separates helix 7a from the CORE and is located at residues 158 and 159.

The two helices in the LID region move approximately as a rigid body (rmsd of 0.87 Å for 20 atoms) between the open and closed form with a total rotation of 18°. However, a more complicated conformational change for the loop connecting them is apparent from its poor superposition after applying the transformation matrix based only on the helices. This implies the existence of internal hinges that operate on the loop of the LID region, but because of poor electron density we were not able to identify these residues with confidence. The described analysis of motions generated upon substrate binding is depicted in a movie (supplemental material).

Substrate Specificity—We were not able to obtain crystals of UMP/CMP kinase with bound ligand, despite the presence of substrates (UMP or CMP) or the bisubstrate inhibitor UP_5A in the crystallization buffer. This was rather surprising because UP_5A binds with a dissociation constant in the low nanomolar range to the human UMP/CMP kinase. The human UMP/CMP kinase crystallized only at low pH (~4.6). Analysis of the crystal packing revealed that the side chain of Asp^{56} is 2.4 Å away from the symmetry-related side chain of Asp^{66}. Thus, we conclude that the low pH was required to protonate one or both of these carboxylic acid groups, and only in this state would the productive crystal contacts be made. However, at this pH, the binding of substrate and of the inhibitor UP_5A could be weakened. One explanation for weak UP_5A binding implicates the lack of ability of the protein to coordinate magnesium at low pH. UP_5A will contribute two phosphate oxygens to ligate this magnesium, and the other ligands are coming from water mol-

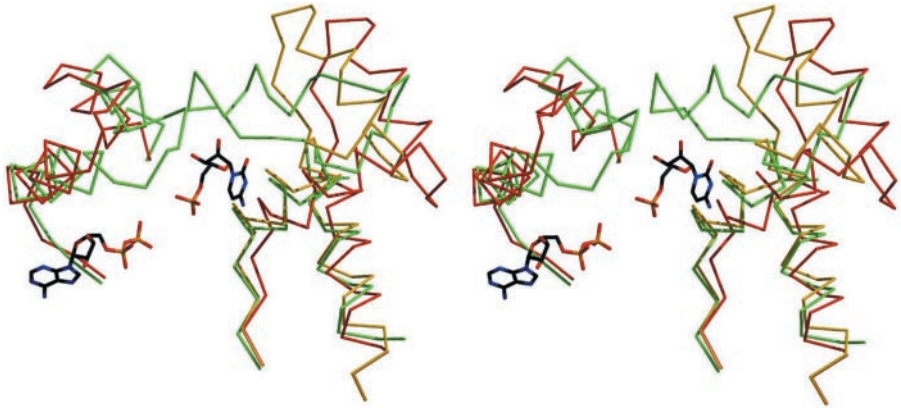

FIG. 4. **Conformation change in UMP/CMP kinase.** Stereo view of UMP/CMP kinase in three states: the open state (*red*), the closed state (*green*), and a hypothetical intermediate state (*yellow*). This hypothetical state was generated by calculating a transformation matrix between the open and closed conformations based on helices 2 and 4. The intermediate structure depicts the effects of hinges I and IV on the whole NMP-binding region. The nucleotides CMP and ADP are depicted in their closed state position. It is evident that both regions act as lids that close around the substrates upon their binding.

ecules, where one of these water molecules is held in position by a conserved carboxylic acid (Asp^{92}). However, at low pH, Asp^{92} will be protonated, abrogating its ability to participate in magnesium binding. In fact, the crystals used to solve the structure of the slime mold UMP/CMP kinase in complex with UP_5A were grown at pH 8.8 (7); in this condition, the magnesium was observed. Yet we were unable to obtain crystals of human UMP/CMP kinase at a higher pH.

As a result, we used the *D. discoideum* structure to homology model the human enzyme in the closed conformation. The purpose was to gain an understanding of the substrate specificity of the human enzyme (3, 4). Analysis of the model of the substrate-bound human UMP/CMP kinase reveals a direct interaction between the enzyme and the 2′-hydroxyl of the NMP (2.8 Å between the 2′-hydroxyl in ribonucleotides and the carbonyl oxygen atom of Lys^{61}) (Fig. 5). This interaction explains the preference of the enzyme for ribonucleotides over 2′-deoxyribonucleotides, which is manifested by a higher K_m value of dCMP in comparison with UMP and CMP. Furthermore, the model rationalizes the high K_m of AraCMP for the human UMP/CMP kinase. Our model predicts that the 2′-ara-hydroxyl group present in this nucleoside analog cannot maintain the interaction with the carbonyl group of Lys^{61}. The 2′-hydroxyl of AraC, modeled either in the C2′-endo conformation, as observed for CMP bound to UMP/CMP kinase (10), or as the C2′-exo, as observed in the deoxycytidine kinase structure (5), is at a distance between 4.5 and 5.5 Å from the carbonyl group of Lys^{61}.

The severe reduction in the phosphorylation rate of pyrimidines modified in the 5-position (*e.g.* TMP *versus* dUMP) can be rationalized because of the close contact (3.1 Å) made by the carbonyl of Gly^{93} with the base C-5 atom. Furthermore, the side chain of Ala^{37} can contribute to the discrimination against TMP. This observation has implications for the development of new nucleoside analogs and suggests that even slight modifications at the 5-position, such as the replacement of a hydrogen atom with a fluorine, could result in reduced phosphorylation efficiency. Concerning the phosphoryl donor-binding site, the preference of the enzyme for ATP (3) becomes evident upon inspection of the model. In our model, we observe a 2.9 Å distance between the carbonyl oxygen of Lys^{179} and the amino group at the 6-position of the adenine base (Fig. 5*a*). GTP would be unfavored because of a carbonyl group at the 6-position.

DISCUSSION

We have obtained the structure of human UMP/CMP kinase in the absence of substrates (open conformation). Comparison of the open and the closed states reveals the substrate-induced conformational changes. The induced fit theory stipulates that substrates can cause an appreciable change in the conformation of an enzyme, and as a result of this change, catalytic residues become correctly positioned to participate in catalysis (30). Structural studies on NMP kinases have supplied us with important observations supporting the induced fit theory (31). The phosphoryl transfer reaction catalyzed by NMP kinases occurs after the two substrates, phosphoryl acceptor and donor nucleotides, are bound at the active site. Data support a random bi-bi mechanism for this reaction (32), being in contrast to nucleoside diphosphate kinase that follows a ping-pong mechanism. This means that at least four different enzyme conformations can be adopted during the catalytic cycle: a substrate-free state, two different states in which either the phosphoryl donor or acceptor nucleotide is bound, and finally, the state in which both substrates are bound. Unfortunately it has proven difficult to obtain crystal structures of all of these states on one particular NMP kinase. Nonetheless, structures of NMP kinases from different species in different states have been obtained. Analysis of these crystal structures made it possible to create a "movie" displaying possible substrate-induced movements during the reaction (8). In the case of guanylate kinase, we have recently analyzed similar movements among the apo, GMP-bound (33), and GMP/ADP-bound states (27). Prior to this work, all structures of UMP/CMP kinase contained nucleotides at both substrate-binding sites. Now having a structure of the uncomplexed form of human UMP/CMP kinase, we demonstrate that during its catalytic cycle this enzyme experiences movements similar to those seen in other NMP kinases.

A change in the conformation of hinge residues results in a rigid body movement of each individual domain relative to the rigid CORE domain. In contrast to guanylate kinase (27), in the case of UMP/CMP kinase, in addition to interdomain hinges, we observe additional hinges that operate within the NMP-binding and LID regions.

For the LID region, these additional hinges are located in the loop connecting its adjacent helices. Thus, although the helical parts of the LID region move as a single rigid body, the loop part undertakes an additional flipping motion. This side chain flipping results in ~18 Å change in the position of the guanidinium moiety of Arg^{140} between the open and the closed conformations, whereas the Cα atoms difference is ~10 Å. Only in its closed conformation can Arg^{140} fulfill its catalytic role by concomitantly interacting with the phosphate of the NMP and the transferred phosphoryl group of ATP (Fig. 5). Similarly,

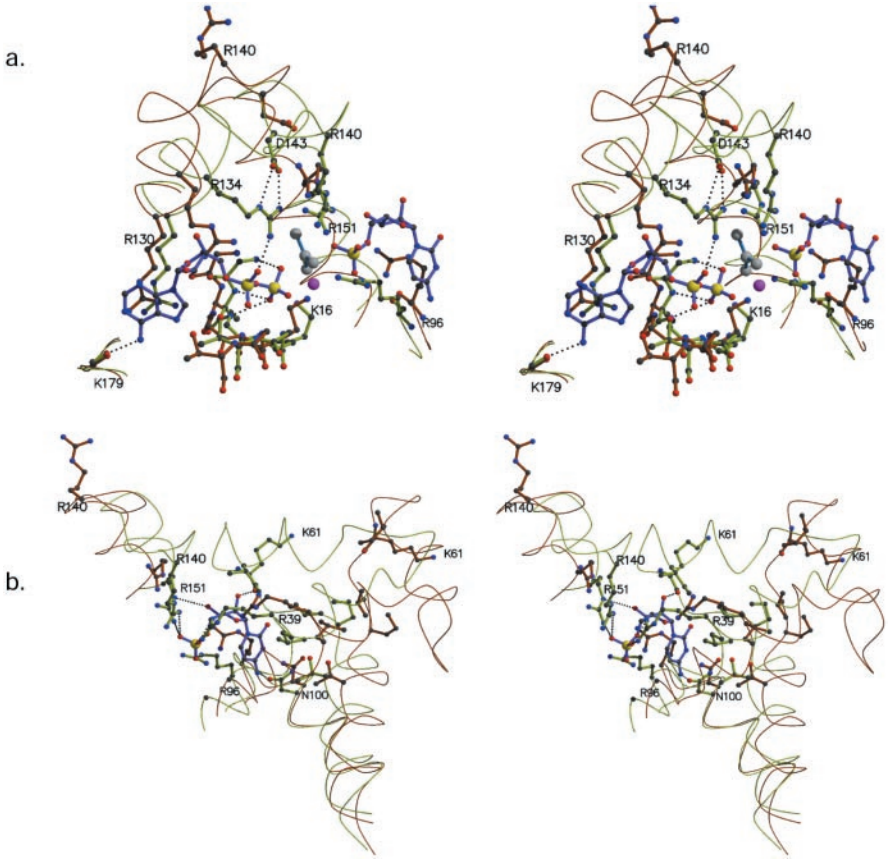

FIG. 5. **The substrate-induced conformational changes generate a competent active site.** Shown is an overlay of the open UMP/CMP kinase (*brown*) and the model of the closed conformation based on the structure of the *D. discoideum* enzyme (*green*). *a*, focus on the LID region movements. Only in the conformation adopted in the presence of both nucleotides (displayed in *purple* are ADP, CMP and the aluminum fluoride as observed in the transition state mimic structure of *D. discoideum* UMP/CMP kinase (10)) can the LID arginines 134 and 140 interact with the substrates. Of special note is the conformational change of Arg^{140} that in the closed form allows it to interact with the AlF_3, which corresponds to the γ-phosphate of ATP. Additionally, interactions that position the catalytic arginines, such as between Arg^{134} and Asp^{143}, are only made possible in the closed state. *b*, focus on the NMP-binding region. Only in the closed form is the UMP/CMP-binding site complete. Here of special note is the interaction made between the CMP ribose and the main chain carbonyl group of Lys^{61}.

another catalytically important residue, Arg^{134}, changes its conformation upon LID closure. Like in the case of Arg^{140}, this residue can only interact with the transferred phosphoryl group in the closed conformation. Moreover, the position of Arg^{134} in the closed conformation is stabilized by Asp^{143}. The distance between the side chains of Arg^{134} and Asp^{143} decreases from 10 Å in the open state to 2.8 Å in the closed state. The flipping motion of the loop in the LID region is facilitated by two conserved glycine residues adjacent to the critical Arg^{134} and Arg^{140} (^{134}**RG**KSS**GR**^{140}). Support for the interpretation of high loop flexibility of the LID region comes from analysis by Hutter and Helms (34) of the UMP/CMP kinase reaction using molecular orbital computations suggesting that conformational flexibility of the LID region is critical for catalysis.

The movement observed for the NMP-binding region occurs via four hinges: two external (I and IV) and two internal (II and III). The motions of the NMP-binding helix 3, with a total rotation of 68°, resemble the closure of a lid. As a consequence of this movement, interactions between residues located in helix 3 and the phosphoryl acceptor nucleotide (UMP or CMP) are made possible. Noteworthy is the interaction made by the 2′-hydroxyl of the NMP substrate and the carbonyl of Lys^{61}. This interaction explains the pronounced selectivity for ribonucleotides as evidenced by the 10–60-fold higher K_m values reported for AraCMP as compared with CMP (3, 4). In addition, hydrophobic interactions with the uridine/cytidine base are formed in the closed conformation by Leu^{38}, Ile^{62}, and Val^{63} of the NMP-binding region.

At the same time the 3′-hydroxyl of the sugar interacts with the LID residue Arg^{151} in its closed conformation. Thus, the sugar of the NMP is not only responsible for the stabilization of the NMP-binding domain in the closed conformation but also in building a bridge between the LID and NMP-binding domains. This explains why dCMP, which lacks a 2′-hydroxyl, and ddCMP, which lacks both 2′- and 3′-hydroxyls, are phosphorylated with poor efficiency corresponding to only 0.5 and 0.1% of that of CMP, respectively (3).

However, not all of the residues involved in substrate binding undergo a dramatic conformational change. For example, in the case of UMP or CMP binding, the surface made by the conserved sequence that forms strand β3 and its following loop changes its conformation only moderately. Asn^{100}, a residue that plays an important role in the base specificity of UMP/CMP kinase, is rather static. In addition, Arg^{96}, responsible for interacting with the UMP or CMP phosphate group, is a part of the static CORE region. A similar situation occurs at the ATP-binding site, where certain residues, such as the P-loop (which binds the α and β phosphates), or Lys^{179} (which interacts with the adenine amino group), experience only a slight conformation change in response to ATP binding. Therefore, these residues are likely to be required for the initial substrate recognition.

The current theory postulates that enzymes oscillate between different low energy states, and the ligands act to stabilize the closed conformation (35, 36). In terms of the induced fit mechanism, we conclude that the binding sites are partially formed in the absence of substrates. After the initial substrate-binding event, which occurs with the enzyme in the open state, closure of the NMP-binding and LID regions produces additional contacts with the substrate that stabilize the closed state. The LID primarily moves as a single rigid body, with an added flipping motion for the loop that connects the two helices of the LID region. In contrast, the NMP-binding domain undergoes a pronounced conformational change that can be divided into two distinct motions: an overall rigid body movement of the whole domain and an additional motion of the central helix relative to the rest of the domain. Together, these movements of the LID and NMP-binding regions build a catalytically competent active site. As was previously pointed out (9),

by utilizing the induced fit mechanism NMP kinases can prevent futile ATP hydrolysis by disabling the active center in the absence of both substrates. Therefore, our results show the major role played by enzyme dynamics in catalysis and substrate recognition in human UMP/CMP kinase.

Acknowledgments—We thank the staff at SER-CAT (Southeast Regional Collaborative Access Team) for help in data collection.

REFERENCES

1. Van Rompay, A. R., Johansson, M., and Karlsson, A. (1999) *Mol. Pharmacol.* **56,** 562–569
2. Labenz, J., Friedrich, D., and Falke, D. (1982) *Arch Virol.* **71,** 235–249
3. Liou, J. Y., Dutschman, G. E., Lam, W., Jiang, Z., and Cheng, Y. C. (2002) *Cancer Res.* **62,** 1624–1631
4. Pasti, C., Gallois-Montbrun, S., Munier-Lehmann, H., Veron, M., Gilles, A. M., and Deville-Bonne, D. (2003) *Eur. J. Biochem.* **270,** 1784–1790
5. Sabini, E., Ort, S., Monnerjahn, C., Konrad, M., and Lavie, A. (2003) *Nat. Struct. Biol.* **10,** 513–519
6. Muller-Dieckmann, H. J., and Schulz, G. E. (1994) *J. Mol. Biol.* **236,** 361–367
7. Scheffzek, K., Kliche, W., Wiesmüller, L., and Reinstein, J. (1996) *Biochemistry* **35,** 9716–9727
8. Vonrhein, C., Schlauderer, G. J., and Schulz, G. E. (1995) *Structure* **3,** 483–490
9. Muller-Dieckmann, H. J., and Schulz, G. E. (1995) *J. Mol. Biol.* **246,** 522–530
10. Schlichting, I., and Reinstein, J. (1997) *Biochemistry* **36,** 9290–9296
11. Schlichting, I., and Reinstein, J. (1999) *Nat. Struct. Biol.* **6,** 721–723
12. Brundiers, R., Lavie, A., Veit, T., Reinstein, J., Schlichting, I., Ostermann, N., Goody, R. S., and Konrad, M. (1999) *J. Biol. Chem.* **274,** 35289–35292
13. Wiesmüller, L., Scheffzek, K., Kliche, W., Goody, R. S., Wittinghofer, A., and Reinstein, J. (1995) *FEBS Lett.* **363,** 22–24
14. Kabsch, W. (1993) *J. Appl. Crystallogr.* **24,** 795–800
15. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) *Nucleic Acids Res.* **31,** 3381–3385
16. Gerstein, M., and Chothia, C. (1991) *J. Mol. Biol.* **220,** 133–149
17. Lesk, A. M., and Chothia, C. (1984) *J. Mol. Biol.* **174,** 175–191
18. Jones, T. A., Zhou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
19. Diamond, R. (1992) *Protein Sci.* **1,** 1279–1287
20. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30,** 1022–1025
21. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
22. Dreusicke, D., Karplus, P. A., and Schulz, G. E. (1988) *J. Mol. Biol.* **199,** 359–371
23. Murshudov, G. N., Vagin, A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255
24. Schulz, G. E., Müller, C. W., and Diederichs, K. (1990) *J. Mol. Biol.* **213,** 627–630
25. Ostermann, N., Schlichting, I., Brundiers, R., Konrad, M., Reinstein, J., Veit, T., Goody, R. S., and Lavie, A. (2000) *Struct. Fold. Des.* **8,** 629–642
26. Ostermann, N., Segura-Pena, D., Meier, C., Veit, T., Monnerjahn, C., Konrad, M., and Lavie, A. (2003) *Biochemistry* **42,** 2568–2577
27. Sekulic, N., Shuvalova, L., Spangenberg, O., Konrad, M., and Lavie, A. (2002) *J. Biol. Chem.* **277,** 30236–30243
28. Hayward, S., and Berendsen, H. J. (1998) *Proteins* **30,** 144–154
29. Lavie, A., Vetter, I. R., Konrad, M., Goody, R. S., Reinstein, J., and Schlichting, I. (1997) *Nat. Struct. Biol.* **4,** 601–604
30. Koshland, D. E. (1994) *Angew. Chem. Int. Ed. Engl.* **33,** 2375–2378
31. Yan, H., and Tsai, M. D. (1999) *Adv. Enzymol Relat. Areas Mol. Biol.* **73,** 103–134
32. Rhoads, D. G., and Lowenstein, J. M. (1968) *J. Biol. Chem.* **243,** 3963–3972
33. Blaszczyk, J., Li, Y., Yan, H., and Ji, X. (2001) *J. Mol. Biol.* **307,** 247–257
34. Hutter, M. C., and Helms, V. (2000) *Protein Sci.* **9,** 2225–2231
35. Kumar, S., Ma, B., Tsai, C. J., Wolfson, H., and Nussinov, R. (1999) *Cell Biochem. Biophys.* **31,** 141–164
36. Gerstein, M., Lesk, A. M., and Chothia, C. (1994) *Biochemistry* **33,** 6739–6749
37. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
38. Merrit, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 869–873