Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 484–491, January 1998
Biochemistry

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on April 30, 1996.

Conserved water molecules contribute to the extensive network of interactions at the active site of protein kinase A

(conformational malleability/substrate recognition/protein crystallography)

SHMUEL SHALTIEL*, SARAH COX†‡, AND SUSAN S. TAYLOR†§

†Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, School of Medicine, University of California, San Diego, CA 92093-0654, and *Department of Biological Regulation, The Weizmann Institute of Science, Rehovot 76100, Israel

Contributed by Susan S. Taylor, November 19, 1997

ABSTRACT **Protein kinases constitute a large family of regulatory enzymes, each with a distinct specificity to restrict its action to its physiological target(s) only. The catalytic (C) subunit of protein kinase A, regarded as a structural prototype for this family, is composed of a conserved core flanked by two nonconserved segments at the amino and carboxyl termini. Here we summarize evidence to show that (*i*) the active site consists of an extended network of interactions that weave together both domains of the core as well as both segments that flank the core; (*ii*) the opening and closing of the active site cleft, including the dynamic and coordinated movement of the carboxyl terminal tail, contributes directly to substrate recognition and catalysis; and (*iii*) in addition to peptide and ATP, the active site contains six structured water molecules that constitute a conserved structural element of the active site. One of these active-site conserved water molecules is locked into place by its interactions with the nucleotide, the peptide substrate/inhibitor, the small and large domains of the conserved core, and Tyr-330 from the carboxyl-terminal "tail."**

The catalytic (C) subunit of cAMP-dependent protein kinase A (PKA), discovered in 1968 (1), is regarded structurally as a prototype for the protein kinase family (2). These enzymes share a considerable degree of homology not only in the sequence of their common kinase core (residues 40–300 in PKA) (3) but also in the three-dimensional structure of the core (4, 5). This core, shown in Fig. 1, consists of two domains with a spatial juxtaposition of conserved amino acids at the active site cleft. The smaller amino-terminal domain is associated primarily with the binding of ATP (6–8), whereas the larger domain, associated with recognition and binding of peptide substrates (5), and contains most of the residues that contribute directly to phosphoryl transfer. ATP is buried at the base of the cleft, whereas the surface of the large lobe at the outer edge of the cleft provides a stable surface on which the peptide docks. Although the core is conserved in all protein kinases, the segments that flank the core, namely the "head" (residues 1–39 in PKA) and the "tail" (residues 301–350) are not (9). These flanking segments seem to have unique structural and functional assignments in each kinase and can play an important role in determining specificity, state of activation, and cellular localization. In PKA, the "head" accommodates an amino-terminal myristyl moiety followed by an A-helix (residues 10–30). This helix provides a complementary scaffold on which the kinase core is docked. Interactions with the core are mostly hydrophobic with Trp-30 inserted into a deep hydrophobic pocket between the two lobes (10). This A-helix helps to stabilize the enzyme and to position the C-helix in the small domain so that it is properly oriented for catalysis at the cleft interface (11).

In contrast to the "head", the C-terminal tail wraps as an extended chain around both lobes of the core (Fig. 1). The predicted flexibility of at least a portion of this tail was based on the following solution studies. In its unliganded state the tail was found to be susceptible to proteolysis by a kinase specific membrane protease (KSMP) (12–14), was modified at Cys-343 by sulfhydryl-specific reagents (15–17), and was labeled at 328–335 by a water-soluble carbodiimide (18). In contrast, the enzyme was protected in the presence of ATP and/or when it was part of a holoenzyme complex (15, 16, 18, 19). Additional evidence illustrating the conformational "malleability" of the tail in free C was based on salt-induced sensitivity of the sulfhydryl groups to covalent modification (15, 17, 19). The conformational flexibility of the tail subsequently was confirmed by x-ray crystallography, which revealed a looser, more "open" conformation and a tightly packed "closed" conformation (20, 21). Identification of the specific cleavage site for KSMP also confirms the earlier predictions for conformational flexibility of the tail (22). Flexibility of the tail in free C also is consistent with low angle neutron scattering (23) and with ligand-induced changes observed by circular dichroism (24, 25).

As shown in Fig. 2, the nonconserved tail is comprised of four segments. The first segment (residues 301–315), firmly anchored to the large domain (21), is followed by an extended segment (residues 317–326), which is anchored to the large lobe in the closed conformation but disordered in the open conformation (20, 21). The third and fourth segments (residues 327–335 and 336–350, respectively) embrace the small domain. The acidic cluster surrounding Tyr-330 (residues 327–335) is anchored onto the small lobe in the closed conformation, but is more exposed to solvent in the open conformation consistent with its availability to proteolysis (12–14) and modification (18). The last segment (residues 336–350) forms a hydrophobic knot that is anchored within the small lobe in both conformations. Ser-338, one of two stable

Abbreviations: C, catalytic subunit; rC, recombinant C subunit; mC, mammalian C subunit; KSMP, kinase specific membrane protease; PKA, protein kinase A; PKI, protein kinase inhibitor; PKS, substrate analog of PKI (5–24); ASCW, active-site conserved waters; LSCW, ligand-specific conserved waters.

‡Present address: DuPont Merck Pharmaceutical Company, Experimental Station Wilmington, DE 19880-0336.
§To whom reprint requests should be addressed. e-mail: staylor@ucsd.edu.

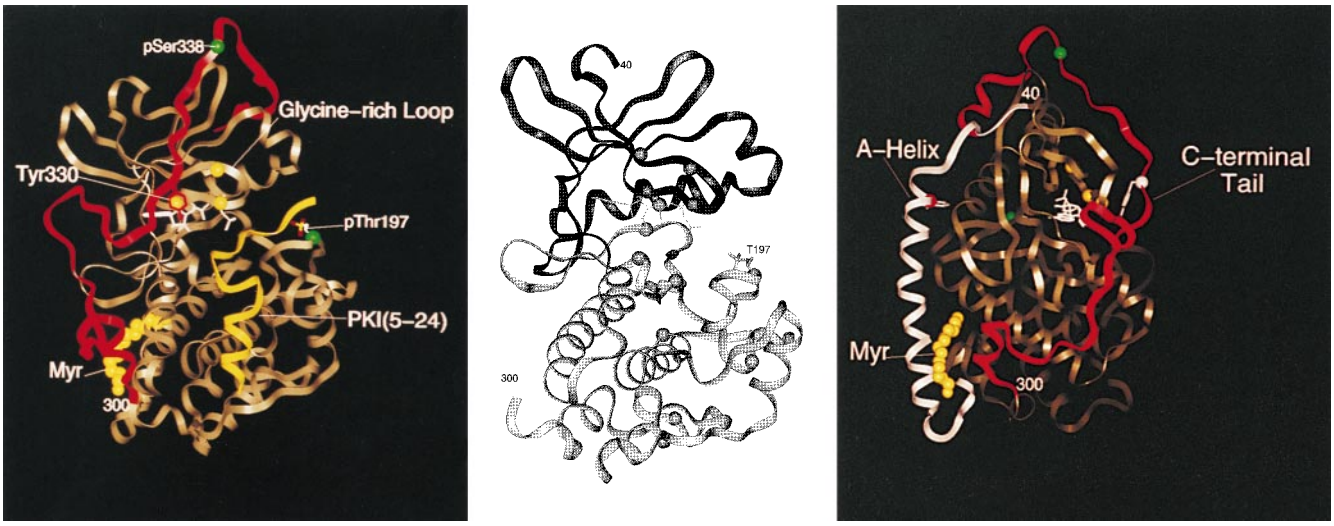

FIG. 1. Ribbon diagram of the catalytic subunit. The conserved core (*Center*) highlights conserved residues (G50, G52, G55, K72, E91, D166, N171, D184, E208, D220, and R280). The small lobe (residues 40–120) is dark. The nonconserved "head" (residues 1–39; white) and "tail" (residues 300–350; red) are highlighted (*Left* and *Right*). The myristyl moiety is yellow, and phosphates (Ser-338 and Thr-197) are green. Side chains of Trp-30 and Tyr-330, two prominent aromatic residues that interact with the core, are shown. The core is tan (*Left* and *Right*) with Gly-50, -52, and -55 shown as yellow balls. *Center* and *Left* show the same view. (*Right*) The view is rotated counterclockwise approximately 90°.

phosphorylation sites in C, and Cys-343, one of two cysteines in C, both are in this segment.

We recently have shown that the flexibility of the tail may well have functional importance, not only by serving as a gate for the nucleotide (26) but also by contributing to substrate recognition and catalysis (27). The particular importance of Tyr-330 for recognition of substrates and inhibitors is based on three complementary lines of evidence: (*i*) cleavage and inactivation at this site by KSMP (13, 22); (*ii*) analysis of the crystal structures of closed vs. open conformations (21); and (*iii*) single-site mutations of Tyr-330, which showed not only a significant increase in the K_m for peptide but also an 1,100-fold decrease in catalytic efficacy (V_{max}/K_m) (27).

Because of the potential importance of Tyr-330, we undertook a comparative analysis of seven different crystal structures of C, representing closed, open, and intermediate conformations. On the basis of this analysis, we conclude that C contains at its active site, in addition to MgATP and the peptide substrate, a set of conserved water molecules. One of the most prominent of these water molecules in the closed conformation interacts with Tyr-330 whereas the others interact with the MgATP and the conserved residues at the active site. Thus Tyr-330, in spite of its location in the nonconserved carboxyl-terminal tail, is a direct part of the network of interactions that extends outward from the active site cleft and links the active site to distant parts of the molecule.

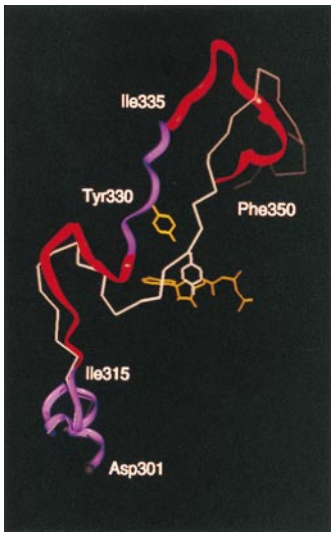

FIG. 2. C-terminal tail in closed and open conformations. The four segments of the tail in the open conformation are shown in purple (301–315), red (316–326), purple (327–335), and red (336–350). The conformation of this segment in the ATP:PKI(5–24) ternary complex is shown in white. ATP is yellow. Tyr-330 is highlighted. Adapted from ref. 26.

EXPERIMENTAL PROCEDURES

Structural Data from X-Ray Crystallography. Seven different crystal structure models of C, summarized in Table 1, were used for the comparison. These models include the closed conformation of the ternary complex containing recombinant (r)C, Mn^{+2}ATP, and heat stable protein kinase inhibitor [PKI(5–24)] (I) (7), the binary complex formed between rC and PKI(5–24) (V) (28); the closed conformation of the ternary complex of rC with adenosine and PKI(5–24) (III) (29), the open conformation of the binary complex formed between mammalian (m) C and iodinated PKI(5–22) (VII) (20, 21), and the closed conformation of mC with AMP-PNP and PKI(5–24) (VI) (8). Structures I, II, IV, V, VI, and VII were crystallized under identical conditions although VI was crystallized independently in a different laboratory. Structure III was crystallized under different conditions (29).

Differences in the method of refinement results in differences in the distribution of B-factors. The five structures showing a closed conformation are directly comparable in this

Table 1. Crystal structures used for comparison

No.	Complex	Ref.	Resolution	O/C	Color
I	rC:ATP:PKI(5–24)	(7)	2.2 Å	C	Red
II	rC:ADP:PKS(5–24AS)	(30)	2.25 Å	C	Purple
III	rC:adenosine:PKI(5–24)	(29)	2.2 Å	C	Yellow
IV	rC:PKS(5–24AS)-(P)	(30)	2.2 Å	C	Blue
V	rC:PKI(5–24)	(28)	2.0 Å	C	Turquoise
VI	mC:AMPPNP:PKI(5–24)	(8)	2.0 Å	C	
VII	mC:PKI(5–22)*	(6)	2.9 Å	O	

Closed and open conformations are designated as C and O. The colors codes are used for Figs. 4 and 7. PKS, protein kinase substrate. *This peptide was iodinated on Tyr-11.

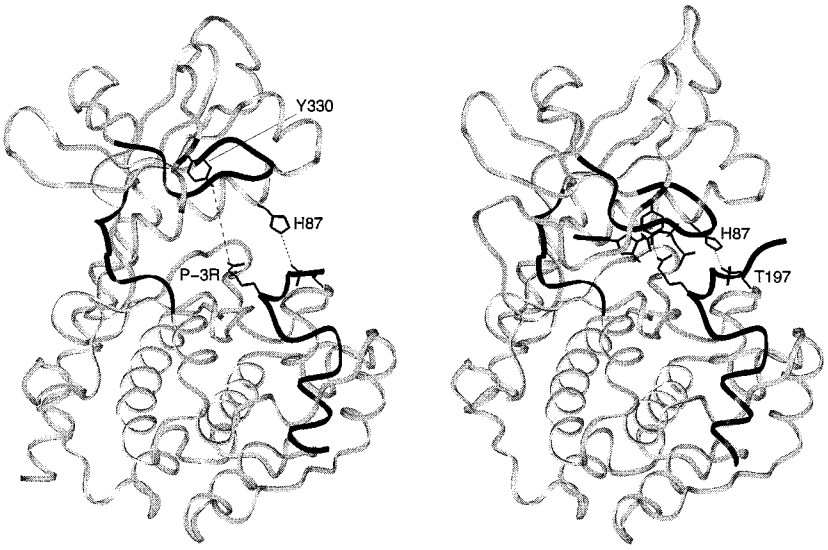

FIG. 3. Open and closed conformations of the C subunit. The binary complex of mC and iodinated PKI(5–22) adopts an open conformation (*Left*). The ternary complex, rC:Mn2ATP:PKI(5–24) is closed (*Right*). The distances between the His-87 side chain and Thr-197 phosphate and between the Tyr-330 and the P-3 Arg of the inhibitor peptide serve as quantitative indicators for the degree of "openness." The glycine-rich loop (residues 47–56), the linker strand (residues 121–127), and PKI(5–24) are highlighted in black.

respect. Initial refinements were carried out by using the X-PLOR program. Final refinement was carried out by using TNT with unrestrained B-factors. The open binary complex used for these analyses was refined by using TNT and FRODO (20). In the final refinement, regions of the molecule that could not be located with confidence were omitted from the model and the refinement parameters restricted B-factors to >16 Å^2.

Structures were superimposed by using backbone atoms of five residues, 144, 168, 208, 244, and 280 in the large lobe. The conserved water molecules identified here were within a distance of <1.5 Å.

RESULTS

Open and Closed Conformations of the Subunit: Criteria for Classification. In the crystalline state, at least two distinct conformational states of C (denoted "open" and "closed"), have been observed (20, 21). They differ in the relative orientations of the small lobe and large lobe, in the extent of opening of the ATP binding cleft between the two lobes, and in the positioning of the carboxyl-terminal tail (20, 26, 27). This conformational change is best characterized by three quantitative molecular parameters:

(*i*) A regression of the glycine-rich loop, which in the open conformation moves away from the larger lobe (Fig. 3). In the fully closed conformations this loop forms an amide bond between the backbone amide of Ser-53 at the tip of the loop and the γ-phosphate of ATP (8).

(*ii*) A regression of His-87 in the small lobe (located at the beginning of the C-helix), away from phospho-Thr-197, (the essential phosphorylation site in the activation loop of the large lobe), by about 3 Å. This regression weakens a key electrostatic/hydrogen bonding interaction between the two lobes.

(*iii*) A regression of Tyr-330, within the cluster of acidic amino acids (^{328}DDYEEEE^{334}) in the carboxyl-terminal "tail," which in the open conformation moves approximately 7 Å away from a key highly conserved biorecognition element of the peptide substrate, the p-3 Arg (27).

Structural Malleability of the Various Structures of C as Reflected by their Temperature (B) Factors. To gain additional insight into the functional role(s) of this conformational change, we carried out a comparative analysis of the structures of C solved so far in our own and other laboratories (4, 5, 7, 8, 20, 21, 28–30). Our conclusions were further validated with one additional structure, which was completed while this

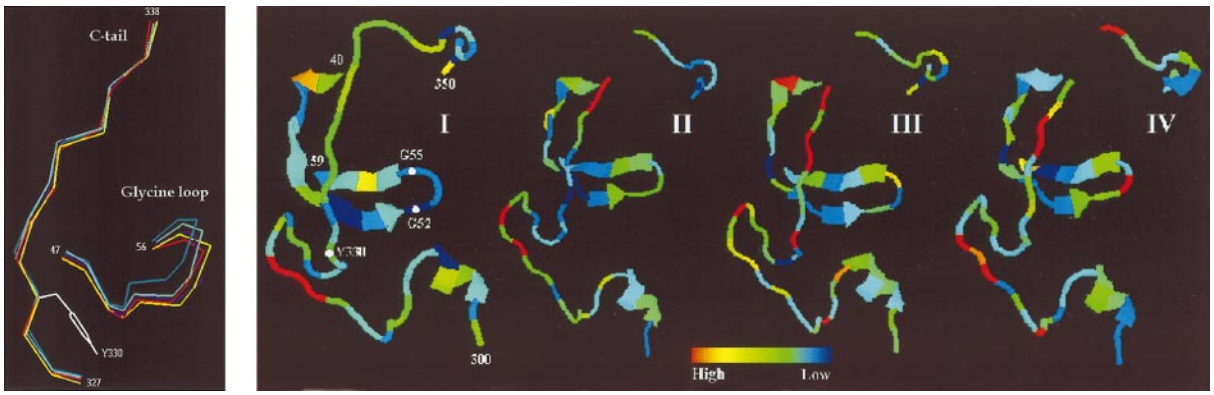

FIG. 4. Temperature (B) factors for four different crystal structures of the C subunit. (*Left*) The Cα trace of the five superimposed structures. Tyr-330 also is shown. Only the tail (residues 301–350) and the glycine-rich loop (residues 47–56) are shown. (*Right*) A ribbon representation of the Cα-trace for four of these structures, colored according to temperature factors. The scale for the colors is also shown. The structures shown are: I (rC:Mn_2ATP:PKI(5–24);red); II (rC:ADP:(P)PKS; purple); III (rC:adenosine:PKI(5–24); yellow);IV (rC:PKS; blue);V: binary complex (rC:PKI(5–24); turquoise).

analysis was being carried out, an adenosine:PKI(5–24) ternary complex with rC (29). Because these structures all represent fully phosphorylated, active forms of C, the observed conformational differences likely represent physiologically important changes that are associated with catalysis.

Although the closed conformation obtained with ATP and PKI(5–24), can be described as a compact, tight structure, the more open conformation obtained with iodinated PKI(5–22) reflects a more malleable (loose) structure both in solution and in the crystalline state. The crystallographic temperature factors can be taken as an indicator of such malleability. In solution there very likely are several open structures loosened to a different extent. The B-factors for the glycine-rich loop and the C-terminal "tail" are indicated in Fig. 4.

As a first step, we examined the quality of the crystallographic data. Four of the structures were solved to a resolution greater than 2.3 Å and were refined by using similar methods. Thus, the data was directly comparable. Because the mammalian binary complex (20, 21) was solved to a resolution of only 2.9 Å, and different procedures were used to refine the coordinates, it was not included in this comparison.

The glycine-rich loop that serves to position the phosphates of ATP shows a wide range of temperature factors and its orientation also varies considerably in the different structures (29). Three conserved loops actually converge at the active site cleft: (*i*) the glycine-rich loop, (*ii*) the catalytic loop, and (*iii*) the Mg positioning loop. Only the glycine-rich loop is mobile (26). The tip of the loop is stable only in the ATP and PKI(5–24) ternary complex.

In the open conformation, the positioning of the glycine-rich loop and the carboxyl-terminal tail result in a considerable ease of access to the active site cleft. The side chain of Tyr-330 is on the surface of the protein and lies across the ATP binding pocket in the closed conformation, forming a sort of a "lid" for this pocket. However, in the open conformation Tyr-330 swings away with the side chain hydroxyl moiety becoming solvent exposed. In all the structures examined, this Tyr-330 side chain is well ordered, as reflected by its low temperature factors.

Functional Groups in C Found in the Vicinity of its Substrate and Inhibitor Recognition Elements. Initially, the specific importance of Tyr-330 was based on the sensitivity of this region to cleavage by KSMP, which resulted in loss of activity (14), and characterization of the specific KSMP cleavage site (22) confirmed the prediction. Subsequent crystal structures confirmed the proximity of Tyr-330 to the active site. It was, however, the kinetic consequences of replacing Tyr-330 (27) that suggested that the role of this residue may not be merely to provide a biorecognition site for the peptide substrate by forming a hydrogen bond with the p-3 Arg in the substrate peptide. In addition, Tyr-330 appeared to be contributing to catalysis. We presumed that this catalysis also could involve other structural elements such as resident water molecules at this site, through which this Tyr residue might interact with additional groups participating in substrate binding and/or catalysis. We, thus, searched for conserved water molecules that interacted specifically with ATP, peptide, and conserved amino acids at the active site cleft.

Fig. 5 highlights some of the functional groups in C that contribute to recognition of substrate and inhibitor peptides. The consensus site is defined as the P-3 through P+1 residues, and its direct interactions with the protein are extensive. In the ternary complex with ATP and PKI(5–24), Tyr-330 approaches the Nε atom of the P-3 Arg residue to a distance of ≈3 Å. The same Nε atom is 2.9 Å from the 2′ hydroxyl of the ribose in ATP. When we examined the immediate surroundings of Tyr-330, we found that there is also a water molecule in this locus that is conserved in most of the solved structures of PKA. We therefore carried out a comparative analysis of the water molecules at the active site of PKA.

Analysis of Water Molecules at the Active Site. Because the placing of water molecules in a crystal structure is dependent on the resolution of the structure and is at the discretion of the crystallographer, the number of water molecules that are built in the structure may vary from one structure to another. In this analysis the structure that has the fewest water molecules (115) built in was the PKS(P) binary complex, compared with 167 for the ADP/PKS ternary complex, 289 for the ATP/PKI ternary

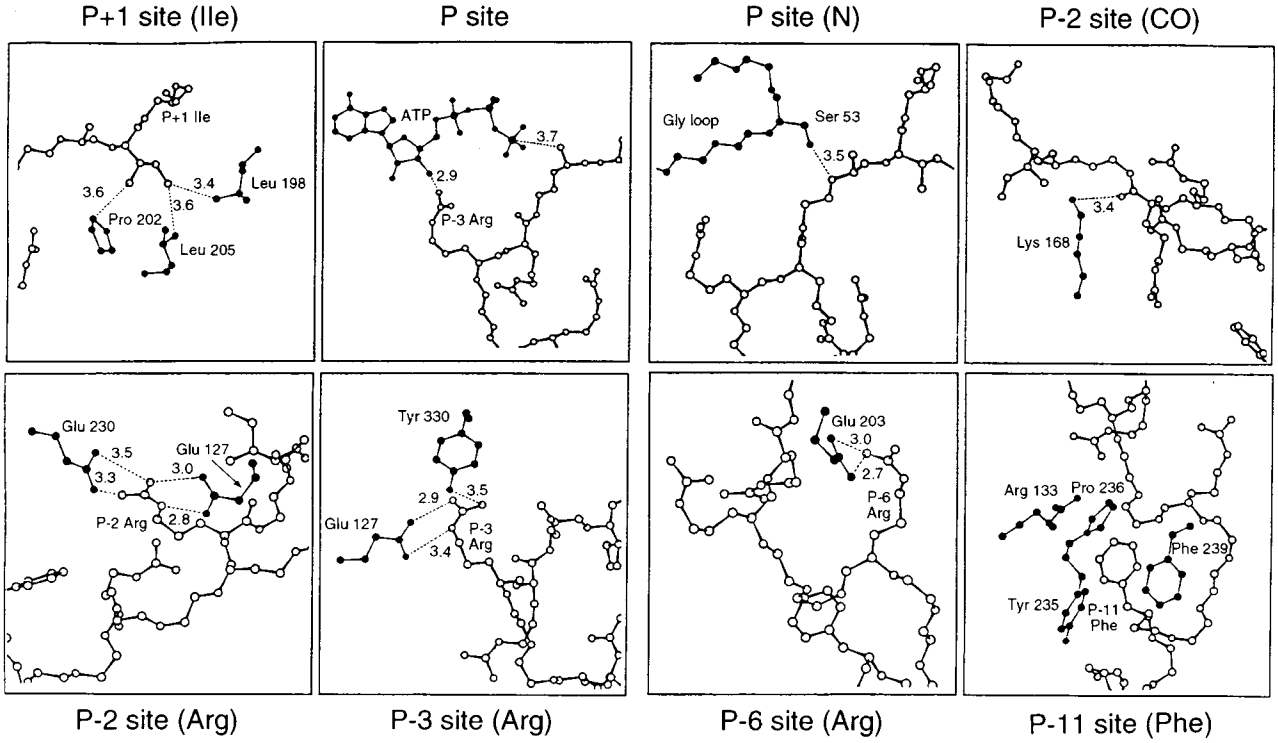

FIG. 5. Specific interactions of the C subunit with PKI(5–24). Distances are taken from the rC:Mn_2ATP:PKI(5–24) complex.

complex, and 534 for the PKI binary complex. On the basis of our survey, we defined two classes of water molecules at the active site: (*i*) active-site conserved waters (ASCWs), which are associated directly with substrates/inhibitors, or with conserved active site residues in the ATP/PKI(5–24) ternary complex and are assumed to be those water molecules that are part of the active complex that participates in catalysis, and (*ii*) ligand-specific conserved waters (LSCWs), which are found only with incomplete complexes such as the adenosine plus peptide complex or with binary complexes. These LSCWs are displaced when substrates bind.

ASCWs. Six conserved water molecules were found in the vicinity of ATP in the ternary complex containing ATP and PKI(5–24) (Fig. 6). The exact positioning of these ASCWs differs slightly when this ternary complex was compared with the other structures. However, when the various structures were superimposed (Fig. 7 and Table 2), these ASCWs were less than 1.5 Å apart. Two of these water molecules make specific hydrogen bonds with the conserved residues Lys-72 and Glu-91. ASCW-a interacts with Glu-91 in every structure except the PKI(5–24) binary complex (V) where it forms a hydrogen bond with Lys-72. ASCW-b forms a hydrogen bond with Glu-91 in every structure except the phosphopeptide-ADP complex (IV). A third ASCW-c is deeply buried in the cleft interface and interacts with the carbonyl backbone of Asp-184.

Two water molecules were associated with the metal binding sites. ASCW-d is close to the α and γ phosphates and comes in close contact with the second metal ion. ASCW-d, however, also is present in every structure and is not dependent on the presence of the metal ion. A fifth ASCW molecule (ASCW-e) is in the vicinity of the primary metal ion locus (Fig. 6 and Table 2). In structures II and VI, which contain Mn and ATP and Mn and AMP-PNP, respectively, there is an additional water molecule. These waters are designated e^1 and e^2. The water structure in this region very likely is influenced by the presence of the metal ions. Both metal ions are hexa coordinated. The primary metal accepts two liganding oxygens from the side chain of Asp-184 and two from the β and γ phosphates, and the coordination is completed by two water molecules, e^1 and e^2. In the four structures where the metal ions are absent there is a single water molecule, ASCW-e, which interacts with the side chain of Asp-184 and occupies a position within the region defined by the triad comprised of the metal and the two liganding waters. The second metal also has six ligands: one from the side chain of Asp-184, one from the side chain of Asp-171, three from the phosphate and the coordination is completed by a water molecule. In this case, a water molecule sits close to the site of the liganding water in all the structures, irrespective of the presence of the nucleotide. With the exception of the binary complex containing C and the phosphorylated substrate analog of PKI(5–24), these ASCWs were seen in all of the structures.

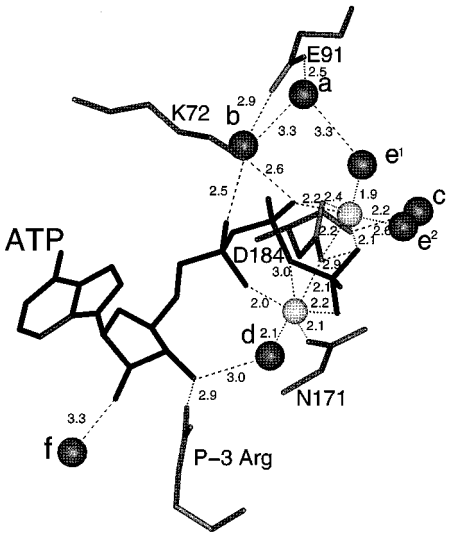

FIG. 6. ASCWs. The specific interactions made by the ASCWs (a–f) at the active site cleft in the ternary complex of rC, MnATP, and PKI(5–24) are summarized. The numerical designations for the ASCWs in the rC:MnATP:PKI(5–24) pdb file are as follows: a (597), b (400), c (477), d (635), e (447), and f (423).

Table 2. ASCWs

ASCW	Nearest neighbors	Atom	Distance	B-factor	I	II	III	IV	V	VI
a	Lys-72	NZ	3.3	20.1	●	●	●	○	●	●
	Glu-91	OE1	2.5	7.4						
b	Glu-91	OE2	2.9	8.2	●	●	●	●	●	●
c	Asp-184	O	2.6	24.0	●	●	●	●	●	●
d	Mn2	Mn	2.1	21.0	●	●	●	●	●	●
	ATP	O2G	2.9	18.0						
	ATP Ribose	3′OH	3.0	33.8						
e^1	Asp-184	OD1	3.0	11.0	●					●
	ATP	O1B	2.9	13.1						
	Mn1	Mn	1.9	23.7						
e^2	Asp-184	OD2	3.2	11.9	●					●
	ATP	O3G	3.1	6.6						
	Mn1	Mn	2.2	23.7						
f	ATP Ribose	2′OH	3.4	48.5	●	●	●	○	●	●
	Tyr-330	OH	2.7	25.9						
	Leu-49	O	2.7	15.2						
	Glu-127	OE1	3.6	25.8						

The nearest neighbors to each water molecule are indicated, including residue, atom, distance, and the B-factor for this atom. Distances and temperature factors are taken from the C:ATP:PKI(5–24) ternary complex. The metal ion and two waters (e^1 and e^2) in the ternary complexes with ATP or AMPPNP are replaced by one water (e) in the other structures. Filled and empty circles indicate the presence or absence, respectively, of the water molecule in each structure.

In addition to these five ASCWs, which interact with ATP and/or with conserved residues at the active site cleft, a sixth conserved water molecule was found. This ASCW-f molecule is at a hydrogen bond-forming distance from the hydroxyl of Tyr-330, the side chain of Glu-127, and the 2′OH of the ribose ring of ATP (Figs. 6 and 8). The importance of the ribose 2′OH group in the binding of ATP to C originally was observed in a study mapping the ATP binding site of this kinase (31). This study showed that on replacing the 2′OH by H, the K_i for ATP increased from 9 to 38 μM (>4-fold). However replacement of the 3′OH by H decreased the K_i from 9 to 1.5 μM (6-fold). In other words, if we compare the relative affinity of two ATP analogs from which the hydrogen bond-forming oxygen in the ribose moiety has been removed, then the 2′H analog is ≈25-fold less efficient as a competitive inhibitor of ATP for C than the 3′H analog (31). Interestingly, the opposite is true when the high-affinity ATP binding to the type I holoenzyme was compared for these two analogs. In this case, the 2′OH binds more tightly than the 3′H analog. The importance of this ASCW denoted (f) was further confirmed by the finding of a water molecule in the same location in a recently solved new crystal structure where C is complexed with a natural product inhibitor, balanol (N. Narayana, personal communication).

As seen in Fig. 8, this conserved water molecule is at a key position in the molecule, from which it can bridge and coordinate a variety of interactions involving: (*i*) ATP (through the 2′OH of the ribose); (*ii*) the small lobe of the enzyme (through the carbonyl oxygen of Leu-49 in the functionally important glycine-rich loop); (*iii*) the large lobe of the enzyme (through the carboxylate of Glu-127); (*iv*) the C-terminal "tail"

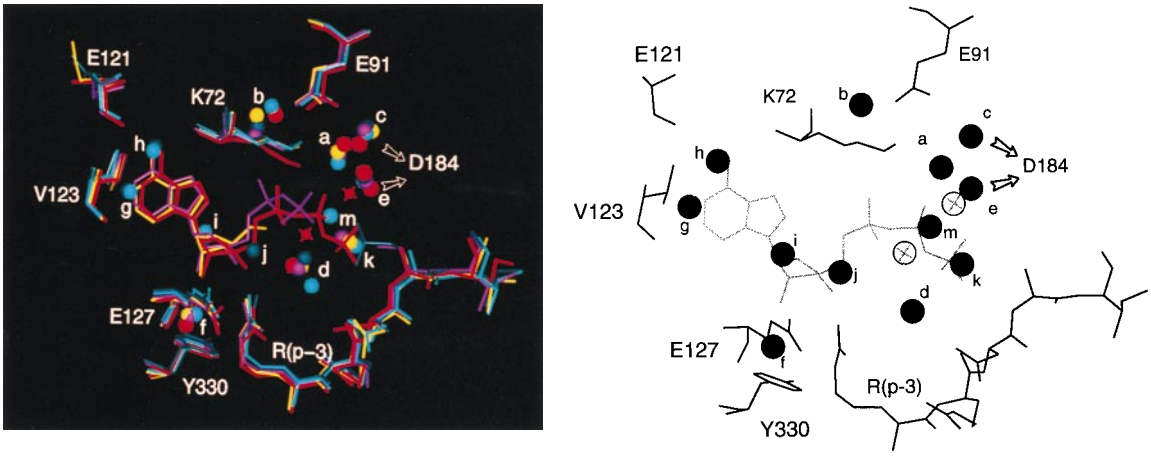

FIG. 7. Summary of the conserved water molecules in the active site cleft. (*Left*) The conserved active site region of five different closed conformation crystal structures (I–IV) are superimposed. ASCWs present in all (or most) of the structures (a–f) and LSCWs that are present when all or part of one of the ligands is absent (g-m) are shown. The residues Glu-91, Lys-72, Glu-127, Glu-121, Val-123, and Tyr-330 from the enzyme and a section of the inhibitor peptide, residues PKI(18–22), and water molecules in the nucleotide binding pocket from the five structures are color-coded according to Fig. 5 and Table 1. (*Right*) Summary of the positions of the ASCWs and LSCWs.

(through the phenolic hydroxyl of Tyr-330); and finally (*v*) the peptide inhibitor PKI (through the guanidino group of the Arg residue at the p-3 position of the inhibitor). This locus is thus a center for a network of interactions. Although the water molecule does not seem to mediate all of them, direct interactions (e.g., electrostatic or hydrogen bonding) are possible between the participants in this network.

LSWMs. ATP can be described as consisting of three parts: the adenine ring, the ribose ring, and the triphosphate moiety (29). ASCW molecules were found to be associated with each of these subsites. However, the presence of some of these water molecules depends on the presence the ligands or, more specifically, on a particular functional group of the ligand.

In both the binary complexes of C, which have a peptide ligand but no nucleotide, water molecules replace the two groups of the adenine ring, which, when a nucleotide is present, interact specifically with the protein. One of these water molecules (h) resides at the locus occupied by the N6 nitrogen of the adenine ring, which normally forms a hydrogen bond with the carbonyl of Glu-121 (Table 3). The other water molecule (g) resides at the locus occupied by the N1 nitrogen of the adenine ring, which normally forms a hydrogen bond with the imido group of Val-123. These two water molecules thus can be regarded as conserved in the structures that do not contain the adenine ring. They are present at these loci both in the binary C:PKI complex and in the C:PKS(P) complex, and can make equivalent hydrogen bonds. It is interesting to note that most of the synthetic or natural product protein kinase inhibitors that have been characterized structurally bind to the active site cleft so that a portion of the inhibitor fills this adenine binding pocket and forms these two hydrogen bonds to the backbone of the linker strand that joins the two lobes (32).

Additional LSCW molecules were found in the C:PKI binary complex. One such water molecule (i) was found in the locus otherwise occupied by the 2′OH of the ribose. In the binary complex with PKS(P), two water molecules were found in the place of the ribose, one at the position of the 2′OH (i) and the other (j) at the position of the 3′OH of the ribose. Again, it is possible that the absence of this water molecule (j) from the PKI binary complex is simply because of its omission in solving the crystal structure. However, because this is the structure with the greatest number of water molecules built in, the absence of this second water molecule may indicate that it is not a tightly bound LSCW molecule. The water molecules vicinal to the triphosphate chain have several features of interest. First, in the binary complex rC:PKI there are two LSCW molecules, which lie in the locus accommodating the β and γ phosphates. In the binary complex C:PKS(P) these water molecules are not present. However, the phosphate of the phosphorylated Ser of the peptide occupies a position very close to that which otherwise would be occupied by the γ-phosphate of ATP. In the ternary complex of C with ADP and PKS there is a water molecule replacing the γ-phosphate.

DISCUSSION

The role of water molecules in proteins, and especially their possible involvement in bioligand recognition and in the catalytic function of enzymes, has been debated for many years (33). Completely buried water molecules were found at conserved sites in structurally homologous proteins, suggesting that they have a distinct role. In some cases such water molecules were found to be associated with functional groups at the active site (34), in other cases they were suggested to act as a "proton wire" during catalysis (35), or to be present at the interface of protein folding domains and involved in interdomain motions (36). Conserved water molecules were studied recently in several enzymes, e.g., in serine proteinases (37), in ribonuclease T1 (38) and A (39), and in triose phosphate isomerase (40).

Our interest in identifying conserved water molecules at the active site of the C subunit of PKA, stemmed from our systematic mapping of the interactions between the carboxyl-terminal tail of the catalytic subunit of PKA, and the active site

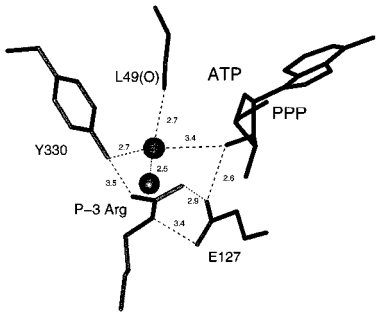

FIG. 8. Interaction of Tyr-330 with ASCW-f. The specific interactions made by ASCW-f with Tyr-330, ATP, peptide, and the small and large lobes are shown. The distances are from the rC:MnATP:PKI(5–24) ternary complex. A second water molecule also is indicated.

Table 3. LSCWs

LSCW	Nearest neighbors	Atom	Distance	B-factor	I	II	III	IV	V	VI	Remarks
g	Val-123	N	3.0	18.7	○	○	○	●	●	○	Replaces adenine N1
h	Glu-121	O	3.0	19.1	○	○	○	●	●	○	Replaces adenine N6
i	Glu-127	OE2	3.5	19.2	○	○	○	●	●	○	Replaces ribose 2′OH
j	Thr-51	O	3.3	48.3	○	○	○	●	○	○	Replaces ribose 5′OH
k	Lys-168	NZ	2.8	8.5	○	●	●	○	●	○	Replaces ATp (γ) P or Ser-P
m	ASCWa	OH2	3.3	—	○	○	○	○	●	○	Replaces ATP (β) P

Distances and temperature factors are taken from the C:PKI(5–24) binary complex except for j, which is taken from the C:PKS(5–24AS)-(P) binary complex. Filled and empty circles indicate the presence or absence, respectively, of the water molecule in each structure.

of this enzyme. While comparing these interactions in the various solved structures of this enzyme (with and without substrates or their analogs), we noticed the persistent residence of water molecules at the active site and decided to define their positions and assess their possible role in the enzyme function. The resulting analysis of these structures leads us to conclude that (*i*) the open and closed conformations are functionally important for catalysis; (*ii*) that there is an extended network of interactions at the active site that links together not only the two domains but also the amino and carboxyl terminal segments that flank the core; (*iii*) that the carboxyl terminal tail, in particular, converges in a dynamic way to recognize and bind substrates, to facilitate a rapid transfer of the phosphate from ATP to the substrate protein, and then to release the products; (*iv*) that there are six structured water molecules that, like the peptide and ATP, constitute a conserved structural element of the active site. Five of these ASCWs coordinate MgATP to conserved residues in the core protein whereas the sixth water molecule (f) is locked into place by its interactions with the nucleotide, the peptide substrate/inhibitor, and the small and large domains of the conserved core and Tyr-330 in the C-terminal "tail."

ASCWs a, b, and c interact directly with one of the conserved residues at the active site cleft. ASCW-b, the most buried, is associated primarily with Glu-91. ASCW-c, also buried, hydrogen bonds to the carbonyl backbone of Asp-184. ASCW-a hydrogen bonds to the side chain of Glu-91. ASCWs a-c reside in the active site both in the presence and absence of substrates. ASCWs d and e interact with Mg ATP and are present in the active site whether or not it is occupied by substrates. ASCW-d interacts with the inhibitory metal ion that bridges the α and γ phosphates of ATP. Whereas ASCW-e interacts with the activating metal ion and to the ATP γ phosphate, ASCW-f resides in the active site cleft only in the closed conformation.

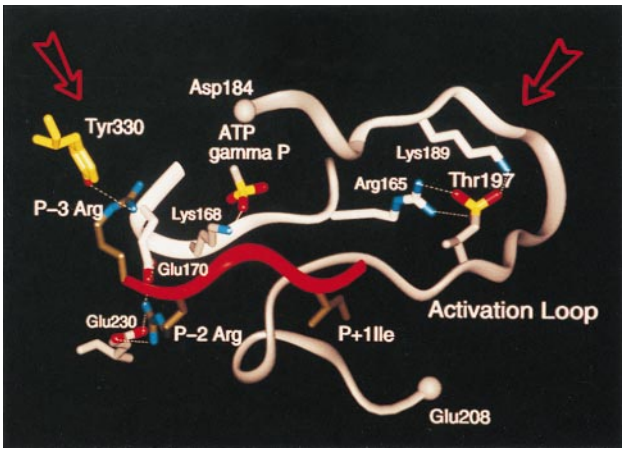

FIG. 9. Extended network of interactions at the active site cleft. Binding of the consensus site peptide and ATP to the active site cleft creates a contiguous network of interactions that extend well beyond the direct site of phosphoryl transfer. The two arrows indicate regions that are influenced by nonconserved parts of the molecule. The left arrow highlights the position of Tyr-330 in the "tail." The right arrow indicates the region that is influenced by the A-helix at the amino terminus. The γ-phosphate of ATP is yellow, the inhibitor peptide, PKI(18–22), is red, and the activation segment from Asp-184 in the DFG loop to Glu-208 is gray with β-strand 9 and the activation loop (residues 165–171) highlighted in white.

The extended network that links these distal regions of the molecule is shown in Fig. 9. Several lines of kinetic evidence confirm the functional importance of this extended network for transfer of the phosphate from ATP to the peptide. Initially, kinetic evidence using rapid quench analysis, demonstrated that the relatively slow k_{cat} (20 s^{-1}) did not represent the true phosphoryltransfer step, k_3, which is 23-fold faster (41, 42). Stopped-flow kinetics using a fluorescent tagged C subunit confirmed this finding and also indicated that the slow step correlates not only with ADP release but also with the conformational changes that allow for release of the nucleotide (43). The distinct conformational flexibility of the C subunit originally observed in solution (17) is thus an integral part of its catalytic mechanism. In addition to these kinetic studies, several site-specific mutations demonstrate that distal residues can indeed effect catalysis. At one end of the extended network is Tyr-330 in the C-terminal "tail." Mutation of Tyr-330 effects not only peptide binding as demonstrated by a decrease in K_m but also phosphoryltransfer as indicated by a significant reduction in k_{cat} (27). A similar long-range effect on k_3 as well as K_m also was seen when Glu-230 was replaced with Gln. At the other end of this extended network is Thr-197. Phosphorylation of this Thr is essential for the correct and stable conformation of the activation loop that is required to support optimal catalysis (44). This phosphate coordinates with the activation loop (Asp-164–Asn-171) via Arg-165, with the Mg positioning loop (Asp-164–Phe-187) via Arg-165 and the backbone carbonyl of Phe-187, with β strand 9 via Lys-189 and in the closed conformation with the C-helix in the small lobe via His-87. In addition, it coordinates indirectly with the nonconserved A-helix at the N terminus via a bridge between Lys-189, Arg-193, and Trp-30. Replacing Thr-197 with Ala leads not only to a 60-fold increase in the K_m for ATP but also to a 200-fold decrease in k_3.

Although the C subunit is one of the simplest and best understood members of the protein kinase family, it represents only one example of how the conserved core is regulated in part by the nonconserved segments that flank it, as well as by other regulatory molecules. Other kinases use different, but equally sophisticated mechanisms for regulating the core. The two recently solved structures of *src* and *hck* provide a strikingly different example of how nonconserved regions that flank the core can contribute to kinase function (45, 46). In these two proteins, the core is preceded by an SH3 domain and an SH2 domain. The short carboxyl-terminus contains an inhibitory phosphate (Tyr-526) whereas the linker strand that joins the SH2 domain to the kinase core is a proline-containing segment. In the absence of the activating phosphorylation, the kinase core of *hck* and *src* is maintained in an inactive state by the flanking domains. (P)Tyr-526 near the carboxyl-terminus specifically binds to its own SH2 domain whereas the proline-rich segment that links the SH2 domain to the kinase core

binds to its own SH3 domain. In this tethered state the core is locked into an inactive conformation. The activation loop is disordered, the C-helix in the small domain is twisted, and conformational flexibility is restricted. In free cdk2, a very small protein kinase, the C-helix and the activation loop are disordered, and the enzyme is not active (47). Binding to cyclin, a necessary event for activation, positions the C-helix for catalysis and orders the activation loop so that it is poised for phosphorylation by a heterologous kinase (48). In the case of mitogen-activated protein kinase, the C-terminal "tail" wraps around the surface of the core that in PKA is masked by the A-helix when the enzyme is dephosphorylated. Phosphorylation of the activation loop at Thr-183 and Tyr-185 causes the carboxyl-terminus to be ordered differently (49). It is now clear that the core-flanking segments in the various kinases can have an important influence on specificity, localization, stability, and regulation. As more structures are solved of protein kinases in both their active and inhibited states, we shall be able to appreciate better the molecular basis for these effects. The active site conserved water molecules described here for PKA, or their equivalents, also may turn out to be important structural and functional components of the active site of protein kinases in general.

We wish to thank Drs. E. Radzio-Andzelm and C. Smith for preparation of the figures. This work was supported by grants from the Minerva Foundation, and by the German–Israeli Foundation for Scientific Research and Development (to S.S.), and by the National Institutes of Health (GM19301) and the American Cancer Society (to S.S.T.). S.S. is the incumbent of the Kleeman Chair in Biochemistry at the Weizmann Institute of Science.

1. Walsh, D. A., Perkins, J. P. & Krebs, E. G. (1968) *J. Biol. Chem.* **243,** 3763–3765.
2. Taylor, S. S., Knighton, D. R., Zheng, J., Ten Eyck, L. F. & Sowadski, J. M. (1992) *Annu. Rev. Cell Biol.* **8,** 429–462.
3. Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) *Science* **241,** 42–52.
4. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Ashford, V. A., Xuong, N.-H., Taylor, S. S. & Sowadski, J. M. (1991) *Science* **253,** 407–414.
5. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Xuong, N.-H., Taylor, S. S. & Sowadski, J. M. (1991) *Science* **253,** 414–420.
6. Zheng, J., Knighton, D. R., Ten Eyck, L. F., Karlsson, R., Xuong, N.-H., Taylor, S. S. & Sowadski, J. M. (1993) *Biochemistry* **32,** 2154–2161.
7. Zheng, J., Trafny, E. A., Knighton, D. R., Xuong, N.-H., Taylor, S. S., Ten Eyck, L. F. & Sowadski, J. M. (1993) *Acta Crystallogr. D* **49,** 362–365.
8. Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H. & Huber, R. (1993) *EMBO J.* **12,** 849–859.
9. Taylor, S. S. & Radzio-Andzelm, E. (1994) *Structure* **2,** 345–355.
10. Veron, M., Radzio-Andzelm, E., Tsigelny, I., Ten Eyck, L. F. & Taylor, S. S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10618–10622.
11. Herberg, F., Zimmermann, B., McGlone, M. & Taylor, S. S. (1997) *Protein Sci.* **6**, 569–579.
12. Alhanaty, E. & Shaltiel, S. (1979) *Biochem. Biophys. Res. Commun.* **39,** 323–332.
13. Alhanaty, E., Patinkin, J., Tauber-Finkelstein, M. & Shaltiel, S. (1981) *Proc. Natl. Acad. Sci. USA* **73,** 3492–3495.
14. Alhanaty, E., Tauber-Finkelstein, M., Schmeeda, H. & Shaltiel, S. (1985) *Curr. Top. Cell. Regul.* **27,** 267–278.
15. Jiménez, J. S., Kupfer, A., Gani, V. & Shaltiel, S. (1982) *Biochemistry* **21,** 1623–1630.
16. Nelson, N. & Taylor, S. S. (1983) *J. Biol. Chem.* **258,** 10981–10987.
17. Kupfer, A., Jiménez, J. S. & Shaltiel, S. (1980) *Biochem. Biophys. Res. Comm.* **96,** 77–80.
18. Buechler, J. A. & Taylor, S. S. (1990) *Biochemistry* **29,** 1937–1943.
19. Kupfer, A., Jiménez, J. S., Gottlieb, P. & Shaltiel, S. (1982) *Biochemistry* **21,** 1631–1637.
20. Karlsson, R., Zheng, J., Xuong, N.-H., Taylor, S. S. & Sowadski, J. M. (1993) *Acta Crystallogr. D* **49,** 381–388.
21. Zheng, J., Knighton, D. R., Xuong, N.-H., Taylor, S. S., Sowadski, J. M. & Ten Eyck, L. F. (1993) *Protein Sci.* **2,** 1559–1573.
22. Chestukhin, A., Litovchick, L., Muradov, K., Batkin, M. & Shaltiel, S. (1997) *J. Biol. Chem.* **272,** 3153–3160.
23. Olah, G. A., Mitchell, R. D., Sosnick, T. R., Walsh, D. A. & Trewhella, J. (1993) *Biochemistry* **32,** 3649–3657.
24. Reed, J. & Kinzel, V. (1984) *Biochemistry* **23,** 968–973.
25. Reed, J., Kinzel, V., Kemp, B. E., Cheng, H.-C. & Walsh, D. A. (1985) *Biochemistry* **24,** 2967–2973.
26. Narayana, N., Cox, S., Xuong, N.-H., TenEyck, L. F. & Taylor, S. S. (1997) *Structure* **5,** 921–935.
27. Chestukhin, A., Litovchick, L., Schourov, D., Cox, S., Taylor, S. S. & Shaltiel, S. (1996) *J. Biol. Chem.* **271,** 10175–10182.
28. Knighton, D. R., Bell, S. M., Zheng, J., Ten Eyck, L. F., Xuong, N.-H., Taylor, S. S. & Sowadski, J. M. (1993) *Acta Crystallogr. D* **49,** 357–361.
29. Narayana, N., Cox, S., Shaltiel, S., Taylor, S. S. & Xuong, N. (1997) *Biochemistry* **36,** 4438–4448.
30. Madhusudan, K. R., Trafny, E. A., Xuong, N.-H., Adams, J. A., Ten Eyck, L. F., Taylor, S. S. & Sowadski, J. M. (1994) *Protein Sci.* **3,** 176–187.
31. Hoppe, J., Freist, W., Marutzky, R. & Shaltiel, S. (1978) *Eur. J. Biochem.* **90,** 427–432.
32. Taylor, S. S. & Radzio-Andzelm, E. (1997) *Curr. Opin. Chem. Biol.* **1,** 219–226.
33. Edsall, J. T. & McKenzie, H. A. (1983) *Adv. Biophys.* **16,** 53–183.
34. Finney, J. L. (1977) *Philos. Trans. R. Soc. London* **278,** 3–32.
35. Meyer, E. (1992) *Protein Sci.* **1,** 1543–62.
36. Meyer, E., Cole, G., Radhakrishnan, R. & Epp, O. (1988) *Acta Crystallogr.* **44,** 26–38.
37. Sreenivasan, U. & Axelson, P. H. (1992) *Biochemistry* **31,** 12785–12791.
38. Malin, R., Zielenkiewicz, P. & Saenger, W. (1991) *J. Biol. Chem* **266,** 4848–4852.
39. Zegers, I., Maes, D., Dao-Thi, M. H., Poortmans, F., Palmer, R. & Wyns, L. (1994) *Protein Sci.* **3,** 2322–2339.
40. Komives, E. A., Lougheed, J. C., Liu, K., Sugio, S., Zhang, Z., Petsko, G. A. & Ringe, D. (1995) *Biochemistry* **34,** 13612–13621.
41. Grant, B. D. & Adams, J. A. (1996) *Biochemistry* **35,** 2022–2029.
42. Adams, J. A. & Taylor, S. S. (1992) *Biochemistry* **31,** 8516–8522.
43. Lew, J., Taylor, S. S. & Adams, J. A. (1997) *J. Biol. Chem.* **36,** 6717–6724.
44. Adams, J. A., McGlone, M. L., Gibson, R. & Taylor, S. S. (1995) *Biochemistry* **34,** 2447–2454.
45. Xu, W., Harrison, S. C. & Eck, M. J. (1997) *Nature (London)* **385,** 595–599.
46. Sicheri, F., Moarefi, I. & Kuriyan, J. (1997) *Nature (London)* **385,** 602–609.
47. deBondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O. & Kim, S.-H. (1993) *Nature (London)* **363,** 595–602.
48. Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. & Pavletich, N. P. (1995) *Nature (London)* **376,** 313–320.
49. Canagarajah, B. J., Khokhlatchev, A., Cobb, M. H. & Goldsmith, E. J. (1997) *Cell* **5,** 859–869.