600 ps Molecular Dynamics Reveals Stable Substructures and Flexible Hinge Points in cAMP Dependent Protein Kinase

Abstract: Molecular dynamics simulations of the catalytic subunit of cAMP dependent protein kinase (cAPK) have been performed in an aqueous environment. The relations among the protein hydrogen-bonding network, secondary structural elements, and the internal motions of rigid domains were examined. The values of fluctuations of protein dihedral angles during dynamics show quite distinct maxima in the regions of loops and minima in the regions of \( \alpha \)-helices and \( \beta \)-strands. Analyses of conformation snapshots throughout the run show stable subdomains and indicate that these rigid domains are constrained during the dynamics by a stable network of hydrogen bonds. The most stable subdomain during the dynamics was in the small lobe including part of the carboxy-terminal tail. The most significant flexible region was the highly conserved glycine-rich loop between \( \beta \) strands 1 and 2 in the small lobe. Many of the main chain dihedral angle changes measured in a comparison of the crystallographic structures of “open” and “closed” conformations of cAPK correspond to the highly flexible residues found during dynamics. © 1999 John Wiley & Sons, Inc.

Keywords: computer simulation; molecular dynamics; fluctuations in proteins; protein kinase

INTRODUCTION

The protein kinases represent a large family of enzymes. All perform the same reaction, namely the transfer of phosphate from ATP to a serine, threonine, or tyrosine residue of a protein substrate. These proteins are involved in cellular regulatory processes and play pivotal roles in diverse signal transduction pathways. Forty-nine
Crystal structures of protein kinases are now available in the Protein Data Bank. These, together with the high degree of sequence similarity between the members of this family, indicate that they share a common structure and mechanism of action. The best studied member of the family is cAMP-dependent protein kinase (cAPK). In the cell the catalytic subunit of cAPK is maintained in an inactive state by forming a tetrameric complex with a regulatory (R) subunit. On addition of cAMP this holoenzyme complex (R2C2) dissociates to yield two active monomeric C-subunits and an R-subunit dimer. The C-subunit of this kinase serves as a prototype for the catalytic domains of the entire family. The catalytic mechanism has been described in some detail, and there are now several published crystal structures of the C-subunit.

The sequence and structural homologies between members of the kinase family are restricted to the kinase core. The crystal structures of the C-subunit complexed with different nucleotides and/or peptides 5–10 reveal that it consists of this core and two flanking regions, a myristylated segment (residues 1–39) dominated by a long N-terminal helix and an extended C-terminal tail (residues 300–350) which wraps over the entire core structure. The conserved core consists of two lobes. The N-terminal lobe (residues 40–120) dominated by a 5-stranded β-sheet is responsible for nucleotide binding. The larger C-terminal lobe is mostly helical, containing a four helix bundle, two additional helices, and four short β-strands. Most of the highly conserved active site residues and most of the residues involved in peptide binding are contributed by this large lobe. MgATP is bound in a deep cleft between the two lobes. The adenosine portion is deeply buried in a hydrophobic pocket. The α and β phosphates of ATP make multiple interactions with residues in the small lobe while the γ phosphate makes contact with both lobes. Beta strands 1 and 2 are linked by a glycine-rich loop motif which is anchored to the phosphates of ATP. The peptide lies on the surface of the large lobe. There are multiple hydrophobic and electrostatic interactions that orient the peptide so that the site of phosphorylation is placed at the cleft interface close to the γ-phosphate of ATP. Phosphoryl transfer takes place at the edge of the cleft.

The crystal structures of the C-subunit fall into two classes. The closed conformation is observed in a binary complex of recombinant C-subunit (rC) with the peptide inhibitor PKI(5–24);6 in ternary complexes of rC with Mn2ATP,11 Mg2ATP,9 AMPPNP,5 or adenosine and PKI(5–24);6 and in a ternary complex of ADP with the substrate analog PKS(5–24, N20A, A21S).7 Many lines of evidence indicate that phosphoryl transfer takes place in this closed conformation.12 However, since the MgATP is completely masked from solvent there must be a conformational change in the enzyme to allow MgATP to enter or the product, MgADP, to leave.

A crystal structure of an “open” conformation binary complex of the mammalian C-subunit with PKI(5–24) has also been solved.10 A detailed analysis of the open and closed molecules, as well as a C-adenosine binary complex which shows an intermediate conformation,8 yields a great deal of information on the conformational flexibility of this molecule and provides insight into molecular motion. This conformational flexibility is an essential feature of the overall catalytic process. Examination of the two structures has suggested that certain residues may represent hinge points allowing the relative motions of the two domains. The glycine rich loop is flexible, usually shows the highest B-factors, and its conformation depends on the presence of the nucleotide phosphates. To achieve the degree of domain reorientation that is indicated by these crystal structures and by solution studies,13 it is likely that several regions of the molecule must display flexibility. Other regions of the molecule that adopt similar conformations in the crystal structures may also be dynamic in solution. Additional data to complement information derived from an analysis of the crystal structures is gained by performing theoretical dynamic calculations.

RESULTS AND DISCUSSION

The energy of the system drops substantially in the first 20 ps of molecular dynamics and then levels off (Figure 1a). Four methods of analysis showed that the secondary structure elements overall are remarkably stable (Figure 1b). First, an analysis of the changes in dihedral angles throughout the run indicates flexible regions of the molecule. Second, rigid substructures within the kinase were identified using a method of rigid domain extraction developed for studying conformational changes in hemoglobin.14 Third, an analysis of hydrogen-bonding interactions which remain stable throughout the dynamics run yielded insights into the mechanism by which these substructures are maintained and how they interact. Finally, a principal component analysis of the correlation matrix showed further information on the coordinated domain motions of the molecule.

Radius of Gyration

Figure 1c depicts the radius of gyration of cAPK as a function of time during the simulation. Figure 1d
shows that the radius of gyration correlates extremely closely with the distance between the centers of mass of the small and large lobes, defined for this calculation as residues before and after residue 121. The radius of gyration for the “closed” crystal structure of cAPK is 19.8 Å and for the open crystal structure is about 21 Å. The MD calculations were started from the closed structure, and the value of the radius of gyration increased through the run up to 21.2 Å. The radius of gyration finally stabilized at 21 Å. Experimental values of the radius of gyration determined by small-angle x-ray scattering are 21.1 ± 0.8 Å for the ternary complex with ATP and PKI(5–22) and 23.1 ± 0.5 Å for cAPK alone. A small-angle neutron scattering result of 18.7 ± 0.5 Å for the binary and ternary complexes and 19.5 ± 0.5 Å for the apoenzyme is reported in ref. 15. The changes in radius of gyration are 2 ± 1.3 Å by small-angle x-ray scattering, 0.8 ± 1.0 Å by neutron scattering, 1.2 Å from the crystal structure determinations, and 1.4 Å in our molecular dynamics simulation. These values agree within experimental error. The conformation was not identical with that of the “open” crystal structure. The small lobe has rotated several degrees further in the dynamics snapshot from 580 ps than the rotation from closed to open conformations, and several important loops (the glycine-rich loop and the activation loop, in particular) have moved substantially from their positions in either structure.

Conserved Water Molecules

This simulation was started without incorporating the crystallographic water molecules. Instead the protein was equilibrated with a water bath. A comprehensive analysis of conserved water molecules in a series of cAPK crystal structures was done by Shaltiel et al. Comparison of the sites reported in that paper with cAPK crystal structures was done by Shaltiel et al. Analysis of conserved water molecules in a series of cAPK crystal structures was done by Shaltiel et al.16 Comparison of the sites reported in that paper with water molecules from the MD simulation is shown in Table I. Ten of thirteen water molecules appear to be conserved in the MD simulation.

The first class of water molecules reported by Shaltiel et al. is the one they denote ASCW, molecules that interact with ATP. Table I shows that waters a, b, d, and f are well preserved even in the absence of ATP. Waters c, e1, and e2 (which have only one interaction each with the protein) have no additional support in the absence of ATP and are not consistently found in our analysis. The second class of water molecules, denoted LSCW, consists of waters found when ATP is not present, which replace atoms of ATP in interactions with protein. We find that all of these water molecules are persistently present throughout the molecular dynamics study.

Analysis of Flexibility

To determine which regions of the C-subunit display flexibility the changes in dihedral angles were examined. This method is more informative than Cartesian coordinate shifts. Significant changes in a dihedral angle often lead to noticeable changes in the overall conformation of a protein. Since the values of both backbone dihedral angles, φ and ψ, define the secondary structure of a molecule, the changes in either of them individually would not be sufficient to determine how the conformation of a molecule was changed during dynamics. The distance between (φ,ψ) and (φ,ψ) on a two-dimensional Ramachandran plot was defined as the parameter L and is a measure of the conformational change of a particular residue. The parameter L has the virtue of incorporating changes in both backbone conformational angles into one scalar quantity that can be analyzed for each residue as a function of time.

Figure 2 depicts the RMS value of L. The lower portion of the large lobe shows greatly reduced variability compared to the rest of the molecule. The significant peaks of flexibility are at residues 36, 44 and 45, 52–56, 62–64, 68 and 69, 106 and 107, 163–166, 190–193, 199–201, 215–217, 256, 309, 314, 317 and 318, 333, and 343. All of the peaks of flexibility correspond to loops between secondary structural elements. The peak at residue 35 is at the end of the A-helix and may indicate a hinge at this point. The A-helix bridges the large and small lobes and remains more closely aligned with the large lobe in the open crystallographic conformation. When the domains reorient, the A-helix must move relative to the small lobe of the kinase. A group of peaks from residues 45 to 55 corresponds to the flexible glycine-rich loop that is strongly associated with ATP binding. Hydrogen bonds between the amide nitrogen of Ser 53 and the γ-phosphate of ATP and Gly 55 and Phe 54 with the β-phosphate of ATP help anchor the loop. In crystal structures where the γ-phosphate of ATP is absent this loop is relatively disordered. The region of flexibility between residues 106 and 107 is of particular interest. The loop 97–107 is the only region in the N-terminal lobe of the core that remains aligned with the large lobe in the open conformation crystal structure. Hinge points at either end of this loop would be required to allow flexing to occur. The peak in the flexibility plot for the 106–107 hinge is clear, and a distinct peak occurs also at residues 98 and 99.

The two lobes are linked by a section of polypeptide in an extended conformation between the end of β-strand 5 in the small lobe and the beginning of
α-helix D in the large lobe. A hinge at residue 126 was proposed on the basis of small-angle x-ray scattering studies which indicated a rotation of the two domains by as much as 20°. Our calculations indicate that there is increased flexibility over a range of residues around 120 and 121 and 124–126. This suggests that the motion is not a simple hinge at one residue.8

In the large lobe significant flexibility of residue 136 corresponds to the loop between helices D and E. Following helix E in the structure there are four short β-strands. These exhibit very little flexibility except in the loop between β-6 and β-7. This loop contains Asp 166, which has been postulated as a catalytic base. The next regions of significant flexibility are within a long section of polypeptide from residue 191

FIGURE 1  Summaries of the dynamic behavior of cAPK throughout the course of the simulation. (a) Energies of the system: Total ($E_{\text{tot}}$), potential ($E_p$), and kinetic ($E_k$). (b) 50 ps snapshots of the α-carbon trace throughout the dynamics run. The solid line is the initial conformation of cAPK. (c) Radius of gyration for cAPK. (d) Plot of the distance between center of mass of the large and small lobes against the radius of gyration for 20 ps snapshots of the dynamics trajectory. Frames a–c indicate that the simulation initially equilibrated properly and that the dynamics are stable. Frame d shows that the two lobes undergo concerted shifts during the dynamics which are reflected in the radius of gyration.
to 217 that contains little secondary structure. This section contains two functional loops. The “activation” loop 190–199 contains the essential phosphothreonine 197. This residue is coordinated by Arg 165 and Lys 189 in the large lobe. None of residues Thr 197, Arg 165, or Lys 189 display significant flexibility, suggesting that they form an anchor, maintaining structural integrity in this part of the molecule. It is of note that this lack of flexibility is seen in this molecular dynamics run which was performed in the absence of the phosphate moiety on the Thr 197. One interpretation is that once this loop is ordered the phosphate is not responsible for maintaining the structure of this loop. In protein kinase C (PKC) initial processing of the enzyme to an active form requires phosphorylation of the equivalent to Thr197 as well as at two other phosphorylation sites. After this processing only a portion of the mature PKC retains the phosphate at the Thr 197 site17. In this case also it has been suggested that once initial ordering of the loop has been induced by phosphorylation the active conformation can be maintained in the absence of phosphorylation. The region between the two anchor points at 189 and 197 does display flexibility as expected for a loop structure.

Following the phosphorylation loop is the P + 1 loop, from residues 198 to 205. In complexes of the kinase with the inhibitory peptide PKI(5–24) this loop forms a hydrophobic pocket to receive a hydrophobic side chain at the P + 1 site in the peptide substrate. This loop also shows some flexibility. Before entering helix F, the peptide goes through another turn. The highest flexibility in the kinase is found here at residues 215 and 216. Flexibility also occurs in the loops between helices G and H and H and I. Two of these loop regions make key interactions with the regulatory subunit of the kinase.

The extensive flexibility of the C-terminal tail is not surprising and is consistent with the crystal structure of the open conformation in which parts of the C-terminal tail are either not seen at all or have high crystallographic atomic displacement factors. This is a region of extended structure, and the most interesting sites in this region are rather the nonflexible points that may be linked to rigid domains (see below).

Conformation Stability (Rigid Domains)

To complement the analysis of the flexible structure, we also searched for “rigid bodies” or rigid domains...
of the protein whose configurations were maintained throughout the whole dynamics run. This can not be recalculated from the flexibility measurement because the RMS analysis of dihedral angles fluctuation does not indicate correlated movement of parts of the molecule. A method of rigid domain extraction was used to locate rigid substructures. The method of Nichols et al. was developed for analysis of conformational changes between two stable states of a macromolecule. Our use of the method in analysis of dynamics calculations is for the identification of conserved substructures throughout the simulation, as opposed to following a time course. Domains were extracted for all pairs of snapshots following the equilibration period and separated by ≥60 ps. The separation was necessary to prevent overly high correlation between consecutive snapshots. Essentially all pairs of snapshots showed the same two domains, but individual residues within each domain showed significant variability as to the fraction of snapshot pairs in which they were present in the corresponding domain. The two domains correspond largely to the previously described large and small lobe structure of cAPK.

Figure 3a,b show the frequency plots for domain 1 (a) and domain 2 (b). Each plot in Figure 3 shows the percentage of the snapshots in which each residue was found associated with either domain 1 or domain 2. Figure 4 shows the rigid domains mapped onto the three-dimensional structure of cAPK. It is clear from Figures 3a and 4 that domain 1 includes the small lobe of cAPK, including the β-sheet structure and the C helix, and also includes the highly conserved Asp 184 and residues from the C-terminal tail. Cross-linking between residues Asp 184 and Lys 72 was consistently found when the catalytic subunit of cAPK was modified by dicyclohexylcarbodiimde (DCCD). The fact that these residues are invariant in all protein kinases supports the idea that each of them plays a significant role in catalytic process. It is known that Asp 184 has an essential ATP binding function in conjunction with residues in the small lobe, and the C-terminal tail wraps around the small lobe of the catalytic core.

<table>
<thead>
<tr>
<th>Water</th>
<th>Neighbors</th>
<th>Crystal Structures</th>
<th>10 ps</th>
<th>80 ps</th>
<th>240 ps</th>
<th>320 ps</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCW</td>
<td>a Lys-72 NZ</td>
<td>3.3</td>
<td>3.4</td>
<td>3.4</td>
<td>3.1</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>Glu-91 OE1</td>
<td>2.5</td>
<td>2.6</td>
<td>2.6</td>
<td>2.7</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>b Glu-91 OE2</td>
<td>2.9</td>
<td>4.4</td>
<td>4.5</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c Asp-184 O</td>
<td>2.6</td>
<td>2.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>d Xxxx OH</td>
<td>1.0</td>
<td>1.0</td>
<td>1.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yyyy OH</td>
<td>1.6</td>
<td>1.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>e Asp-184 OD1</td>
<td>3.0</td>
<td>2.6</td>
<td>3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>f Tyr-330 OH</td>
<td>2.7</td>
<td>2.7</td>
<td>2.8</td>
<td>3.1</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>g Leu-49 O</td>
<td>2.7</td>
<td>2.6</td>
<td>2.8</td>
<td>4.9</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>h Glu-127 OE</td>
<td>3.6</td>
<td>2.8</td>
<td>2.6</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>i Lys-168 NZ</td>
<td>2.8</td>
<td>3.3</td>
<td>4.2</td>
<td>3.2</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>j MOLSOL</td>
<td>1.4</td>
<td>1.5</td>
<td>1.7</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>LSCW</td>
<td>k Lys-123 N</td>
<td>3.0</td>
<td>2.9</td>
<td>3.7</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>l Glu-121 O</td>
<td>3.0</td>
<td>2.7</td>
<td>3.0</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>m Glu-127 OE2</td>
<td>3.5</td>
<td>3.4</td>
<td>2.6</td>
<td>2.8</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>n Thr-51 O</td>
<td>3.3</td>
<td>2.8</td>
<td>4.1</td>
<td>4.9</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>o Lys-168 NZ</td>
<td>2.8</td>
<td>3.3</td>
<td>4.2</td>
<td>3.2</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>p MOLSOL</td>
<td>1.4</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

* The nearest neighbors of each water molecule that positioned itself during MD run in vicinity of defined in (ref. 16) conserved water sites are indicated, including residue, atom, and distance to OH2 atom of closest water molecule.

* In this row the distances between the water molecule from MD run to the conserved water ASCWd (ref. 16) is measured.

* During MD run OE atoms of Glu-127 exchanged their positions. OE1 took the positions of OE2 and vice versa.

* The same water molecule as in previous row.
Domain 2 contains a number of highly important regions of the large lobe of the protein kinase catalytic core. The conservation of conformation in these regions is particularly interesting. The catalytic loop region is prominent. Arg 165, which flanks the catalytic base Asp 166, is a member of the domain, but Asp 166 itself is included in domain 2 less than 20% of the time. Glu 170 and Asn 171, both of which are found in domain 2 40% of the time, are involved in ATP binding. Residues 197–211, which include the P₁ substrate recognition site (residues 197–206), are among the most strongly conserved structural features of the large lobe. Thr 197 is the essential phosphorylation site. It has been previously noted that this residue shows little flexibility even though it was not phosphorylated in this model. A stretch from residues 219 to 230 on the F helix is also strongly conserved. These residues are involved in stabilization of the catalytic loop conformation and in substrate recognition. Asp 220 forms hydrogen bonds back to Arg 165 and Tyr 164 in the crystal structures, stabilizing the conformation of the catalytic loop. The nearly invariant residue Arg 280 interacts strongly with Glu 208, so it is not surprising that both are included in domain 2. Glu 208 is, as previously noted, in the section that includes the P + 1 substrate binding site. It is also very interesting to note that the activation loop, residues 193–196, is not included in either domain.

### Hydrogen Bond Network

Hydrogen bonds (HB) were calculated for all snapshot conformations of the protein, and the frequency of occurrence of those bonds during the period from 100 to 610 ps was calculated. Figure 5a shows the stable hydrogen bonds that persist throughout the dynamics run. The diagonal of this graph is consistent with the neighbor-neighbor interactions along the sequence and reflects secondary structure features. This part of the plot contains information about helices and β-strands in the protein. Stable hydrogen bonds that lie off the diagonal indicate interactions between secondary structural elements and extended structures. These fall into six main regions: (a) intra-small lobe
interactions (residues 15–124); (b) intra-large lobe interactions (residues 124–300); (c) small lobe interactions with large lobe (residues 15–124 versus residues 125–300); (d) small lobe interactions with the C-terminal tail (residues 15–124 versus 300–350); (e) large lobe interactions with the C-terminal tail (residues 125–300 versus 300–350); (f) intra-tail interactions (residues 125–300 versus 300–350). Many of these hydrogen bonds are also present in the crystal structures of cAPK. In particular, all of the secondary structural elements evident in the diagonal portions of Figure 5 are present in the 1ATP coordinate set, as expected. In addition, all of the persistent off-diagonal hydrogen bonds in the small–small, small–large, and large–large interaction regions of Figure 5 are found in the crystal structure.

Significant shifts are observed in the interactions of the small and large lobes of the enzyme with the C-terminal tail, most notably in the small lobe. Threonine 48 occupies a central position at the base of the glycine-rich loop and in the crystal structure forms hydrogen bonds from O$_{g1}$ to Arg 56 (N$_{h2}$) and Glu 332 (O$_{e1}$) and from N to Glu 332 (O$_{e2}$). Arg 56, which is at the end of the glycine-rich loop opposite Thr 48, is further anchored by a hydrogen bond from N$_{h1}$ to Glu 333 O$_{e1}$. There was a dramatic switch during the dynamics run such that Glu 332 hydrogen bonds to Arg 56 N$_{h2}$ in 87% of the dynamics snapshots, while retaining the hydrogen bond to Thr 48 N 86% of the time and the hydrogen bond to Thr 48 O$_{g1}$ 45% of the time. Lysine 47, which forms a hydrogen bond from N to Thr 48 O$_{g1}$ in the crystal structure, forms hydrogen bonds with the carboxylate group of Asp 329 in 76% of the dynamics snapshots. These shifts in specific interactions involving residues at the base of the glycine-rich loop may indicate a mechanism for coupling opening of the loop (required for release of ADP and binding of ATP) to mobility of the C-terminal tail. Much of the C-terminal tail is poorly localized in all crystal structures of cAPK, as shown by high B-factors, and the flexibility analysis shown in Figure 2 indicates considerable conformational flexibility in this region. The hydrogen-bonding interactions between Asp 329, Glu 332, and Glu 333 and residues Lys 47, Thr 48, and Arg 56 may provide a bias toward an open position of the glycine loop that is counteracted by the interactions between residues at the end of the loop and the phosphate residues of ATP once ATP binds.

**Correlated Motion**

The eigenvalues of the correlation matrix clearly showed two and only two dominant modes of correlated motion over the period of this calculation. The two largest eigenvalues are 177.3 and 121.2; the third largest is 16.8. 1 order of magnitude smaller than the largest eigenvalue. The remaining eigenvalues decrease rapidly to complete insignificance. The eigenvectors for the two strongest modes are shown in Figure 6. Note that the strongest mode is systematically positive in the small lobe, and systematically negative in the large lobe, except for positive values at residues 170, 184, and the C-terminal tail. Asp 184 was also found to move with the small lobe by the rigid domain analysis. Glu 170 and Asn 171 combine with Asp 184 to form part of the ATP binding pocket. It is particularly interesting that this correlation persists in the dynamics calculation in the absence of ATP. The C-terminal tail was also shown, as expected from the structure, to be part of domain 1.

The second strongest mode correlates well with domain 2. All of the major peaks in domain 2 (Figure...
3b) match peaks in the corresponding eigenvector. However, there are also several parts of the eigenvector in the small lobe which do not correspond clearly to any of the parts of either domain. These may represent additional correlations in the small lobe which have not yet been identified.

It is interesting to note that nearly all of the points in the plots of the eigenvectors where there are major changes of slope or sharp transitions between positive and negative regions occur at the boundaries of secondary structural elements. The only exception to this is in the C-terminal tail.

**CONCLUSIONS**

Conformational variability of the active form of cAPK has been demonstrated experimentally in several crystal structures. The molecular dynamics study reported here has demonstrated important features of the conformational states available to this molecule. The obvious domain motions expected from the bilobal crystal structure have been shown to contain important details which are closely related to the functional properties of the protein. These details have been confirmed by four separate methods of analysis.

Conservation of interatomic distances shows that residues in the region of 98–105 in the small lobe tend to be correlated with movement of the large domain (Figure 3b), while the highly conserved Asp 184 in the large lobe is correlated with the small domain (Figure 3a) despite the fact that its immediate neighbors Phe 185 and Gly 186 are correlated with the large domain (Figure 3b). These cross-lobe correlations are

**FIGURE 4** Residues included in the first (red) and second (blue) rigid domains. Residues not included in the domains are shown in yellow. The N and C termini of the chain are marked, and residue numbers are provided to aid in correlation between the overall structure and the text.
surprising because the region from about 120 to 126 is a clear candidate for a hinge point between domains, and the obvious presumption from the crystal structure would be that the two domains would act independently. Analysis of the matrix of correlation coefficients further showed that the motion of Glu 170 in the catalytic loop of the large lobe is also correlated with the motion of Asp 184 and the small lobe. Asp 184, Glu 170, and Asn 171 form part of the ATP binding pocket; most of the rest of the pocket is made up of residues from the small lobe. These correlations are detected even in the absence of ATP.

Analysis of variability of dihedral angles and conservation of interatomic distances also showed that the crucially important activation loop (residues 193–196) does not correlate with either domain. As expected, most of the variability shown by this measure is at points between secondary structure elements. This view of the architectural basis of the molecular flexibility was confirmed by the eigenvectors of the vector correlation matrix, which show that the changes in sign of correlation all occur at boundaries of secondary structural elements. This result, while not unexpected, also showed that the loop motions show patterns of correlation and anticorrelation.

Analysis of persistent hydrogen bonds confirmed the preservation of secondary structure and the domain structure of the protein but also showed an interesting conformational switch that may have functional implications. Lys 47, Thr 48, and Arg 56 are at the boundaries of the glycine-rich loop. This loop binds the phosphate moieties of ATP and must move for ADP to be released. Analysis of the hydrogen bonds during the dynamics simulation showed that these residues at the base of the loop are persistently coupled to acidic residues in the C-terminal tail. This suggests a possible mechanism for coupling the opening of the glycine loop with mobility of the C-terminal tail.

This dynamics simulation, performed without ATP or phosphorylated amino acids, has demonstrated a remarkable coherence of structural features known to have functional significance. The activation loop, the glycine-rich loop, the domain structure, the ATP binding site, and the catalytic loop are all features that have been identified as persistent by several analysis techniques. These conclusions are further supported by chemical studies which defined not only the “malleability” of the C-terminal tail but also the interactions between Lys 72 in the small lobe and Asp 184 in the large lobe.

METHODS

The AMBER version 4.1 molecular dynamics program was used for calculations on the Cray C90 and Intel Paragon supercomputers of the San Diego Supercomputer Center (SDSC). The calculations utilized the all-atom force field with a single nonbonded cutoff of 12 Å for energy minimization and molecular dynamics. The protein was positioned in a box of 5958 TIP3P water molecules. The closest position of a water’s oxygen to the protein was 2.4 Å. The simulations were performed at constant volume. The whole system was loosely coupled from 0 K to a heat bath of 300 K with the time constant 0.2 ps. The bond lengths of hydrogen atoms were constrained using the SHAKE algorithm.

The crystallographic structure of cAPK (PDB ID code 1ATP) was modified by substitution of all three phosphorylated amino acid residues with the unphosphorylated analogs. As a first step toward understanding how changes in the residues including phosphorylation and possible mutations can affect the dynamics of the C-subunit, it was necessary to carry out the first set of calculations with the structure consisting of a fully parametrized standard set of amino acids, so the difference impacted by all possible changes could be further defined. The inhibitory peptide PKI (5–24), ATP, and crystallographic water molecules were also eliminated from the structure. These calculations, including only the protein molecule, are a foundation for further exploration of the dynamic features of C-subunit structure.

The modified sidechains were subjected to energy minimization in vacuum. The protein was placed in a large
water box (75 × 60 × 60 Å). The system containing the protein and water was then energy minimized for 2000 iterations. Water molecules were then energy minimized for 2000 iterations while the protein was held fixed. Molecular dynamics of waters only was then performed for 25 ps followed by energy minimization of the whole system for 2000 iterations. After this preparation molecular dynamics of the whole system was run for about 600 ps. Snapshots of all parameters were saved at 0.5 ps intervals.

Statistics of dihedral angles during the run were calculated using the program CARNAL, which is included in the AMBER 4.1 package. Calculations of rigid domain stability during molecular dynamics were performed using the difference-distance matrix\(^1\) between corresponding \(\alpha\) carbons of the protein at different times of dynamics. We extracted the rigid domains for pairs of conformations of cAPK taken from 90 to 600 ps at 20 ps intervals. Thirty-six pairs were calculated. The frequency of occurrence of a residue within the same rigid domain throughout all pairs was used as one indicator of correlated motion.

The program H-plus\(^2\) was used to find the hydrogen bonds in the protein. Snapshots from 100 to 610 ps were used, and hydrogen bonds were calculated for each of them. The frequency of occurrence of each hydrogen bond throughout the dynamics run was calculated as another indicator of persistence of conformation.

To calculate correlated motions of residues for 100–610 ps we used the vector correlation coefficient

\[
C_{ij} = \frac{\langle r_i \cdot r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{\langle (r_i \cdot r_i) \rangle (r_j \cdot r_j)}}
\]

where \(r_i\) and \(r_j\) are vectors defining positions of residues \(i\) and \(j\) as the average positions of their backbone or heavy atoms. The inner products are summed over all coordinate snapshots in the specified time interval. The correlation matrix so defined was then diagonalized. The eigenvalues identify modes of motion and the eigenvectors show which sets of atoms have correlated motion.

The authors thank Drs. J. Andrew McCammon and David A. Case for helpful discussion and the reviewers for several excellent suggestions.

REFERENCES

3. Adams, J. A.; Taylor, S. S. Energetic limits of phosphotransfer in the catalytic subunit of cAMP-dependent

**FIGURE 6** Correlated motion within cAPK as shown by analysis of the vector correlation coefficients between pairs of residues throughout the simulation. The correlation matrix was diagonalized and shown to have two dominant eigenvalues. The eigenvectors corresponding to these eigenvalues show the correlated motions associated with the two modes, as described in the text.