A Structural Model of the Catalytic Subunit-Regulatory Subunit Dimeric Complex of the cAMP-dependent Protein Kinase*

Received for publication, October 26, 2001, and in revised form, December 19, 2001
Published, JBC Papers in Press, January 17, 2002, DOI 10.1074/jbc.M110298200

Chang-Shung Tung‡, Donal A. Walsh§, and Jill Trewhella¶¶

From the ‡Theoretical Biology and Biophysics Group and §Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545 and the ¶Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, California 95616

Previous neutron scattering studies elaborated the topographical relationship of the regulatory (R1α) and catalytic (Cγ) subunits of the cAMP-dependent protein kinase. We present here the results of a set of computations that lead to an atomic model of the cAMP-dependent protein kinase heterodimer, Δ1–90αRIα–Cγ. The first step in the modeling utilized the crystal structures for the porcine Cγ and bovine Δ1–90αRI, or rat Δ1–111αRII, to homology-model structures of the species and isoforms that had been used in the neutron scattering experiments (bovine Cγ subunit and murine Δ1–93αRI, subunit, respectively). A docking procedure, constrained by the dimensions and positions of the ellipsoids in the neutron-derived R-C model as well as mutagenesis data, was used to develop “best fit” models for the heterodimer. Simulated annealing, molecular dynamics, and energy minimization were then used to refine the side chain packing at the heterodimer interface. For comparison, the calculations were done using the homology models derived from both the RIα and RIβ crystal structures. Both resultant models had many similarities. Each predicted similar interfaces. The RIα-based model has 25% more hydrogen bonds than that based on RIβ with seven of these potential bonds in common. The distribution of hydrophobic, polar, and charged residues at the interface was similar for both models, with a distribution more characteristic of the exposed surface residues than those in the protein interior. The calculated interface area in each is relatively small (<2000 Å²). The RIα-based model, however, has a significantly better fit with the scattering data and is therefore the one of distinctly higher probability. With its small interface area that has a high proportion of charged and polar residues, the complex appears poised for dissociation, and each subunit existing as a stable entity. This result is consistent with the known physiological events required for cAMP-dependent activation of the kinase.

The cAMP-dependent protein kinase (PKA) 1 is a multifunctional kinase that serves as a prototype for understanding second messenger signaling and protein phosphorylation (1–3). In the absence of a cAMP signal, the enzyme is inactive and exists as a dimer of dimers, having two identical regulatory (R) and two identical catalytic (C) subunits. The C subunit has the conserved catalytic core structure that is common to the majority of protein kinases. The R subunit has a dimerization domain, followed by a pseudosubstrate sequence that inhibits C subunit activity and two in tandem cAMP-binding domains (in order from the N to C termini). When two cAMP molecules bind to each of the R subunits, the C subunit is activated, presumably via some sort of release of the pseudosubstrate sequence. The activation of the C subunit has been assumed to involve dissociation of the C subunits from the R2 homodimer (4–6). The strongest evidence for the dissociation of the C subunit is the fact that the C subunit, but not Rα, is known to cross the nuclear membrane in order to phosphorylate a number of its target proteins (7–10). Fluorescence energy transfer measurements performed under equilibrium conditions have indicated, however, that full dissociation of C from R2 may not be required for activation and catalysis of phosphorylation (11, 12).

There are three isoforms of C (α, β, and γ) and two major isoforms of R (I and II) that are further subdivided into subforms (α and β) (13). The crystal structure of Cγ shows that it has the now classical protein kinase bilobal structure, with the catalytic cleft located between the two lobes (14–16). Both small angle solution scattering (1, 17) and crystallographic studies (18) have shown that binding of peptides that model either substrate or pseudosubstrate sequences results in closure of the catalytic cleft. This closure is achieved via a hinge movement about a conserved Gly residue (Gly125) located in the sequence that joins the two lobes.

The crystal structures of the cAMP-binding domain of both the bovine R1α subunit and the rat R1β have been solved using the truncation mutants Δ1–90RIα and Δ1–111RIβ, respectively, each of which is missing the dimerization domain but contains part of the pseudosubstrate sequence and the two in tandem cAMP-binding domains (19, 20). Only the cAMP-binding domains plus a small part of the interconnecting peptide to the pseudosubstrate sequence are seen in the crystal structures. These crystal structures reveal very similar structures for the individual cAMP-binding domains for each R isoform but with the orientation between the two sites in Rα differing from that seen with R1β, resulting in the latter having a more compact structure.

Based upon deuterium labeling and neutron scattering con-
trast variation experiments, we previously derived low resolution models for both an R-C heterodimer and the PKA holoenzyme (R₂C₂) (1). A two-ellipsoid model was developed for the Δ¹⁻⁹¹R₁₂₅-C₂₅₉ heterodimer that provided the overall dimensions of each subunit within the complex and their topographical relationship. Similarly, a model for the R₂C₂ holoenzyme was derived. The holoenzyme was shown to have a dumbbell shape, with the R₂ dimerization domain forming the bar of the dumbbell connecting the two globular cAMP-binding lobes, each associated with a C subunit ellipsoid. The centers of mass of the two C subunits are 120 Å apart such that they do not interact.

The previously solved crystal structures of the C subunit and of the cAMP-binding domain for R fit reasonably well into the neutron models, showing good agreement between the model ellipsoid dimensions and the overall dimensions of the crystal structures. Of note, based on recent evaluations, the more extended R₃ crystal structure (19) fits significantly better into the neutron ellipsoid model compared with the more compact structure recently solved for R₁₂₅ (20).

In this paper, we describe the results of a set of computations that arrive at an atomic model of the Δ¹⁻⁹¹R₁₂₅-C₂₅₉ heterodimer. The data used to develop this model include the neutron scattering-derived two-ellipsoid model for the heterodimer, crystal structures of the R and C subunits, and mutagenesis data identifying points of subunit-subunit contact. The model provides a detailed mapping of the subunit interface and an understanding as to why both nondissociated PKA and the dissociated subunits are each stable in aqueous solution.

MATERIALS AND METHODS

Sequence Alignment—Sequence alignments were performed by running the program ALIGN from GeneStream (available on the World Wide Web at www2.igh.cnrs.fr/bin/align-guess.cgi).

Homology Modeling—Homology modeling is a technique that uses the known structure (the template) of a protein with a homologous sequence to predict the structure of a protein of interest (the target). The details of our homology modeling method are described elsewhere (21). To model the main chain structure of the target molecule, those regions aligned with no insertions or deletions between the target and the template molecules are assigned the structure of the corresponding regions in the template molecule. To model the insertions, we used an algorithm we developed previously to model loops (22, 23). This algorithm is particularly efficient in the sampling of loop structures with fixed ends (22). To model the side chains, we initially assume that the residues in the target occupy the same space as those in the template. This consideration is very useful when the target protein is folded into a compact structure.

Molecular Docking—In the docking procedure, the two subunits, as structured by the homology modeling, were treated as rigid bodies, and only the C-α coordinates were used. Amino acids were represented as 4-Å radius spheres centered at their respective C-α coordinate positions. The docking procedure was divided into two steps: the first being the coarse docking of the two subunits into the ellipsoid model for the complex derived from the neutron scattering study (1), while the second step is a detailed search in the conformational space around the structures derived from the coarse docking procedure.

The coarse docking was achieved by matching the centers of mass and principal moment axes between the two homology-modeled subunits and the ellipsoid of the neuron-derived model. The three principal moment axes of a molecule were derived by diagonalizing the transformation matrix defined as follows (24),

$$I = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}$$

(Eq. 1)

where the diagonal elements are known as the moment of inertia coefficients, and have the form

$$I_{ii} = \sum_{i=1}^{n} m_i (r_i^2 - x_i^2)$$

(Eq. 2)

where $m_i$ represents the mass of the $i$th atom, and $r_i$ is its position. The off-diagonal elements are designated as products of inertia, a typical one being the following,

$$I_{ij} = -\sum_{i=1}^{n} m_i x_i y_i$$

(Eq. 3)

Because of the rotational symmetries of the ellipsoids representing the subunits in the neutron-derived model, there are 16 possible ways to dock the subunits in this coarse docking step.

To further delineate those selected structures, we defined three empirical energy terms: $E_{vdw}, E_{ele}$, and $E_{c2}$. $E_{vdw}$ is a measure of the van der Waals energy with parameters corresponding to a van der Waals radius of 4.0 Å and a depth of 0.12 Kcal/mol. $E_{c2}$ can be calculated as follows,

$$E_{c2} = 100 \cdot (d_0 - 6)^2 + 10 \cdot \sum_{i=1}^{6} q_i$$

(Eq. 4)

where $d_0$ is the distance between the two residues that the experiment has determined to have ion pair (25), and $d_i$ to $d_6$ are the shortest distances between the six residues identified as being in close contact with the interface region (R: Asp141, Lys347; C: Thr135, Trp296, Thr397, Lys317; see below) and the other subunit with $q_i$ defined as follows,

$$q_i = (d_i - 8)^2 \text{ if } d_i > 8$$

(Eq. 5)

$$q_i = 0 \text{ if } d_i \leq 8$$

(Eq. 6)

$E_{c2}$ is a parameter that measures the goodness of fit between the structure and the ellipsoid model. $E_{c2}$ is defined to be the number of residues that are external to the ellipsoid model, in the structure that is external to the ellipsoid model is attributed with 1 Kcal/mol of energy.

Mutagenesis Data Used to Constraine or Evaluate Models—Residues were identified from mutagenesis studies with either yeast or mammalian enzyme as either involved or not involved in the R-C interaction; the following residue numbers are for bovine C₂ and murine R₃ sequences. Residues on the R₂ (R₁₂₅, R₁₈₀) subunit affecting the R-C interaction (25, 33) are as follows: Asp141, Asp160, Asp187, Glu143/Glu144 (Glu143, Glu150/Glu159), Asp171 (Asp170, Asp187), Lys247 (Lys242, Lys263), Glu263 (Asp260, Lys275). The asterisk indicates Glu residues immediately following the Gln residues that would be aligned with Glu143 of RI. What are the key findings and conclusions of this study?
ing the target structure based on the template structure was therefore simple and straightforward. The porcine crystal structure (16) is for C complexed with PKI peptide, and hence the catalytic cleft is in its “closed” conformation (17). Since the binding of the R subunit pseudosubstrate domain to the C subunit also results in a closed catalytic cleft (1), this model is the best choice for constructing an R-C heterodimer.

**Modeling the Regulatory Subunit**—There are two published crystal structures of the R subunit: bovine $\Delta^{1-90}R_{\text{II}}$ (19) (1RGS in the Protein Data Bank) and rat $\Delta^{1-111}R_{\text{II}}$ (20) (1CX4 in the Protein Data Bank). The initial residues from the N terminus of these deletion mutants, including those of the pseudosubstrate sequence, are missing from the crystal structure coordinates, presumably because they are disordered in the crystals. The homology modeling therefore only included residues starting from 112 for bovine $R_{\text{II}}$ and 139 for rat $R_{\text{II}}$. Each of these structures was used, in turn, to develop an $R_{\text{II}}$-$C$ model in order to compare and contrast the derived heterodimer models and to optimize the final structure against the combined experimental data. The murine $\Delta^{1-90}R_{\text{II}}$ subunit used for previous neutron scattering experiments of PKA heterodimer and holoenzyme was the target sequence (29) (GenBank™ accession number J02935) of the cAMP-dependent protein kinase is aligned with those of the bovine $R_{\text{II}}$ and rat $R_{\text{II}}$ (the templates; sequences and crystal structures from the Protein Data Bank 1RGS and 1CX4, respectively). The murine $R_{\text{II}}$ and bovine $R_{\text{II}}$ show 43% sequence identity, 22% conservative substitutions, and three insertions, while murine $R_{\text{II}}$ and rat $R_{\text{II}}$ show 73% identity, 8% conservative substitutions, and two insertions. The homology-modeled structures of the $R_{\text{II}}$, subunit paired with each crystal structure ($R_{\text{II}}$, (left pair) and $R_{\text{II}}$, (right pair)) are shown at the bottom, with the inserted sequence segments in yellow. For orientation, the site of interaction with the C subunit, as determined by this work and our previous study (1), is on the lower back side of the R as it is shown here.

![Image](https://example.com/image.png)

**Fig. 1.** The sequence of the murine $R_{\text{II}}$ subunit (the target, sequence from GenBank™, accession number J02935) of the cAMP-dependent protein kinase is aligned with those of the bovine $R_{\text{II}}$ and rat $R_{\text{II}}$ (the templates; sequences and crystal structures from the Protein Data Bank 1RGS and 1CX4, respectively). The murine $R_{\text{II}}$ and bovine $R_{\text{II}}$ show 43% sequence identity, 22% conservative substitutions, and three insertions, while murine $R_{\text{II}}$ and rat $R_{\text{II}}$ show 73% identity, 8% conservative substitutions, and two insertions. The homology-modeled structures of the $R_{\text{II}}$, subunit paired with each crystal structure ($R_{\text{II}}$, (left pair) and $R_{\text{II}}$, (right pair)) are shown at the bottom, with the inserted sequence segments in yellow. For orientation, the site of interaction with the C subunit, as determined by this work and our previous study (1), is on the lower back side of the R as it is shown here.

The alignments of the template and target sequences are shown in Fig. 1 (upper panel). The homology-modeled structures of the target molecule as well as the structure of the template molecules are shown at the lower panel of Fig. 1. The root mean square deviations between the main chain atoms in the aligned residues of the template and the target molecules are small (<1 Å) for both the $R_{\text{II}}$ and $R_{\text{II}}$ templates. The insertions in the target molecules are shown in yellow. The insertions are not close to the interface area and hence will not influence the modeling of the interface. The template R subunits used in both cases here were the cAMP-bound crystal structures (19, 20). Our previous x-ray scattering data show that there is no change in the overall shape of $R_{\text{II}}$, upon cAMP binding (1), suggesting that the cAMP-bound protein is a suitable template structure to determine the target R subunit structure.

The structure of $R_{\text{II}}$, derived from neutron scattering studies appears to better match the crystal structure for $R_{\text{II}}$. The crystal structures for $R_{\text{II}}$ and $R_{\text{II}}$ differ predominantly in the relative orientations of the two cAMP-binding domains, which are otherwise individually similarly structured in the two forms. This difference in orientation results in $R_{\text{II}}$ having a significantly more compact structure. Overlay of the neutron-derived ellipsoid model for $R_{\text{II}}$, with the two crystal structures shows a better fit for $R_{\text{II}}$, with the more bent $R_{\text{II}}$ structure protruding more out of the highly asymmetric, cigar-shaped ellipsoid model for $R_{\text{II}}$. Thirty-five more residues were external to the neutron scattering-derived ellipsoid for $R_{\text{II}}$. The structure of $R_{\text{II}}$, derived from neutron scattering studies appears to better match the crystal structure for $R_{\text{II}}$. The crystal structures for $R_{\text{II}}$ and $R_{\text{II}}$ differ predominantly in the relative orientations of the two cAMP-binding domains, which are otherwise individually similarly structured in the two forms. This difference in orientation results in $R_{\text{II}}$ having a significantly more compact structure. Overlay of the neutron-derived ellipsoid model for $R_{\text{II}}$, with the two crystal structures shows a better fit for $R_{\text{II}}$, with the more bent $R_{\text{II}}$ structure protruding more out of the highly asymmetric, cigar-shaped ellipsoid model for $R_{\text{II}}$. Thirty-five more residues were external to the neutron scattering-derived ellipsoid for $R_{\text{II}}$. The structure of $R_{\text{II}}$, derived from neutron scattering studies appears to better match the crystal structure for $R_{\text{II}}$. The crystal structures for $R_{\text{II}}$ and $R_{\text{II}}$ differ predominantly in the relative orientations of the two cAMP-binding domains, which are otherwise individually similarly structured in the two forms. This difference in orientation results in $R_{\text{II}}$ having a significantly more compact structure. Overlay of the neutron-derived ellipsoid model for $R_{\text{II}}$, with the two crystal structures shows a better fit for $R_{\text{II}}$, with the more bent $R_{\text{II}}$ structure protruding more out of the highly asymmetric, cigar-shaped ellipsoid model for $R_{\text{II}}$. Thirty-five more residues were external to the neutron scattering-derived ellipsoid for $R_{\text{II}}$.
pared with R_{II}. The compaction of R_{II} can be seen in the radius of gyrations ($R_g$) and maximum linear dimensions ($d_{max}$) calculated for the crystal structures ($R_{II} R_g = 19$ Å and $d_{max} = 63$ Å; $R_{II} R_g = 21$ Å and $d_{max} = 72$ Å, respectively). The higher values for $R_{II}$ also better match the values for $R_{II}$ obtained from the scattering studies for $R_{II}$ ($R_g = 23$ Å and $d_{max} = 86$ Å) (1), although the scattering data include contributions from the pseudosubstrate region that is missing from the crystal structures; hence, direct comparison is not possible. Despite these apparent differences, we nonetheless chose to pursue modeling the $R_{II}$-C$_\alpha$ interaction based upon both R crystal structures at this stage for the purposes of comparison and to see how robust our modeling approach is.

**Mutagenesis Data Define Contact Residues between Two Subunits**—From studies with either homologous mammalian or yeast enzymes (25, 30–32), five residues on the C$_\alpha$ subunit have been identified as probably being involved in the R-C interaction: Thr$^{195}$, Trp$^{196}$, Thr$^{197}$, Lys$^{213}$, and Lys$^{217}$ (see “Materials and Methods”). These are residues that are not part of the pseudosubstrate binding domain; nor do they directly contribute to phosphotransferase activity. Similar mutation studies have identified five residues on the R$_{II}$ subunit as involved in binding C (25, 33): Asp$^{140}$, Glu$^{143}$, Asp$^{172}$, Lys$^{242}$, and Asp$^{263}$ (see “Materials and Methods”). From the aligned sequences, these residues map to Asp$^{141}$, Glu$^{144}$, Asp$^{171}$, Lys$^{247}$, and Glu$^{263}$ on the R$_{II}$ subunit and Asp$^{157}$, Glu$^{160}$, Asp$^{197}$, Lys$^{263}$, and Glu$^{279}$ on the R$_{II}$ subunit. Compensatory mutation studies further indicate a direct electrostatic bridge between the subunits involving Glu$^{143}$ on R$_{II}$ and Lys$^{213}$ on C$_\alpha$ (25). Residue Glu$^{143}$ of R$_{II}$ maps to Glu$^{144}$ of R$_{II}$ and Glu$^{160}$ of R$_{II}$ in the aligned sequences, and because Glu is not charged, it cannot form an ion pair. However, the residue immediately 5’ to this Glu residue is a Gln on both R$_{II}$ isoforms (Glu$^{143}$ on R$_{II}$ and Glu$^{159}$ on R$_{II}$) and therefore in both an ideal candidate for forming an ion pair with Lys$^{213}$ on C$_\alpha$. This ion pair was used as a key anchor in the modeling to determine the initial orientation of the R and C subunits with respect to each other.

**Modeling the Heterodimer**—Two full sets of computations were completed to develop alternate models for the R$_{II}$-C$_\alpha$ interaction. The first used the R$_{II}$ crystal structure based homology model for R$_{III}$, while the second followed the same protocol but used the R$_{II}$ crystal structure to model R$_{II}$. Described now in detail are the computations for R$_{II}$-based structure.

The first step in docking the R and C subunits was a coarse docking in which the centers of mass and principal moment axes were determined for the homology-modeled murine $\Delta^{1–90}$R$_{II}$ and bovine C$_\alpha$ coordinates (see “Materials and Methods”). These axes and centers are illustrated in Fig. 2 in relationship to the C$_\alpha$ main chain structures for R (upper left) and C (upper right). Each subunit could then be aligned with the center of mass and principal moment axes of its respective ellipsoid in the heterodimer model derived from the neutron scattering data (1). This type of relationship is illustrated in the lower center of Fig. 2. With its highly asymmetric ellipsoid, the placement of R could be done with more precision compared with C, which is more globular and symmetric overall.

Due to the rotational symmetries of the ellipsoids in the neutron model, each of the subunits can be placed in four different orientations (by rotating 180° with respect to each of the principal moment axes), giving a total of 16 possible configurations for the dimer. Of these 16 configurations, however, only four allow the R-Glu$^{143}$ and C-Lys$^{213}$ ion pair to form, thereby restricting the number of possible configurations from this first modeling iteration to the four configurations (A–D) illustrated in the lower part of Fig. 2.

The second step in the docking was a detailed search in the configuration space around the four heterodimer structures...
derived from the coarse docking procedure and the ion pair constraint. To speed up the computations, we used a reduced coordinate representation for the subunit structure that includes an extended atom representation (at the C-α coordinate positions) for the residues. Each subunit is treated as a rigid body and is sampled on a grid associated with the six degrees of freedom (three translational and three rotational with respect to \( x, y, \) and \( z \) axes).

The grid step size for each degree of freedom was chosen such that movements from grid point to grid point result in a displacement of either of the two residues that form the ion pair of \( \sim 2 \text{ Å} \). The initial grid search was performed with one of the subunits moving to different grid points, while the other subunit was held in its initial orientation obtained from the coarse docking. The ranges explored for the six degrees of freedom for the R subunit were as follows: \(-12 \) to \( 12 \text{ Å} (2 \text{ Å})\), \(-6 \) to \( 6 \text{ Å} (2 \text{ Å})\), \(-6 \) to \( 8 \text{ Å} (2 \text{ Å})\), \(-20 \) to \( 40 \)° (10°), \(-25 \) to \( 25 \)° (5°), and \(-42 \) to \( 42 \)° (6°), with the numbers in parentheses indicating the step sizes. A total of 840,840 grid positions for the R subunit were sampled with the C subunit fixed. We further define the ion pair distance, \( d_\alpha \), as the C-α to C-α distance between R-Glu and C-Lys and define \( N_{1E} \) to be the number of residues in the R subunit that are external to the ellipsoid model in the docked configuration. Therefore, the parameter set \((d_\alpha, i, \Delta N_{1E}(i))\) provides a measure for the goodness of docking with an index \( i \) corresponding to the \( i \)th configuration of the R subunit. The parameter \( \Delta N_{1E}(i) \) is defined to be \( N_{1E}(i) - N_{1E}(0) \) where \( N_{1E}(0) \) corresponds to the R subunit orientation and position from the initial coarse docking. If we define a set of constraints such that \( d_\alpha \leq 20 \text{ Å} \) and \( \Delta N_{1E}(i) \leq 10 \) residues, only 2353 of the configurations evaluated for the R subunit satisfied them.

The same kind of initial position and orientation search was done for the C subunit. The ranges for the six degrees of freedom for the C subunit were as follows: \(-10 \) to \( 12 \text{ Å} (2 \text{ Å})\), \(-8 \) to \( 10 \text{ Å} (2 \text{ Å})\), \(-8 \) to \( 10 \text{ Å} (2 \text{ Å})\), \(-45 \) to \( 45 \)° (5°), \(-45 \) to \( 45 \)° (15°), and \(-48 \) to \( 48 \)° (6°), with the numbers in parentheses indicating the step sizes. A total of 2,713,200 conformations for the C subunit were sampled while the R subunit remained in its original orientation obtained from the coarse docking; the larger number compared with the search for R reflected the greater uncertainty in placing the more symmetric C subunit in its ellipsoid. When the same set of constraints (20 and 10 for \( d_\alpha \) and \( \Delta N_{2E}(i) \), respectively) was applied as for the R calculation, 10,051 C subunit configurations were selected. These two sets of subunit configurations (2353 for the R subunit and 10,051 for the C subunit) give a total of 23,650,003 possible configurations for the heterodimer. Each of these configurations was then tested against mutagenesis data that give information on close contacts between the subunits.

For the first stage of screening the >23 million possible configurations for R and C, we imposed three simple structural considerations: 1) the van der Waals exclusion rule that says no pairwise C-α distance should be less than 3 Å; 2) the close contact rule requiring that the minimum distances, \( d_1-d_6 \), between the close contact residues and the corresponding subunit should be less than 12 Å; and 3) the distance between the residues that form the ion pair should be less than 10 Å. No heterodimer derived from initial selected configuration A or B (see Fig. 2) passed this initial screening. The same initial screening produced 161 and 50,150 heterodimer structures from initial selected subunit configurations C and D (Fig. 2), respectively.

To further delineate the remaining structures, empirical energetic parameters \( E_{vdw}, E_{C1}, E_{C2} \) (as described under “Materials and Methods”) were used for the second stage of screening. With the limits of these three parameters arbitrarily set at 200 each, 260 structures were selected. All of these 260 structures were derived from initial structure configuration D (Fig. 2). By setting the pairwise root mean square deviation for the representative structures to be <5 Å, the 260 structures could be collated into a set of five families, each member of which had with a very closely related structure. One structure representing each family is shown in Fig. 3. Of the 260 structures, the one that has the best fit in terms of best satisfying the mutagenesis data defining interacting residues (corresponding to the smallest \( E_{C1} \)) was chosen as the optimum R \(_\alpha\)-based structure derived up to this stage in the modeling, which was then be subjected to further refinement. This structure is presented in the upper half of Fig. 4. The root mean square deviation between this best fit and the next best fit structure was 1.1 Å, with the differences in \( d_\alpha \) ranging from 0.1 to 0.5 Å. Of note, this structure belonged to the family with the most structures (family D-4 in Fig. 3, having 111 structures of the total 260).

The same procedure as described above to optimize the modeling of the structure based upon the R \(_\alpha\) crystal structure was also used to derive a best fit structure for an R \(_\alpha\)-C \(_\alpha\) heterodimer based upon the R \(_\alpha\) crystal structure. Because the R \(_\alpha\) homology model derived from R \(_\alpha\) is a significantly worse fit to the neutron-derived ellipsoid model compared with R \(_\alpha\) (with 35 more residues external to the ellipsoid for R \(_\alpha\) compared with R \(_\alpha\) and the already noted divergence in \( R_p \) values...
Evaluation of Models with Respect to Other Data—The proposed site for the R-C interaction in the R_{1α}-based (and R_{1γ}-based) model is supported by the observation that that anti-peptide antibodies to C subunit fragments 187–199 and 187–205 blocked the R-C interaction (34). The proposed interface site is also compatible with the extensive set of mutation studies of yeast PKA C subunit homolog that identified a group of five residues involved in R subunit interaction and 31 C subunit surface residues unlikely to be at the R-C subunit interface (30, 31) (see “Materials and Methods”). Fig. 4 shows the two derived models with residues identified by these mutagenesis studies highlighted as gray spheres (involved) or orange spheres (noninvolved), respectively. In both heterodimer structures, all but one (Asp^{276}) of these residues on the C subunit identified as unlikely to be involved in the R-C interaction (see “Materials and Methods”), are at least 10 Å away from those in the R subunit. The minimum distances between the C-α of Asp^{276} and all C-α atoms in the R subunit is 9.5 Å in the R_{1α}-based (and 8.3 Å for the R_{1γ}-based model). Close inspection of the detailed C subunit structure also indicates that the two hydrogen bond acceptors of Asp^{276} form hydrogen bonds with Thr^{277} and Lys^{279} of the same subunit; thus, it is unlikely that Asp^{276} can be involved in specific recognition between the two subunits. As we have noted previously, two of the R subunit residues (Asp^{171} and Glu^{265}) identified (1) by the mutation studies as important contributors to R-C interaction are buried in the R subunit away from the RC interface. This occurs in both models (Fig. 4). Most likely, these two residues are critically involved in the transmission of the signal from cAMP binding rather than in direct R-C interaction. Based upon studies using the yeast two-hybrid screening approach, Huang and Taylor (35) identified the sequence R_{1α}^{236–244} (equivalent to R_{1α}^{240–248}) as an important contributor to C subunit binding; probably by direct C subunit interaction. In either model depicted in Fig. 4, these residues are at the R-C interface (denoted by magenta).

Construction of the Full Atomic Models—Using the models shown in Fig. 4, the full atomic structures for the R_{1α}-based and R_{1γ}-based models of the R-C heterodimer were constructed by adding to the defined C-α coordinates the side chain coordinates for the R and C subunits that had been obtained by homology modeling of the bovine C_{1α} and murine Δ^{391–91}R_{1γ} (Fig. 1) subunits. These full atomic models of the complex were then subjected to energy minimization using AMBER. The energy-minimized structure for the preferred R_{1α}-based complex is
shown in two orientations in Fig. 5. A barrel and ribbon depiction for the Rα-based complex and, for comparison, the Rββββ-based complex are shown in Fig. 6. A detailing of the ion pair formed between Glu143 of the R subunit and Lys213 of the C subunit for the Rα-based structure is shown in Fig. 7. The coordinates for the Rα and Rββββ-based models of the Rα-Cα heterodimer are in the Protein Data Bank (Rα-based model, Protein Data Bank entry 1KMU; Rββββ-based model, Protein Data Bank entry 1KMW).

Characterizing the R-C Interface—To gain insight into the nature of the R-C interactions, we examined three aspects: the contact surface area, the amino acid composition, and number of hydrogen bonds at the dimer interface. Contact surface area is often used for evaluating binding strength between two subunits in a complex. In general, the larger the contact surface area is, the stronger the binding. The total surface areas of each of the individual subunits as well as the complex were calculated using the program NACCESS (36). The molecular surface area calculated by NACCESS is defined according to Richards (37). The default radius of the water molecule used as the probe sphere for tracing out the surface was 1.4 Å. The contact surface area between R and C to be 1680 Å² for the Rα-based model and 2046 Å² for the Rββββ-based model. When the R subunit is divided into two structurally distinct cAMP-binding domains (domain A, residues 114–250; domain B, residues 251–383), the interface surface areas between the domains and the C subunit in the Rα-based model are 1277 and 403 Å², respectively. In the Rββββ-based model, the two corresponding interface surface areas are 1427 and 673 Å². In both of the models therefore, domain A would be the high affinity C subunit binding domain.

The distribution of residues was examined for the R-C interface as well as for the interior and exterior of the protein. Interface residues were identified as those with significant changes (>10%) in relative solvent-accessible surface areas upon complex formation. Surface and interior residues were defined as those with greater or less than a 50% relative solvent exposure. The data are summarized in Table I. Both the interface surfaces between the subunits and the surface of the subunits as part of the heterodimer contain a high proportion of charged residues and a clear predominance of hydrophilic (charged plus polar) over hydrophobic residues. In contrast, the interiors of both subunits are strongly hydrophobic. In the Rα-based model, the interface residues on the R subunit carry a net zero charge, whereas those on the C subunit carry a net +5 charge. In the Rββββ-based model, the net charges of the interface residues on the R and the C subunits are −1 and +5, respectively.

The structure of the heterodimer was allowed to relax at room temperature (300 K) in order to evaluate more detailed interactions (e.g., hydrogen bonding between the subunits). This relaxation is accomplished by subjecting the complex to a 10-ps run of molecular dynamics using AMBER. During the simulation, hydrogen bonds associated with the secondary structures (helices and sheets) were maintained through a set of distance constraints (the distance between the amino hydrogen and carboxyl oxygen that formed a secondary structural hydrogen bond to be maintained at 2 Å). This procedure allows for those residues at the interface to sample more of the conformational space without unfolding either of the two subunits. A total of 10 energy-minimized structures were obtained for each of the Rα- and Rββββ-derived models from this annealing procedure. Those hydrogen bonds between the subunits existing in the simulation all the time are tabulated in Table II. By this methods, 20 interface hydrogen bonds were identified for the Rα-based model, most of which involved side chain interactions (Table II). The residues involved included R-Asp141,
that only one of the 23 positions is deeply buried in the protein interior, and this position (109) substitutes a Ser for an Ala, two relatively small residues that are likely to be accommodated easily in the same folded structure. Thus, the sequence differences between the C isoforms are unlikely to significantly impact the R-C interaction.

The R isoforms show more sequence variability than do the C isoforms. Considering first the R\textsubscript{I}-based model, we note there are 16 residues on the R subunit involved in hydrogen bonds. The R\textsubscript{I} and R\textsubscript{II} sequences (Protein Data Bank entry 1RGS and GenBank\textsuperscript{TM} accession number NP032949) show 89% sequence identity with no insertions or deletions, and the predicted hydrogen bonds in the model are all conserved. Alignment of R\textsubscript{I} and R\textsubscript{II} (GenBank\textsuperscript{TM} accession number J02935) shows greater variability; nonetheless, 7 of the 16 residues involved in hydrogen bonds are identical, there is one conservative replacement (Lys to Arg), and 4 residues are flanked by residues that have the potential to form a hydrogen bond. The remaining 4 residues (at positions 240, 254, 300, and 328) are substituted by residues that cannot form hydrogen bonds.

Examining the R\textsubscript{II}-based model, there are 13 residues on the R subunit involved in hydrogen bonds. An alignment of R\textsubscript{I} and R\textsubscript{II} (Protein Data Bank entry 1CX4) shows 10 of the 13 residues to be identical, one conservative replacement (Lys to Arg), and the remaining 2 charged residues (Glu\textsuperscript{276} and Asp\textsuperscript{353}) replaced by a polar residue (Thr in each case) that can form a hydrogen bond.

Thus, overall the sequence comparisons between different isoforms of R and C, when evaluated in the context of either the R\textsubscript{I}-based or R\textsubscript{II}-based model, indicate a high degree of conservation at the interface and that the majority of intersubunit hydrogen bonds are therefore conserved.

The heterodimer model predicts a relatively small interface (<2000 Å\textsuperscript{2}), with a high percentage of charged and polar residues at the interface, a similar distribution of charged and polar residues, similar to what is observed for the solvent exposed surfaces, and a network of at least 20 hydrogen bonds. Jain et al. (39) reviewed 23 oligomeric proteins, evaluated their solvent-accessible surfaces, and calculated the areas of their interfaces. Oligomers with small surfaces have globular subunits with accessible surface areas similar to those of monomeric proteins, and these authors suggested that a small interface (700–2000 Å\textsuperscript{2}) correlates with subunits that maintain their structures when isolated from the oligomer, whereas when there are large interface areas (3000–10,000 Å\textsuperscript{2}), the subunits are probably not stable by themselves. The highly polar and charged interface would allow both for tight interaction and for stability of the individual subunits once separate.

This entire set of factors match the physiological requirements for PKA. Both the PKA holoenzyme and its subunits, once
dissociated, must be stable. The interface appears to be designed to be poised for dissociation and stabilization of the subsequently exposed interface residues in a polar, aqueous environment. The dissociation would be triggered by the release of some key interactions, at the interface and/or perhaps subsequently exposed interface residues in a polar, aqueous dissociated, must be stable. The interface appears to be de-

10. Fantozzi, D. A., Taylor, S. S., Howard, P. W., Maurer, R. A., Feramisco, J. R., and Krebs E. G. (1984) Trends Biochem. Sci. 9, 62–68.
11. Johnson, D. A., Leathers, V. L., Martinez, A. M., Walsh, D. A., and Fletcher, W. H. (1993) Biochemistry 32, 6402–6410.
12. Yang, S., Fletcher, W. H., and Johnson, D. A. (1995) Biochemistry 34, 6267–6271.

REFERENCES
1. Zhao, J., Hoye, E., Boylan, S., Walsh, D. A., and Trewella, J. (1998) J. Biol. Chem. 273, 30448–30459.
2. Walsh, D. A., Perkins, J. P., and Krebs, E. G. (1968) J. Biol. Chem. 243, 3763–3766.
3. Taylor, S. S., Knighton, D. R., Zheng, J., Sowadski, J. M., Gibbs, C. S., and Zoller, M. J. (1993) Trends Biochem. Sci. 18, 84–89.
4. Tsuzuki, J., and Kiger, J. A. (1978) Biochemistry 17, 2961–2970.
5. Chau, V., Huang, L. C., Guillermo, R., Biltonen, R., and Huang, C.-H. (1980) Biochemistry 19, 924–928.
6. Builder, S. E., Beavo, J. A., and Krebs, E. G. (1980) J. Biol. Chem. 55, 3514–3519.
7. Sassone-Corsi, P. (1998) Int. J. Biochem. Cell Biol. 30, 27–38.
8. Lalli, E., and Sassone-Corsi, P. (1984) J. Biol. Chem. 269, 17359–17362.
9. Fantozzi, D. A., Harootunian, A. T., Wen, W., Taylor, S. S., Feramisco, J. R., Tsien, R. Y., Meinkoth, J. L., (1994) J. Biol. Chem. 269, 2076–2086.
10. Fantozzi, D. A., Taylor, S. S., Howard, P. W., Maurer, R. A., Feramisco, J. R., and Meinkoth, J. L. (1992) J. Biol. Chem. 267, 16824–16828.
11. Johnson, D. A., Leathers, V. L., Martinez, A. M., Walsh, D. A., and Fletcher, W. H. (1993) Biochemistry 32, 6402–6410.
12. Yang, S., Fletcher, W. H., and Johnson, D. A. (1995) Biochemistry 34, 6267–6271.

Structure of cAMP-dependent Protein Kinase R-C Heterodimer 12431

13. Dosselkand, S. O., Maronde, E., and Gjersten, B. (1993) Biochim. Biophys. Acta 1178, 249–258.
14. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Ashford, V. A., Xuong, N., Taylor, S. S., and Sowadski, J. M. (1991) Science 253, 417–414.
15. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Xuong, N., Taylor, S. S., and Sowadski, J. M. (1991) Science 253, 414–420.
16. Bossemeyer, D., Engh, R. A., Kinzel, V., Poningting, H. and Huber, H. (1993) EMBO J. 12, 849–859.
17. Olah, G. A., Mitchell, R. D., Sossnick, T. R., Walsh, D. A., and Trewella, J. (1993) Biochemistry 32, 3649–3657.
18. Cox, S., Radzio-Andzelm, R., and Taylor, S. S. (1994) Curr. Opin. Struct. Biol. 4, 893–901.
19. Su, Y., Destmann, W. R., Herberg, F. W., Durick, K., Xuong, N. H., Ten Eyck, L., Taylor, S. S., and Varughese, K. I. (1995) Science 269, 807–813.
20. Diller, T. C., Xuong, N. H., and Taylor, S. S. (2001) Structure 9, 73–82.
21. Tung, C. S. (1999) J. Biol. Struct. Dyn. 17, 347–354.
22. Tung, C. S. (1997) Biochem. J. 32, 876–885.
23. Ryu, K. S., Lee, H. Y., Kim, S. P., Beauchamp, J., Tung, C. S., Issacs, N. W., Ji, L., and Ji, T. H. (1998) J. Biol. Chem. 273, 6285–6291.
24. Goldstein, H. (1950) Classical Mechanics, pp. 149–151, Addison-Wesley Publishing Co., Reading, MA.
25. Gibson, R. M., Ying, J. B., and Taylor, S. S. (1997) Protein Sci. 6, 1825–1834.
26. Weiner, S. J., Kollman, P. A., Case, D. A., Chandra Singh, U., Ghio, C., Algon, G., Profeta, S. Jr., and Weiner, P. (1984) J. Am. Chem. Soc. 106, 765–784.
27. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) J. Comput. Chem. 7, 230–252.
28. Wiemann, S., Kinzel, V., and Pyerin, W. (1992) Biochim. Biophys. Acta 1171, 93–96.
29. Scott, J. D., Glacecum, M. B., Zoller, M. J., Uhler, M. D., Helfman, D. M., McGinty, G. S., and Krebs E. G. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 5192–5196.
30. Gibbs, C. S., and Zoller, M. J. (1991) J. Biol. Chem. 266, 8923–8931.
31. Gibbs, C. S., Knighton, D. R., Sowadski, J. M., Taylor, S. S., and Zoller, M. J. (1992) J. Biol. Chem. 267, 4806–4814.
32. Orellana, S. A., Amieux, P. S., Zhao, X. Y., and McKnight, G. S. (1993) J. Biol. Chem. 268, 6843–6846.
33. Gold, H. M., Ji-Buchler, Y., and Taylor, S. S. (1997) J. Biol. Chem. 272, 16343–16350.
34. Sahara, S., Sato, K., Kaise, H., Mori, K., Aoto, M., and Tokmakov, A. (1996) FEBS Lett. 384, 138–142.
35. Huang, L. J. S., and Taylor, S. S. (1998) J. Biol. Chem. 273, 26739–26746.
36. Hubbard, S. J., Thornton, J. M. (1993) NACCESS, Version 2.1, Department of Biochemistry and Molecular Biology, University College, London.
37. Richards, F. M. (1977) Annu. Rev. Biophys. Bioeng. 6, 151–176.
38. Street, A. G., and Mayo, S. L. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 9074–9076.
39. Janin, J., Miller, S., and Chothia, C. (1988) J. Mol. Biol. 204, 155–164.