# Regulatory Subunit of Protein Kinase A: Structure of Deletion Mutant with cAMP Binding Domains

Y. Su, W. R. G. Dostmann,* F. W. Herberg,† K. Durick, N-h. Xuong, L. Ten Eyck, S. S. Taylor,‡ K. I. Varughese

In the molecular scheme of living organisms, adenosine 3′,5′-monophosphate (cyclic AMP or cAMP) has been a universal second messenger. In eukaryotic cells, the primary receptors for cAMP are the regulatory subunits of cAMP-dependent protein kinase. The crystal structure of a 1–91 deletion mutant of the type Iα regulatory subunit was refined to 2.8 Å resolution. Each of the two tandem cAMP binding domains provides an extensive network of hydrogen bonds that buries the cyclic phosphate and the ribose between two β strands that are linked by a short α helix. Each adenine base stacks against an aromatic ring that lies outside the β barrel. This structure provides a molecular basis for understanding how cAMP binds cooperatively to its receptor protein, thus mediating activation of the kinase.

Protein phosphorylation and dephosphorylation is one of the principal mechanisms by which cellular functions are regulated in eukaryotic cells in response to external stimuli (*1*). The enzymes catalyzing these phosphorylations, the protein kinases, are tightly regulated and maintained in an inactive state in the absence of the specific activating signal. The mechanism for maintaining the inhibited state of any protein kinase is at least as critical for its function as is catalytic efficiency.

Although the family of protein kinases now includes several hundred members (*2*), cyclic AMP–dependent protein kinase (cAPK) was among the first to be characterized (*3*). The demonstration of the activation of cAPK by cAMP introduced the hormone second messenger concept, whereby a hormone binding to the extracellular surface led to the generation of a cytoplasmic second messenger (*4*). Of the protein kinases, cAPK is also one of the simplest and best understood biochemically (*5*, *6*), largely because the regulatory (R) and catalytic (C) components are coded for by separate genes, and the proteins can be readily separated upon activation. The inactive holoenzyme is an $R_2C_2$ tetramer. Cyclic AMP binding cooperatively causes the complex to dissociate, thereby releasing an $R_2$ $(cAMP)_4$ dimer and two free and active C subunits. Both cytoplasmic and nuclear proteins are substrates for cAPK, and when C is not anchored to R in the $R_2C_2$ complex, it can enter the nucleus (*7*).

There are two general classes of R subunits, types I and II, and within each class are at least two distinct gene products (*2*, *6*). All R subunits nevertheless retain a well-defined domain structure. At the amino terminus is a dimerization domain, followed by an autoinhibitor site that resembles either a substrate or an inhibitor. This autoinhibitor segment binds to the active site of the catalytic subunit. The R subunits are thus competitive inhibitors of substrate proteins. The carboxyl terminus is comprised of two tandem homologous cAMP binding domains, A and B. Site A is masked in the holoenzyme so that the cooperative activation is mediated by cAMP binding first to site B (*8*). This triggers a conformational change that makes site A more accessible. Cyclic AMP binding to site A then mediates dissociation of the complex. The two cAMP binding sites can be readily distinguished by several criteria. Site A has a faster off-rate and has a preference for N6-substituted analogs. Site B, with a slower off-rate, is preferred by C2- and C8-substituted analogs (*9*). The cAMP binding domains are also homologous to the cAMP binding domain of the catabolite gene activator protein (CAP) in *Escherichia coli* (*10*).

The domain structure of the R subunits, first characterized by limited proteolysis, was subsequently probed with deletion mutants (*5*). One of the most stable of these mutants has a deletion of 91 residues at the $NH_2$-terminus (*11*). Although monomeric, it retains the autoinhibitor site as well as the two cAMP binding sites. It forms a tight complex rapidly with C, and cAMP mediates activation in a manner similar to the tetrameric holoenzyme. Full understanding of the molecular basis for cAMP binding and for holoenzyme activation requires high resolution structures. Crystal structures of the C subunit have been determined (*11a*). We describe here the structure of the Δ1–91 deletion mutant of the recombinant bovine RIα subunit (rRIα). Ever though the dimerization domain is absent, this structure nevertheless reveals the detailed features of each cAMP binding site and provides a molecular basis for the cooperative binding of cAMP and activation of the holoenzyme.

**Structure solution.** The Δ1−91 rRIα subunit, (Δ1−91)rRIα, was expressed in *E. coli* as described (*11*). Unlike the full-length R subunit, this deletion mutant is resistant to proteolysis. A typical yield is 200 mg from 4 liters of culture. To generate a heavy atom substitution site, we replaced the Cys residue with Ser at position 145 by the Kunkel method (*12*). The expression, biochemical properties, and crystallization of (Δ1−91)-rRIα(S145C) are similar to those of (Δ1−91)rRIα.

Crystals of (Δ1−91)rRIα and (Δ1−91)-rR(S145C) were grown by the vapor diffusion method with the hanging-drop procedure in Linbro plastic tissue culture plates at 22°C (*13*). The drops contained 5 μl of protein stock (8 mg/ml) and 5 μl of reservoir solution. Hexagonal crystals were produced from a reservoir solution of 1.1 to 1.2 M $(NH_4)_2SO_4$ (grade III, Sigma), 10 to 12.5 percent glycerol (Sigma) and 10 mM dithiothreitol buffered with 80 to 100 mM sodium acetate, pH 5.5; they grew to their maximum size, 0.15 by 0.15 by 1.5 mm, over 3 weeks. They belonged to space group $P6_122$ ($P6_522$), and the unit cell dimensions were $a = b = 88.9$ Å, $c = 179.9$ Å. There was one molecule per asymmetric unit with a Matthews coefficient, $V_m$, of 2.9 Å$^3$ per dalton. The solvent content (*14*) was approximately 57 percent.

Diffraction data for the native and heavy atom derivative crystals were collected initially with the Xuong-Hamlin multiwire area detector system (*15*) at the NIH National Research Resource at UCSD. The native crystal diffracted to 2.9 Å resolution. Three heavy atom derivatives, namely, two mercury derivatives and a gold derivative, were prepared, and diffraction data to 3.5 Å were collected. The diffraction data for the native protein and the derivatives were again measured to higher resolutions with an MAR image plate scanner at the Stanford Synchrotron Radiation Laboratory (SSRL). The statistics of data collection and the multiple isomorphic replacement (MIR) phases computed with the program

Y. Su, W. R. G. Dostmann, F. W. Herberg, K. Durick, and S. S. Taylor are in the Department of Chemistry and Biochemistry and N-h. Xuong and K. I. Varughese are in the Department of Biology, University of California, San Diego, La Jolla, CA 92093–0654, USA. L. Ten Eyck is at the San Diego Supercomputer Center, La Jolla, CA 92186, USA.

*Present address: Institut für Pharmakologie und Toxicologie, Technische Universität München, Biedersteiner Straße 29, 80802 München, Germany.

†Present address: Ruhr-Universität Bochum, Medizinische Fakiltät, Institut für Physiologische Chemie I, 44780 Bochum, Germany.

‡To whom correspondence should be addressed.

package PHASES (*16*) are listed in Table 1. Solvent-flattening calculations (*17*) with 2.8 Å MIR phases gave an interpretable map. The space group was assigned as $P6_522$ from the right-handedness of the α helices.

After the map was computed with the solvent-flattened phases, model building was done with the graphics programs TOM (*18*) and O (*19*). The program O was used to construct the backbone and side chains based on the Cα trace that was manually generated with the use of TOM. The map showed well-defined electron density for cAMP bound to both domains A and B (Fig. 1). The initial model included two molecules of cAMP and three segments of peptide chains consisting of 190 residues. The coordinates of this model were improved by X-PLOR (*20*) refinement, which gave an *R* factor of 34.1 percent for data between 10.0 to 3.0 Å. At this stage, using the program SIGMAA (*21*), we combined phases from the refined partial model with the MIR phases, and the resultant map revealed better electron density for the unbuilt regions. Seventy-four more residues were incorporated into the model after several rounds of model building and phase combinations. The refinement was done iteratively with X-PLOR and omit maps. At present, model consists of 2020 nonhydrogen atoms. The $NH_2$-terminal residues 92 to 112 and the COOH-terminal residues 377 to 379 could not be traced. Five other regions of the chain consisting of residues 113 to 118, 275 to 279, 285 to 288, 302 to 311, and 375 to 376 showed up in the map with densities less well defined compared to the rest of the structure; however, these residues could be traced and were included in the refinements. These regions were all solvent-exposed. Residues 302 to 311, located in the surface loop connecting β4 and β5 of domain B, have high *B* factors.

Side chains were assigned as follows. $Cys^{145}$ and $Cys^{360}$ were identified from the positions of the mercury atoms in the derivatives. In each domain, conserved residues interacting with cAMP, $Glu^{200}$, and $Arg^{209}$ in domain A and $Glu^{324}$ and $Arg^{333}$

**Table 1.** Diffraction data and structure solution statistics. *N*, number of observations, $R_{deriv} = \Sigma|F_p - F_{ph}|/\Sigma F_p$. $R_{sym} = \Sigma_h\Sigma^N{}_{i=1}|I(h) - I(h)_i|/\Sigma_h\Sigma^N{}_{i=1}I(h)_i$, where $I(h)_i$ is the *i*th measurement of reflection h and I(h) is the mean value of the *N* equivalent reflections. The *R* derivative is calculated with respect to a merged data set obtained by merging the two data sets. $F_H$, average root-mean-square (rms) heavy atom structure factor amplitude. $F_{anom}$, average anomalous dispersion structure factor amplitude of heavy atoms. *E*, rms closure error. Centric $R$, $|F_{ph}(obs) - F_{ph}(calc)| \times 100/\Sigma|F_{ph}(obs) - F_p|$; $E_{anom}$, rms anomalous closure error. The Au derivative was formed by soaking the crystal in solutions of 2 mM $KAuCl_4$. The Hg derivative was obtained by cocrystallizing with $C_2H_5HgCl$. The protein has two cysteines, one at residue 345 and the other at residue 360; however, only $Cys^{360}$ reacted with Hg. For obtaining another Hg derivative, a new mutant was created by replacing $Ser^{145}$ with Cys. Soaking of crystals of this mutant with 6 mM *p*-chloromercuribenzene sulfonate (PCMBS) yielded a derivative with two Hg sites. Diffraction data from two native crystals were measured at UCSD (λ=1.518 Å). These two data sets were merged with a native data set measured at SSRL (λ = 1.08 Å) on a MAR image plate; the program MOSFLM (*45*), was used to reduce the data from SSRL. The Hg atom position in the $C_2H_5HgCl$ derivative was located from the isomorphous difference Patterson map. The Au position and the positions of the two Hg atoms in the S145C mutant derivative were located with cross-difference electron density maps. The solvent-flattening calculations were done according to Wang *et al.* (*17*), on the basis of an assumed solvent content of 50 percent. After convergence, the mean figure of merit was 0.84 with the map inversion *R* factor of 29 percent. An analysis of the diffraction pattern revealed that the overall temperature factor of the crystal is anisotropic. Hence an anisotropic correction was applied. After the correction, the refinement provided an *R* factor of 22.1 percent from the previous *R* of 25 percent. Further refinement with the program TNT (*46*) with all the data yielded the final *R* of 22.9 percent. As the high resolution data were affected by decay, we confirmed our present refinement to 2.8 Å data.

Data sets used in the structure determination

| Data set | Reflections (*N*) | Observations (*N*) | $d_{min}$ (Å) | $R_{sym}$ (%) | Overall complete (%) | $R_{deriv}$ (%) | Device |
|---|---|---|---|---|---|---|---|
| Native 1 | 8904 | 61857 | 3.0 | 4.8 | 99.3 | | Multiwire |
| Native 2 | 15809 | 37561 | 2.4 | 5.5 | 88.5 | | MAR(SSRL) |
| $C_2H_5HgCl$ | 14656 | 53743 | 2.5 | 7.5 | 95.0 | 7.8 | MAR(SSRL) |
| $KAuCl_4$ | 9494 | 19257 | 2.7 | 6.9 | 83.8 | 11.8 | MAR(SSRL) |
| Hg derivative of S145C mutant | 9801 | 52070 | 2.8 | 8.9 | 88.7 | 10.7 | MAR(SSRL) |

MIR statistics of the heavy atom derivatives

| Derivative | Measurement | Average resolution of the shell (Å) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Overall | 9.16 | 5.57 | 4.63 | 4.12 | 3.77 | 3.51 | 3.31 | 3.15 | 3.01 | 2.87 |
| $C_2H_5HgCl$ | $F_H/E$ | 1.96 | 1.81 | 2.29 | 1.90 | 1.72 | 1.92 | 1.96 | 2.07 | 1.97 | 2.08 | 2.11 |
| | $F_{anom}/E_{anom}$ | 2.02 | 1.89 | 2.15 | 1.76 | 1.83 | 2.05 | 2.03 | 2.24 | 2.21 | 2.40 | 1.95 |
| | Centric *R* | 0.68 | | | | | | | | | | |
| $KAuCl_4$ | $F_H/E$ | 1.40 | 1.58 | 1.84 | 1.54 | 1.36 | 1.24 | 1.26 | 1.44 | 1.27 | 1.21 | 1.29 |
| | Centric *R* | 0.58 | | | | | | | | | | |
| Hg derivative of S145C mutant | $F_H/E$ | 2.08 | 2.32 | 2.36 | 1.99 | 1.88 | 2.01 | 2.07 | 2.15 | 1.96 | 2.02 | 1.98 |
| | Centric *R* | 0.64 | | | | | | | | | | |
| Figure of merit | | 0.66 | 0.737 | 0.714 | 0.684 | 0.686 | 0.674 | 0.648 | 0.656 | 0.622 | 0.576 | 0.563 |

Refinement

| Model | Final *R* factor | *B* factor | Data selection | Programs |
|---|---|---|---|---|
| 190 residues + 2 cAMP | 34.1 | Overall | 10.0–3.0 Å $I/\sigma>2$ | X-PLOR |
| 247 residues + 2 cAMP | 25.0 | Individual | 10.0–2.8 Å $I/\sigma>2$ | X-PLOR |
| 264 residues + 2 cAMP | 22.1 | Individual | 8.0–2.8 Å $I/\sigma>2$ | X-PLOR |
| 264 residues + 2 cAMP | 22.9 | Fixed | 15.0–2.8 Å all *F* | TNT |

in domain B, were identified. $Trp^{260}$ was identified by its proximity to the adenine ring in domain A. With these residues as markers, side chains were eventually assigned for all the remaining residues. At present, the *R* factor for the model is 21.7 percent for all the data between 8 to 2.8 Å with $I/\sigma \geq 2.0$ (Table 1). The root-mean-square (rms) deviations from ideal bond length and angles are 0.019 Å and 2.8°, respectively. Ramachandran plots of $\phi,\Psi$ angles (*22*) showed that most of nonglycine and nonproline residues are within the allowed regions with 82.3 percent of the 232 residues in the energetically most favored areas. Only eight nonglycine residues fall into the generously allowed regions. Five of these residues, $Pro^{117}$, $Lys^{118}$, $Arg^{304}$, $Glu^{306}$, and $Glu^{308}$, are located in regions where electron density is less ordered, as indicated above. The other three residues, $Glu^{187}$, $Ser^{319}$, and $Asn^{330}$, are in surface loops.

**Molecular architecture.** The molecule is a monomer with an α/β structure consisting of two cAMP binding domains, each having a similar folding pattern to the cAMP binding domain of CAP (Fig. 2). To simplify comparisons, we have adopted a nomenclature based on CAP, and we aligned the sequences of the cAMP binding domains in CAP, RIα, and RIIα (Fig. 3). The first 21 residues of this deletion mutant are not seen, presumably because this region is disordered. This segment includes the autoinhibitor site that mimics the substrate recognition site and binds to the active site of C. The structure begins with $Arg^{113}$ followed by a helix (α-X:N) (see legend to Fig. 2 for nomenclature). An extended chain then turns into cAMP binding domain A. The COOH-terminus of domain A goes immediately into domain B. The last three residues are disordered and not seen. The molecule is elongated with overall dimensions of 65 Å by 45 Å by 34 Å consistent with earlier estimates of dimensional asymmetry based on the Stokes radius (27.5 Å) (*11*). The distances between the two cAMP molecules is approximately 26 Å.

The specific topography of each cAMP binding domain (Fig. 4) consists of three major α helices and eight β strands. The eight β strands form a flattened β barrel, consisting of two antiparallel β sheets, each with four strands, connected in a jelly-roll topology. The V-shaped pocket in this jelly-roll β barrel forms a major part of the cAMP binding site. The three helices are connected to the ends of the β barrel. Helix A is at the $NH_2$-terminus and helix B, followed immediately by helix C, is at the COOH-terminus. Helices A and B are antiparallel. Between strands 6 and 7 is a short helix, α-B′, sitting on the top edge of the jelly roll. The phosphate group of the cAMP is located near the $NH_2$-terminus of this helix and, as discussed later, the capping of this helix is integrally dependent on cAMP.

**Cyclic AMP binding sites.** In both domains cAMP is bound in a *syn* conformation. The phosphate and the ribose ring interact with the protein through several hydrogen bonds and electrostatic contacts. These interactions mostly occur among highly conserved residues in the segment linking β strands 6 and 7. In contrast, the adenine ring interacts primarily through hydrophobic and stacking interactions with residues in and near the C helix. The nucleotide is thus sandwiched between the β barrel and the C helix. The conserved residues, $Glu^{200}$ and $Arg^{209}$ in domain A and $Glu^{324}$ and $Arg^{333}$ in domain B, participate directly in cAMP binding. The Glu residues interact with the 2′-OH of the ribose ring while the Arg residues interact with the phosphate. In both sites, an aromatic side chain, $Trp^{260}$ in domain A and $Tyr^{371}$ in domain B, stacks with the adenine ring.

While Fig. 4, A and B, shows the general features of each cAMP binding domain, the multiple contacts between cAMP and the protein are summarized in detail in Fig. 4, C and D. The region extending from $Gly^{199}$ to $Ala^{210}$ in site A and from $Gly^{323}$ to $Ala^{334}$ in site B are linked by an extensive network of contacts that are synergistically dependent on the presence of cAMP. For example, $Glu^{200}$ makes multiple contacts. In addition to its interaction with the ribose 2′-OH, it also hydrogen bonds to the Nε1 of $Trp^{260}$ and is stabilized by electrostatic contacts with $Arg^{241}$. $Glu^{200}$ also begins the one-and-a-half-turn helix. Most of the unpaired amides at the beginning of this helix are capped by interactions with protein side

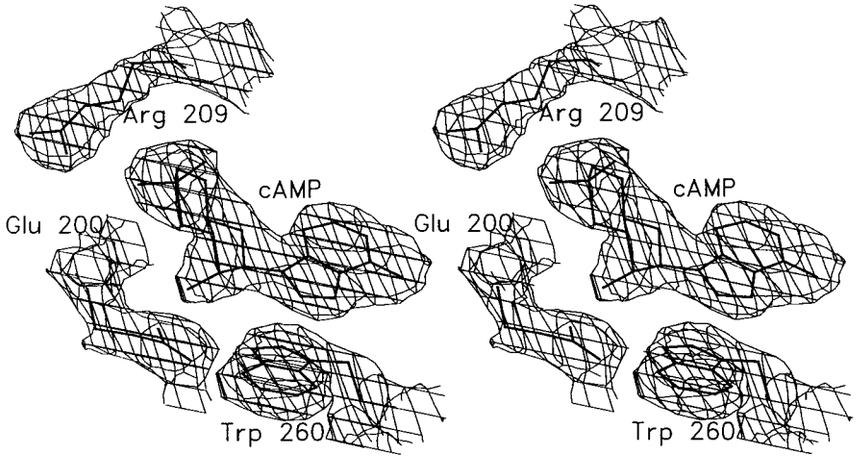


**Fig. 1.** Stereo view of the $(2F_o - F_c)$ electron density of the cAMP binding regions of domain A at 2.8 Å resolution. The map is contoured at 1.5 sigma; cAMP and three residues, $Glu^{200}$, $Arg^{209}$, and $Trp^{260}$, with prominent interactions are shown.

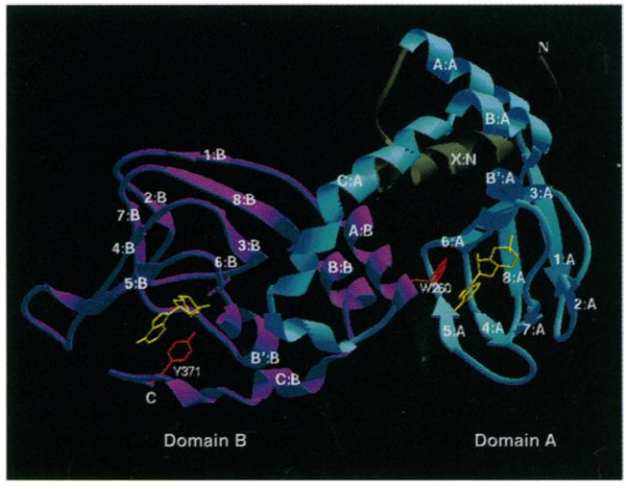


**Fig. 2.** General architecture of (Δ1–91)rRIα. The three domains, shown as a ribbon representation of the backbone, are colored with the $NH_2$-terminal region in gray, domain A in cyan, domain B in magenta, and cAMP in yellow. $W^{260}$ and $Y^{371}$, affinity labeled by 8-N3–cAMP, are shown in red. For nomenclature we have adopted the numbering used for the cAMP binding domain of CAP. The three major helices in domain A are named αA:A, αB:A, and αC:A, and the corresponding helices in domain B are named αA:B, αB:B, αC:B, (Fig. 3). Each β strand is likewise designated β1:A, β2:A. . .β8:A in domain A and β1:B, β2:B, and so forth in domain B. In that the $NH_2$-terminus is missing in this deletion mutant, the first helix is designated at present as α-X:N.

chains or by cAMP. For example, the αNH of $Glu^{199}$, conserved so far in all of these cAMP binding sites, hydrogen bonds to the 2′-OH of cAMP. The αNH of $Glu^{200}$ hydrogen bonds to the Oϵ2 of its own side chain. The αNH of $Ala^{202}$ hydrogen bonds to the equatorial oxygen of the phosphate in cAMP. The αNH of $Leu^{203}$ hydrogen bonds to the α carbonyl of $Leu^{201}$. Thus, the architecture of this site, and specifically the secondary structure that includes this short helix, is integrally dependent on the presence of cAMP itself. The stacking of the adenine ring with $Trp^{260}$ also depends on this network of interactions involving $Glu^{200}$. The $Arg^{209}$ plays a major structural role in addition to binding cAMP. It contributes to cAMP binding by interacting with the equatorial exocyclic oxygen of the cAMP phosphate. This same Nη nitrogen also is only 3.5 Å from the backbone carbonyl of $Gly^{199}$. Thus $Arg^{209}$ bridges the conserved segment that links β strands 6 and 7. However, by contacting the backbone carbonyl of $Asn^{171}$ in β strand 3 and the side chain carboxylate of $Asp^{170}$, it bridges the cAMP binding domain and transmits a signal that extends beyond the immediate cAMP binding site. The interaction with $Asp^{170}$ also contributes to the neutralization of the charge on $Arg^{209}$.

A similar network of contacts is found in site B (Fig. 4D). $Glu^{324}$ hydrogen bonds with the 2′-OH of the ribose ring, to the phenolic OH of $Tyr^{371}$, and interacts with its own backbone amide. The aromatic side chain at the end of the C helix is thus positioned both by its stacking with the adenine ring of cAMP and by hydrogen bonding to a conserved residue in the β barrel. Likewise, the backbone amide of $Gly^{323}$ also interacts with the 2′-OH of the ribose, and $Glu^{324}$ is followed by a short helix. Most of the unpaired αNH groups at the beginning of this helix also hydrogen bond in a manner similar to that in domain A. Unlike $Glu^{200}$, $Glu^{324}$ does not interact with an Arg comparable to $Arg^{241}$. As in domain A, conserved $Arg^{333}$ forms a single ion pair with the equatorial phosphate oxygen of cAMP but also plays a structural role by interacting with the backbone carbonyls of $Gly^{323}$ and $Glu^{289}$, which is in β strand 3, similar to domain A. In domain B there is no carboxylate near $Arg^{333}$ to help neutralize its charge; but there are still potential interactions with residues that extend from $Gly^{287}$ to $Glu^{289}$.

In addition to the electrostatic interactions, the tight binding of cAMP with $K_d$'s of 10 to 100 nM involves a number of strong hydrophobic interactions. In site A, one side of the adenine base faces a hydrophobic pocket formed by the side chains of $Ala^{189}$, $Val^{184}$, $Val^{182}$, $Ala^{210}$, and $Ala^{211}$. On the other side, the five-member ring of the adenine base stacks with the indole ring of $Trp^{260}$ while the six-member ring has van der Waals contacts with $Leu^{201}$. In cAMP binding site B, the adenine ring is sandwiched by hydrophobic interactions between the side chains of $Leu^{316}$, $Tyr^{371}$, $Ser^{373}$, and $Ile^{325}$ on one side and $Val^{300}$, $Val^{313}$, and $Ala^{335}$ on the other. The phenol ring of $Tyr^{371}$ also stacks with the adenine ring. The stacking of the aromatic rings of $Trp^{260}$ and $Tyr^{371}$ with the adenine ring of cAMP also orients the dipole moments of the adenine, Tyr, and Trp rings in an antiparallel alignment as predicted (*23*). Whereas the dipole moments for the adenine and Trp rings are fixed, the dipole moment of Tyr depends on the rotation of the phenolic hydroxyl group. The hydroxyl group of $Tyr^{371}$ is fixed by a hydrogen bond to $Glu^{324}$ so that the dipole of $Tyr^{371}$ cannot rotate.

In prokaryotes the primary receptor for cAMP is CAP where cAMP binding directly mediates gene expression (*24*). As predicted by sequence alignments (Fig. 3) (*10*), the general features of the cAMP domains in CAP and R are conserved although we were unable to solve the structure by molecular replacement with the coordinates of CAP. Each cAMP binding domain has three main helices and an eight-stranded β barrel. There are three invariant Gly residues in each domain. One of them ($Gly^{166}$ in domain A and $Gly^{285}$ in domain B) lies between β2 and β3 and is situated at the third corner of a type II β turn. Type II β turns prefer a Gly residue at this position. The second conserved Gly is located in a loop connecting β3 and β4. The third invariant Gly begins the active site pocket, $Gly^{199}$ in A and $Gly^{323}$ in B. A larger side chain here could collide with the adenine ring. Although the fold is conserved, the length of the C helix and its position relative to the β barrel differ in these three structures (Fig. 5). Most significant is the displacement of the C helix in domain A away from the β barrel. As discussed below, two residues, $Trp^{260}$ and $Arg^{241}$, specifically keep this helix extended away from the barrel.

The different conformation of cAMP, *syn* in R and *anti* in CAP, is due most likely to differences in the environment surrounding the adenine rings. The stacking of the adenine ring with aromatic side chains comparable to $Trp^{260}$ and $Tyr^{371}$, for example, is missing in CAP. Another difference is that the short helix found between β strands 6 and 7 in the R subunits is missing in CAP. CAP has one extra residue in the region that links β6 and β7, and this may prevent the helix from forming. Cyclic AMP binding is three orders of magnitude tighter in R than CAP. The presence of the phosphate:helix interactions and the stacking interactions between the adenine rings and the aromatic side chains probably account for this difference, at least in part.

In eukaryotes, homologous cAMP binding motifs are conserved in the R subunits of cAPK, in the guanosine 3′,5′-monophosphate (cGMP)–dependent protein kinase (cGPK), and in the cyclic nucleotide-gated channels (*25*). The difference in relative specificity for cAMP compared to cGMP is approximately a factor of 100 for cAPK and cGPK (*26*). One residue that partially accounts for this specificity is the equivalent

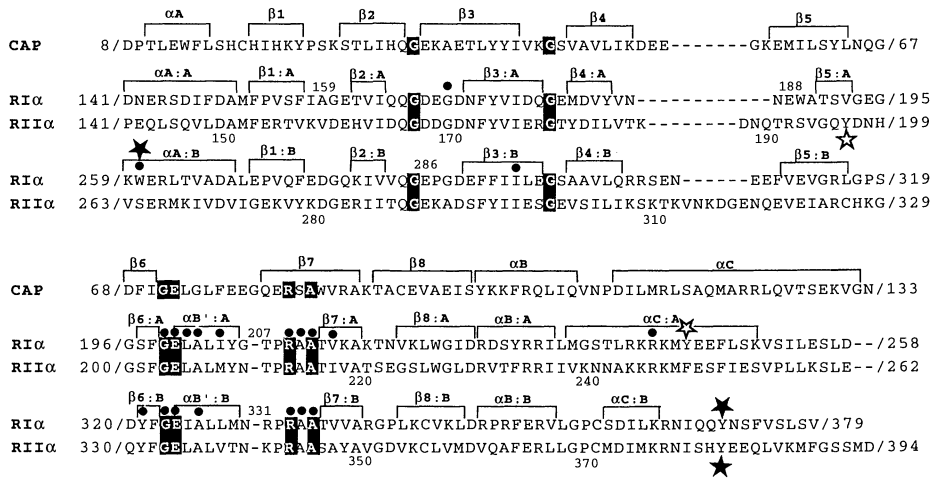


**Fig. 3.** Sequence alignment of the cAMP binding domains of RI, RII, and CAP. RI corresponds to bovine RIα and RII corresponds to bovine RIIα. Residues that are conserved in all three proteins are indicated by black boxes. Sites of affinity labeling with 8-N3-cAMP in the native protein are indicated by filled stars. Sites affinity labeled in the proteolyzed protein are indicated by open stars. Residues identified as important for cAMP-mediated activation by genetic screening are indicated by black dots. Single-letter abbreviations for the amino acid residues are A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

of $Ala^{210}$ (domain A) and $Ala^{334}$ (domain B) in RIα. This is always a Thr in cGPK and an Ala in R. Replacement of these Ala residues with Thr in RIα did improve specificity for cGMP but did not weaken the affinity for cAMP (*27*). On the basis of the current structure, replacing $Ala^{210}$ with Thr and adding an $NH_2$ at the position 6 of cAMP gives a distance of 2.4 Å between the OG1 of Thr and the N2 of cGMP.

**Correlation of the structure with chemical data.** Affinity labeling with 8-$N_3$-cAMP was used to identify residues near the cAMP binding sites. Two sites were labeled in RIα: $Trp^{260}$ in Site A and $Tyr^{371}$ in Site B (*28*). In the RII subunit, $Trp^{260}$ is replaced with a Ser and only one site was labeled, $Tyr^{381}$, the equivalent of $Tyr^{371}$ in RI (*29*).

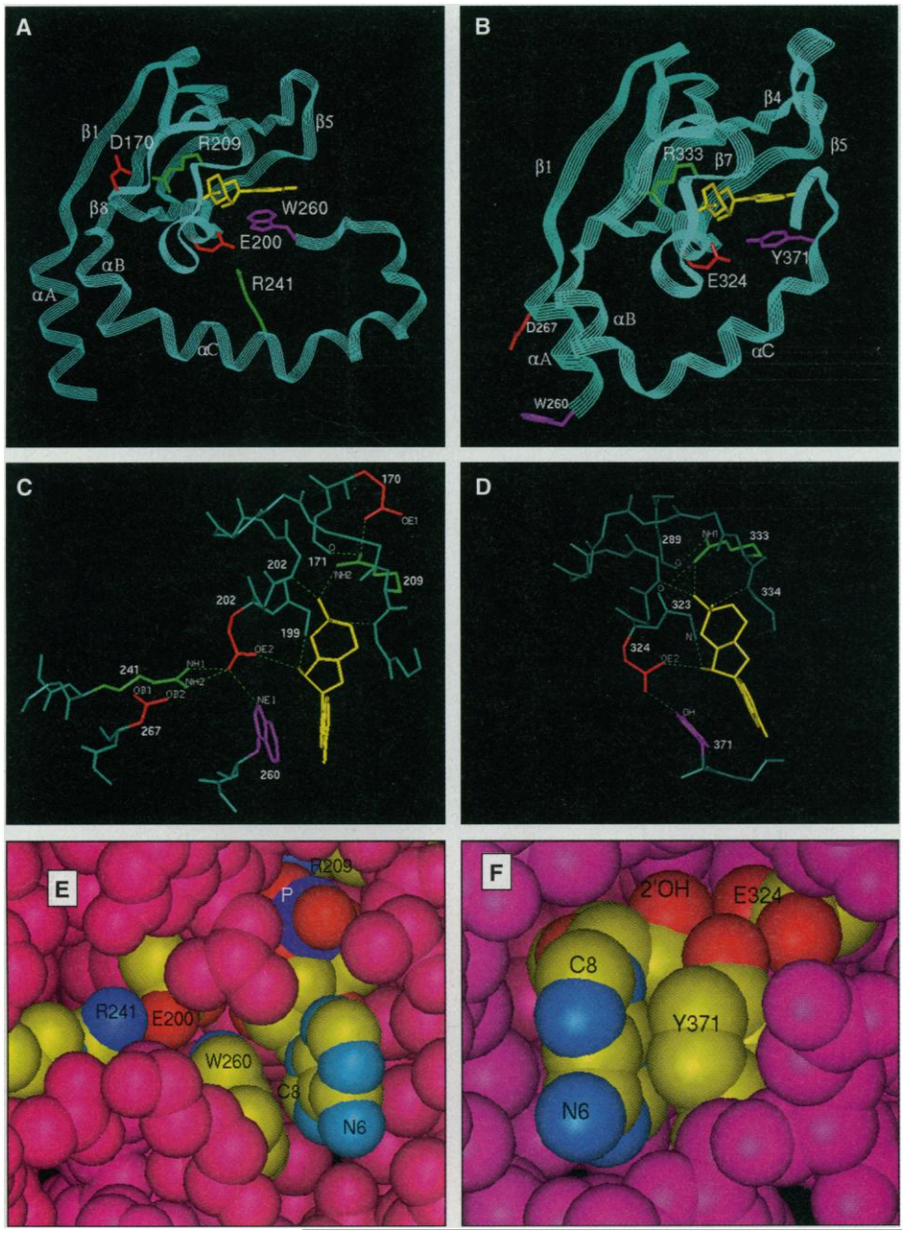


**Fig. 4.** (**A** and **B**): Overall features of the cAMP binding domains. Ribbon diagrams of domain of cAMP binding domains A and B are shown in (A) and (B), respectively. The general orientation of cAMP (yellow), relative to the β barrel and the C helix is shown. Conserved residues ($R^{209}$ and $E^{200}$ in domain A and $R^{333}$ and $E^{324}$ in domain B) are indicated. Stereo views of the hydrogen bonding interactions between cAMP and the protein are shown in (**C**) (site A) and (**D**) (site b). Additional residues, $R^{241}$, $E^{267}$, and $D^{170}$, that interact directly with cAMP domain A are also shown. Possible H bonds are indicated by dashed lines (distances < 3.3 Å). Space-filling models of each cAMP binding site are shown in (**E**) (site A) and (**F**) (site B). Atoms of cAMP and residues involved in cAMP binding are colored differently: carbon in yellow, nitrogen in blue, phosphorus in dark blue, oxygen in red, the others in pink. C8 is the site of attachment of the photoreactive azido moiety.

These two aromatic side chains in RIα are optimally aligned, which presumably accounts for the exceptionally high efficiency in labeling (Fig. 4, E and F).

Affinity labeling of proteolytic fragments and deletion mutants provide an indication of conformational flexibility. For example, when domain B (residues 260 to 379) in RIα is deleted, all the residues normally labeled are absent. In this protein, $Tyr^{244}$ was labeled by 8-$N_3$-cAMP (*30*), indicating that, in the absence of domain B, the C helix in domain A moved closer to the β barrel, thus resembling more closely the orientation in both domain B and CAP (Fig. 5). When the $NH_2$-terminus of the RII subunit was removed by proteolysis (residues 1 to 94), the pattern of photoaffinity labeling also was altered. In this case, $Tyr^{196}$ in domain A was labeled, in addition to $Tyr^{381}$, indicating that the $NH_2$-terminus also imposed some structural constraints on the COOH-terminal part of the molecule (*31*).

The chemical features of the two cAMP binding sites were also mapped with cAMP analogs (*32*). As predicted, no strong hydrogen bonding interactions exist between the adenine ring and the protein; most of the interactions are hydrophobic and stacking. The closest potential hydrogen bond is between the N6 in cAMP bound to domain B and the backbone carbonyl of $Asn^{372}$. Whereas three H bonds between the 2′-, 3′-, and 5′-ribose oxygens were predicted, only the 2′-OH hydrogen bond was observed in the structure. The importance of the exocyclic phosphate oxygens was also correctly predicted. Although the axial oxygen interacts with the amide of $Ala^{210}$ and $Ala^{334}$, respectively, in site A and B, the equatorial oxygen binds to the side chain of $Arg^{209}$ and to the amide of $Ala^{202}$ in site A and to $Arg^{333}$ and $Ala^{326}$ in site B. There is no evidence, however, for a pentacovalent intermediate as was suggested. Analogs also predicted correctly that cAMP binds to R in a *syn* conformation, in contrast to the *anti* conformation seen in CAP.

Analogs also can discriminate between sites A and B (*33*). Site A can accept analogs having substituents at the C6 position whereas substituents at the C8 position are not well tolerated. The N6 of cAMP bound to site A is exposed, but the accessibility of the C8 position is blocked by the six-member ring of $Trp^{260}$, as well as by $Val^{182}$ and $Val^{192}$. In contrast, site B prefers analogs substituted at the C8 position, and in this site, the accessibility of the N6 position is blocked by van der Waals contacts with four residues, $Tyr^{371}$, $Asn^{372}$, $Ser^{373}$, and $Val^{313}$, the N7-C8 edge of cAMP is accessible (Fig. 4F). Another feature that distinguishes the two sites is the relative off-rates for cAMP. In site B with its slower

off-rate, the cyclic nucleotide is buried more deeply and is packed tightly against $Tyr^{371}$. At the base of the cAMP binding pocket, for example, $Arg^{333}$ is completely buried by the nucleotide (Fig. 4F). In contrast, cAMP binds to domain A with a relatively fast off rate. In the crystal structure, this cAMP binding site is more open.

Of the hundreds of cAMP analogs tested, only one group served as antagonists, the chiral phosphorothioate analog (Rp)cAMPS, and its analogs (*34*). When the sulfur is in the equatorial position, as in (Rp)cAMPS, dissociation of the holoenzyme was blocked, whereas (Sp)cAMPS was an agonist. Sulfur has a larger van der Waals radius (1.70 Å) than oxygen (1.35 Å) (*35*), and the P-S bond length of the phosphorothioate group is accordingly longer (1.95 Å compared to 1.5 Å). Because most of the charge resides on the sulfur, the resonance of the electrons is also reduced in these chiral analogs (*36*). Sulfur also does not form hydrogen bonds as well as oxygen. In (Rp)cAMPS, the sulfur replaces the equatorial oxygen that interacts with $Arg^{209}$ and the amide backbone of $Ala^{202}$ in Site A. In the crystal structure the charge of $Arg^{209}$ is neutralized in part by $Asp^{170}$. With (Rp)-cAMPS bound, the charge of R209 would be neutralized predominantly by the negative charge on the sulfur leaving no anchor for $Asp^{170}$. The hydrogen bonding to $Ala^{202}$ would also be disturbed. The larger size of the sulfur in both of the chiral analogs probably causes steric hindrance and accounts for their low affinity. At this point, it is not clear why (Sp)cAMPS, and not (Rp)cAMPS, is an antagonist for CAP.

Genetic approaches also identified functionally important residues. The most extensive genetic mapping has been done in RIα of S49 mouse lymphoma cells (*37*). For these cells cAMP is toxic. By isolating mutants that were resistant to cAMP, a family of RIα mutants defective in cAMP binding were identified. Most of these dominant negative mutations (Fig. 3) are located in the highly conserved β6-β7 loops. Additional residues that lie outside this loop are $Gly^{169}$ and $Arg^{241}$ in domain A and $Trp^{260}$ and $Ile^{293}$ in domain B. $Gly^{169}$ precedes $Asp^{170}$, which ion pairs with $Arg^{209}$, and $Arg^{241}$, which is critical for cooperativity (*38*), pairs with $Glu^{200}$, as predicted (*39*). In domain B, $Trp^{260}$ stacks with cAMP, and $Ile^{293}$ is at the hydrophobic interface between the two domains. Other mutations, genetically engineered into the RIα subunit, have provided insights into the chemical structure of the cAMP binding sites and the cooperativity between the two sites. Extensive in vitro analysis of mutants such as Arg209Lys, Gly199Glu, and Ala335Asp, has confirmed the importance of these residues for cAMP binding and signaling (*40*).

**Molecular basis for cooperativity.** A primary region of contact between domains A and B is an extended hydrophobic surface. Specifically, the hydrophobic surface in domain B formed by $Val^{265}$, $Leu^{269}$, $Val^{346}$, $Leu^{294}$, $Ile^{292}$, $Tyr^{321}$, $Leu^{364}$, $Leu^{357}$, $Val^{356}$, $Cys^{360}$, and $Ile^{363}$ is covered by $Tyr^{244}$, $Phe^{247}$, $Leu^{248}$, $Val^{251}$, $Ile^{253}$, $Leu^{254}$, and $Leu^{257}$ from domain A (Fig. 6). In addition, the ends of αA:B are anchored by further electrostatic and hydrophobic interactions with domain A. $Trp^{260}$, which is a critical residue that links the two domains, was first indicated to be important by affinity labeling (*28*). It lies at the beginning of αA:B and is thus part of the secondary structure of domain B, yet its side chain interacts directly with cAMP binding site A by stacking with the adenine ring and by hydrogen bonding to the conserved $Glu^{200}$ (Fig. 6). Thus it is an integral part of the network of contacts that define cAMP

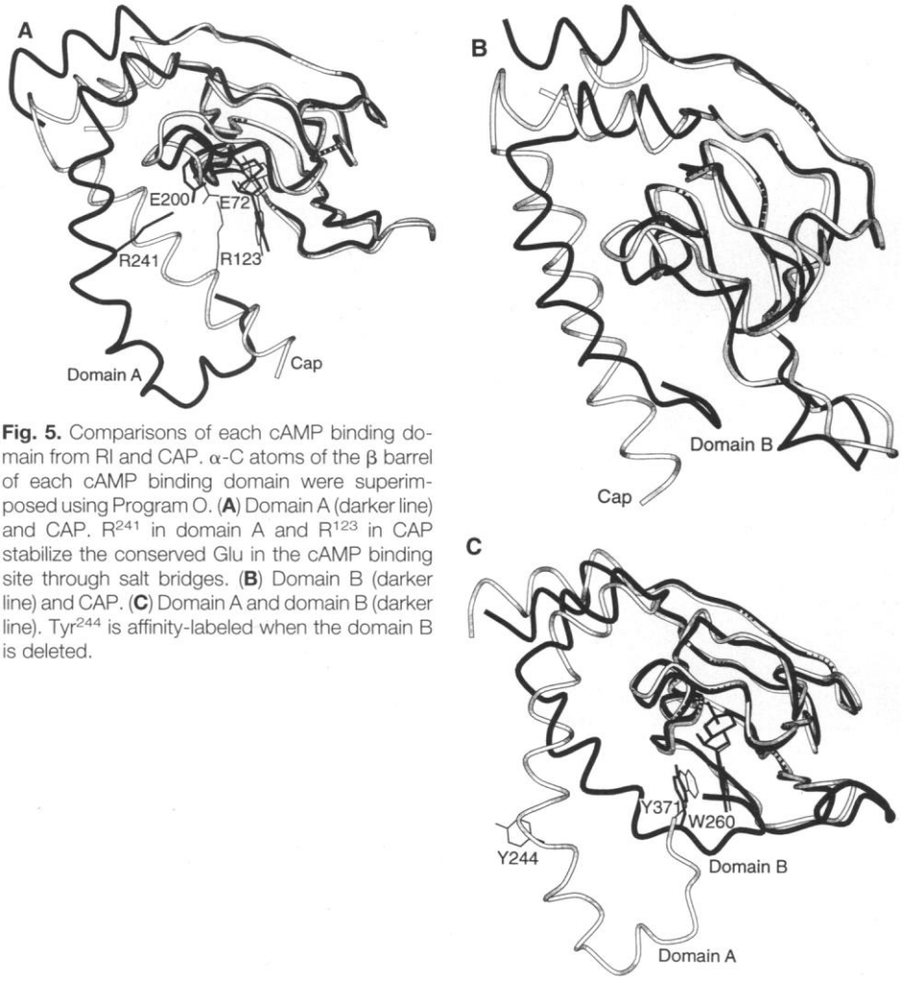


**Fig. 5.** Comparisons of each cAMP binding domain from RI and CAP. α-C atoms of the β barrel of each cAMP binding domain were superimposed using Program O. (**A**) Domain A (darker line) and CAP. $R^{241}$ in domain A and $R^{123}$ in CAP stabilize the conserved Glu in the cAMP binding site through salt bridges. (**B**) Domain B (darker line) and CAP. (**C**) Domain A and domain B (darker line). $Tyr^{244}$ is affinity-labeled when the domain B is deleted.

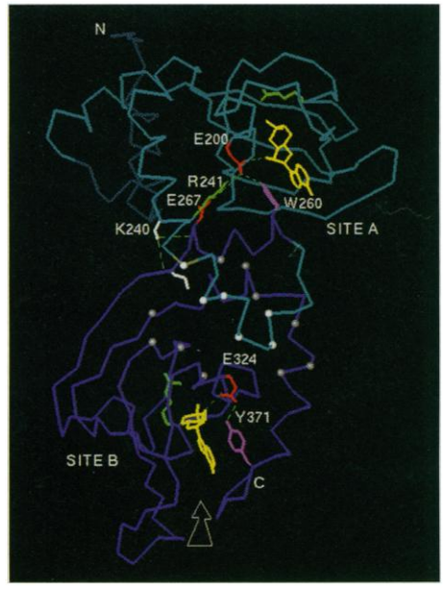


**Fig. 6.** Interactions between domains A and B. The α carbons of residues that participate in hydrophobic interactions between the two domains are indicated as balls. Additional residues that could play a role in mediating allosteric interactions are indicated ($E^{200}$, $R^{209}$, $K^{240}$, $R^{241}$, $W^{260}$, $E^{267}$, $E^{324}$, $R^{333}$, $Y^{371}$), as are the bound cAMP molecules. The arrow indicates the site where cAMP binds first in the holoenzyme.

binding site A. Another interdomain hydrogen bond interaction involves the side chain of $Lys^{240}$ in αC:A, which hydrogen bonds to the backbone carbonyls of $Asp^{267}$ and $Leu^{269}$ at the COOH-terminus of αA:B. Finally, there is an electrostatic interaction between $Arg^{241}$ and $Asp^{267}$. Thus the A helix of domain B is anchored firmly to key regions of site A by hydrophobic, electrostatic, and hydrogen bonding interactions.

One question relating to the R subunits of cAPK is how the cooperative binding of cAMP leads to the dissociation, and thus activation, of the holoenzyme. Although we cannot fully understand this process until a structure of the holoenzyme is solved, we can begin to understand how communication between sites A and B and the C subunit might be mediated from the structure of the R subunit. A deletion mutant lacking domain B (Δ260–379) and even a double mutant, where both the $NH_2$-terminus (Δ1–91) and domain B are deleted, still bind C tightly (*11*, *41*), demonstrating that domain B is not required for high affinity binding to C. On the basis of kinetic arguments, site A is masked in the holoenzyme (*8*). Thus, in the sequential cooperative pathway for activation of cAPK, cAMP binds first to site B, which "opens up" site A allowing cAMP to bind and C to be released.

There are only two stable conformations of the R subunit; the cAMP-saturated, dissociated R and the holoenzyme. The interactions between domains A and B as well as the immediate environment of each cAMP binding site must be different in these two structures. Since no structure of any cAMP-free domain is yet available, we can only speculate on the conformational changes that take place. The C helices, however, probably play an important role in both domains. The aromatic side chains of $Trp^{260}$ and $Tyr^{371}$, for example, are packed tightly against the cAMP ligands and have no contacts on the sides away from cAMP. They must therefore either collapse into the cAMP binding pocket in the absence of cAMP or otherwise rearrange. The C helix in domain A also is slightly bent suggesting a strained conformation. When cAMP binds to site B in the holoenzyme and stacks with $Tyr^{371}$, the orientation of the C helix relative to the β barrel has to change. The fact that cooperativity is lost when $Tyr^{371}$ is replaced with Phe (*42*) confirms the importance of this initial binding of cAMP to domain B and suggests that the strong dipole-dipole interaction between cAMP and $Tyr^{371}$ is important. The cAMP binding to site B will also influence the position of the $NH_2$-terminus of αC:B. Two residues here, $Cys^{360}$ and $Ile^{363}$, are directly involved in the hydrophobic interactions between the two domains. The equivalent of $Cys^{360}$ in RII is more accessible in the holoenzyme than in the dissociated $R_2(cAMP)_4$ (*43*) (Fig. 6). As indicated earlier, this hydrophobic region is linked in multiple ways to cAMP binding site A with $Asp^{267}$ and $Trp^{260}$ being of particular importance. Replacement of $Arg^{241}$, which binds directly to both $Asp^{267}$ in domain B and $Glu^{200}$ in domain A, with Ser or Lys demonstrated that this residue is a key feature for the allosteric coupling between sites A and B (*38*). Thus communication between the two domains is likely carried out through the C helix of domain B to the interdomain hydrophobic interaction region and transmitted to site A through residues $Asp^{267}$, $Arg^{241}$, $Glu^{200}$, and $Trp^{260}$ or vice versa.

When holoenzyme forms, the initial docking of C to R involves interactions of the autoinhibitor site in R with the active site cleft in C. This autoinhibitor site in the free R subunit is extremely labile to proteolysis (*5*, *6*) and, in our structure, the first 21 residues are disordered. This initial interaction is essential; however, it is not sufficient to convey high affinity binding. To achieve high affinity binding requires interactions involving the surface of C that lies COOH-terminal to the consensus site peptide. This is the surface that surrounds and includes the essential phosphorylation site, $Thr^{197}$ (*44*). Point mutation of the P-Thr itself as well as the basic residues that bind the phosphate all interfere with R subunit binding. The second step in forming holoenzyme would thus be the docking of the C subunit, with its active site cleft occupied, to a region on cAMP binding domain A. With this structure we can now begin to model these interaction sites.

## REFERENCES AND NOTES


1. E. G. Krebs, *Bioscience Rep.* **13**, 127 (1993).
2. S. K. Hanks, *Curr. Biol.* **1**, 369 (1991).
3. D. A. Walsh, J. P. Perkins, E. G. Krebs, *J. Biol. Chem.* **243**, 3763 (1968).
4. E. W. Sutherland and T. W. Rall, *ibid.* **232**, 1077 (1958).
5. S. S. Taylor, J. A. Buechler, W. Yonemoto, *Annu. Rev. Biochem.* **59**, 971 (1990).
6. J. D. Scott, *Pharmacol. Ther.* **55**, 123 (1991).
7. E. A. Nigg, *Adv. Cancer Res.* **55**, 271 (1990). D. A. Fantozzi *et al.*, *J. Biol. Chem.* **269**, 2676 (1994).
8. D. Øgreid and S. O. Døskeland, *FEBS. Lett.* **129**, 287 (1981).
9. D. Øgreid, R. Ekanger, R. H. Suva, J. P. Miller, S. O. Doskeland, *Eur. J. Biochem.* **181**, 19 (1989).
10. I. T. Weber, T. A. Steitz, J. Bubis, S. S. Taylor, *Biochemistry.* **26**, 343 (1987).
11. L. Saraswat, G. E. Ringheim, J. Bubis, S. S. Taylor, *J. Biol. Chem.* **263**, 18241 (1988). F. W. Herberg, W. R. Dostmann, M. Zorn, S. J. Davis, S. S. Taylor, *Biochemistry* **33**, 7485 (1994).

11a. D. R. Knighton *et al.*, *Science* **253**, 407 (1991); *ibid.*, p. 414.

12. T. A. Kunkel, K. Benebek, J. McClary, *Methods Enzymol.* **204**, 125 (1991).
13. Y. Su, S. S. Taylor, W. R. Dostman, N.-h. Xuong, K. I. Varughese, *J. Mol. Biol.* **230**, 1091 (1993).
14. B. W. Matthews, *ibid.* **33**, 491 (1968).
15. N.-h. Xuong, C. Nielson, R. Hamlin, D. Anderson, *J. Appl. Crystallogr.* **18**, 342 (1985).
16. W. Furey and S. Swaminathan, *Am. Crystallogr. Assoc. Mtg. Abstr. Ser. 2.* **18**, 73 (1990).
17. B.-C. Wang, *Methods Enzymol.* **115**, 90 (1985).
18. C. Cambillau and E. Horjales, *J. Mol. Graphics* **5**, 174 (1987).
19. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr.* **47**, 110 (1901).
20. A. T. Brünger, *XPLOR Version 3.1 Manual* (Yale University, New Haven, CT, 1993).
21. R. Read, *Acta Crystallogr.* **A42**, 140 (1986).
22. R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **26**, 283 (1993).
23. K. Baldridge and W. R. G. Dostmann, unpublished data.
24. A. Kolb, S. Busby, H. Buc, S. Garges, S. Adhya, *Annu. Rev. Biochem.* **62**, 749 (1993).
25. J. B. Shabb and J. D. Corbin, *J. Biol. Chem.* **267**, 5723 (1992).
26. T. Lincoln and J. D. Corbin, *Adv. Cyclic Nucleic Res.* **15**, 139 (1983).
27. J. B. Shabb, B. D. Buzzeo, L. Ng, J. D. Corbin, *J. Biol. Chem.* **266**, 24320 (1991).
28. J. Bubis and S. S. Taylor, *Biochemistry.* **24**, 2163 (1985); *ibid.* **26**, 3478 (1987).
29. A. R. Kerlavage and S. S. Taylor, *J. Biol. Chem.* **255**, 8483 (1980).
30. G. E. Ringheim, L. D. Saraswat, L. D. Bubis, S. S. Taylor, *ibid.* **263**, 18247 (1988).
31. J. Bubis and S. S. Taylor, *Biochemistry.* **26**, 5997 (1987).
32. R. J. W. DeWit *et al.*, *Eur. J. Biochem.* **152**, 255 (1984); T. S. Yagura and J. P. Miller, *Biochemistry* **20**, 879 (1981); B. Jastorff, J. Hoppe, M. Morr, *Eur. J. Biochem.* **101**, 555 (1979).
33. J. D. Corbin *et al.*, *Eur. J. Biochem.* **125**, 259 (1982); D. Øgreid *et al.*, *ibid.* **150**, 219 (1985); S. R. Rannels and J. D. Corbin, *J. Biol. Chem.* **255**, 7085 (1980).
34. P. J. M. Van Haastert *et al.*, *ibid.* **259**, 10020 (1984); W. Dostmann *et al.*, *ibid.* **265**, 10484 (1990).
35. T. E. Creighton, *Proteins* (Freeman, New York, 1984), p. 139.
36. P. A. Frey and R. D. Sammons, *Science* **228**, 541 (1985).
37. G. S. McKnight *et al.*, *Rec. Prog. Hormone Res.* **44**, 307 (1988). K. A. Gorman and R. A. Steinberg, *Somat. Cell Mol. Genet.* **20**, 301 (1994).
38. M. M. Symcox, R. D. Cauthron, D. Ogreid, R. A. Steinberg, *J. Biol. Chem.* **269**, 23025 (1994).
39. R. A. Steinberg, K. B. Gorman, D. Ogreid, S. O. Doskeland, I. T. Weber, *ibid.* **266**, 3547 (1991); T. A. Woodford, L. A. Correll, G. S. McKnight, *ibid.* **262**, 13321 (1987); E. Duprez *et al.*, *ibid.* **268**, 8832 (1993).
40. J. J. Neitzel, W. R. G. Dostmann, S. S. Taylor, *Biochemistry* **30**, 733 (1991).
41. G. E. Ringheim and S. S. Taylor, *J. Biol. Chem.* **265**, 4800 (1990).
42. J. Bubis, L. D. Saraswat, S. S. Taylor, *Biochemistry* **27**, 1570 (1988).
43. N. Nelson and S. S. Taylor, *J. Biol. Chem.* **258**, 10981 (1983).
44. C. S. Gibbs, D. R. Knighton, J. M. Sowadski, S. S. Taylor, M. J. Zoller, *ibid.* **267**, 4806 (1992); S. A. Orellana, P. S. Amieux, X. Zhao, G. S. McKnight, *ibid.* **268**, 6843 (1993).
45. A. G. W. Leslie in joint CCCP4 and ESF-EACMB Newsletter on Protein Crystallography No. 26 (Daresbury Laboratory, Warrington, UK, 1992).
46. D. E. Tronrud, L. F. Ten Eyck, B. W. Matthews, *Acta Crystallogr.* **A43**, 489 (1987).


47. Supported in part by NIH grants GM34921 (S.S.T.) and RR01644 (N-h.X.), the Lucille P. Markey Charitable Trust (S.S.T., L.T.E., N-h.X.), NSF grant BIR-9223760 (L.T.E.), and Public Health Service Training Grant GM07313 (Y.S.). We thank the San Diego Supercomputer Center for use of the Advanced Scientific Visualization Laboratory and the Cray C90, the Stanford Synchrotron Radiation Laboratory for providing us beam time for data collection, and Xiaoping Dai for his help in data collection and data processing. Coordinates have been deposited on the Brookhaven Protein Data base, with tracking number IRGS.

15 February 1995; accepted 28 June 1995