LA-UR-01- 963

Approved for public release; distribution is unlimited.

Title:

STRUCTURAL MODELING OF THE CATALYTIC SUBUNIT-REGUALTORY SUBUNIT DIMERIC COMPLEX OF THE CAMP-DEPENTENT PROTEIN KINASE

Author(s):

Chang-Shung Tung, T-10, LANL
Stephen C. Gallagher, B-Division, LANL
Donal A. Walsh, UC Davis
Jill Trewhella, B-Division, LANL

Submitted:

BIOPHYSICAL SOCIETY MEETING
BOSTON, MA
FEB. 17-22, 2001

Group Address:

Theoretical Biology and Biophysics Group Los Alamos National Laboratory Los Alamos, New Mexico 87545

Los Alamos
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. The Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee it's technical correctness.

Form No. 836 R5
ST 2629 10/96

STRUCTURAL MODELING OF THE CATALYTIC SUBUNIT–REGULATORY SUBUNIT DIMERIC COMPLEX OF THE CAMP–DEPENDENT PROTEIN KINASE.

C.S. Tung(1), S.C. Gallagher(2), D.A. Walsh(3), J. Trewhella(2). (1)Theoretical Division, LANL
(2) Bioscience, LANL, Los Alamos, NM 87545, (3)Dept of Biological Chemistry, UC Davis, CA 95616.

Abstract: The cAMP–dependent protein kinase (PKA) is a multifunctional kinase that serves as a prototype for understanding second messenger signaling and protein phosphorylation. In the absence of a cAMP signal, PKA exists as a dimer of dimers, consisting of two regulatory (R) and two catalytic (C) subunits. Based on experimentally derived data (i.e., crystal structures of the R and C subunits, mutagenesis data identifying points of subunit–subunit contacts), the neutron scattering derived model for the heterodimer (Zhao et al., 1998) and using a set of computational approaches (homology modeling, Monte Carlo simulation), we have developed a high–resolution model of the RIIα–Cα dimer. The nature of the subunit–subunit interface was studied. Our model reveals an averaged size dimer interface (2100 Angstrom^2) that is distant from the pseudo–substrate binding site on the C subunit. The additional contacts made by the pseudosubstrate increases the stability of the dimeric complex. Based on a set of R–C dimer structures derived using a simulated annealing approach, specific interactions (hydrogen bonds) between the two subunits were identified.

Homology modeling the catalytic subunit: The previous neutron scattering experiments that provided the structures of the PKA heterodimer and holoenzyme were done using bovine Cα subunit. As the template for homology modeling of the structure of this C–subunit form we chose the crystal structure (1CDK in the Protein Data Bank, PDB) of the porcine Cα subunit of PKA. The sequence of the target protein has been determined (accession number X67154, GenBank). The sequence identity between the target bovine Cα–subunit and the porcine Cα–subunit template is extremely high with no insertions or deletions and with the two proteins differing by only by a single conservative amino acid change (Fig. 1, upper panel). Modeling the target structure based on the template structure was therefore simple and straightforward. The porcine crystal structure is for the C–subunit PKI–peptide complex with C in its "closed" conformation. Since the binding of the R–subunit pseudosubstrate domain to the C–subunit also results in a closed conformation, this model is the best choice for constructing the R–C heterodimer. The homology modeled structure of the target C–subunit, as well as the structure of the template, are shown in the lower panel of Fig. 1.

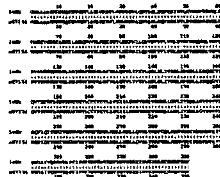


Figure 1. The sequence of the bovine C–subunit (the target) of the cAMP–dependent protein kinase is aligned with that of the porcine C–subunit (the template). The homology modeled structure of the C–subunit (the molecule on the right) and the crystal structure of the C–subunit (the molecule on the left) are shown in the lower part of the figure.

Modeling the regulatory subunit: The crystal structure of the Δ(1–90) bovine RIα subunit of the cAMP–dependent protein kinase (1RGS in PDB), which is the only 3D structure so far resolved for any PKA regulatory subunit, was used as the template. The homologous murine Δ(1–91)RIIα subunit, as used for previous neutron scattering experiments of PKA heterodimer and holoenzyme, was the target sequence. Its sequence has been reported (accession number J02935 in GenBank). This form of RII contains the pseudo–substrate and cAMP– binding domains but not the dimerization domain and thus in combination with C–subunit produces an R–C heterodimer. The target and the template sequences are 43% identical, with an additional 22% conservative residue substitution. Three insertions provide a total of eleven additional residues. The alignment of the template and target sequences is shown in Fig. 2, upper panel. The homology modeled structure of the target molecule, as well the structure of the template molecule, are shown at the lower panel of Fig. 2 (right and left images). The three insertions in the target molecule are shown on the plot in dark gray. The low resolution neutron scattering–resolved structure of the heterodimer, coupled to mutagenesis studies, show that none of the target R–subunit insertions are involved in the binding of the C–subunit. The C–subunit binding surface on the R–subunit is located on the lower back side of the molecule (as depicted in Fig. 2); further conformation of this binding site is provided by this current study. Our past x–ray scattering data have shown that there is no prominent conformational difference between R in the presence and absence of cAMP, indicating that this is a suitable template structure to determine the target RII–subunit structure; upon which to then construct the R–C heterodimer.

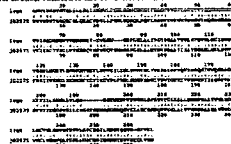


Figure 2. The sequence of the murine regulatory–subunit (the target) of the cAMP–dependent protein kinase is aligned with that of the bovine regulatory subunit (the template). The homology modeled structure of the R–subunit (molecule on the right) and the crystal structure of the R–subunit (molecule on the left) are shown in the lower part of the figure.

Modeling the heterodimer: Modeling the complex structure of the R/C dimer is essentially a docking problem. We have divided the docking procedure into two steps. The first being the course docking of the two subunits into the modeled complex derived from neutron scattering study (Zhao et al., 1998). The course docking involves the matching of centers–of–mass and principal–moment–axes between the two homology modeled subunits and those in the neutron model. Due to rotational symmetries, each of the subunits can be matched in four different configurations giving a total of 16 complex conformations. To reduce the number of possible complex conformations, we use the observed ion pair between Glu–143 of the R–subunit and Lys–213 of the C–subunit (Gibson et al., 1997) as a structural constraint. For each of the subunits, two configurations are having the residue that form the observed ion–pair located at the far side of the complex interface, therefore must be discarded. This leaves 4 different conformations of the R/C complex as the potential solution (see fig. 3).

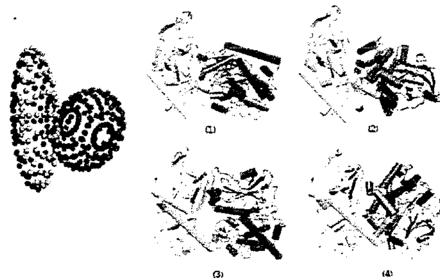


Figure 3. Four different R/C heterodimer structures derived from a coarse docking procedure. The neutron derived model of the heterodimer is shown at the left.

The second step of the docking involves a detailed search in the conformational space around the 4 heterodimer structures derived from the coarse docking procedure. Each of the subunit structure is sampled at a conformational grid associated with the six degrees of freedom (2 Angstrom grid for translational, 10 degree grid for rotational). Taking the size and shape into consideration, we will sample 3,600 configurations for the R subunit and 6,000 configurations for the C subunit, giving a total of almost 22 million conformations to be sampled for each of the four initial heterodimer structures. To speed up the calculation, we will use a reduced–coordinate representation that includes an extended atom (at C–alpha) per residue. Each of the subunits is treated as a rigid body in the docking procedure. In addition to the ion pair, six residues (Asp–141(R), Lys–247(R), Thr–195(C), Trp–196(C), Thr–197(C), Lys–217(C)) were identified to be important to the complex formation of the dimer, therefore should be physically close to the corresponding subunit. Using van der Waals exclusion (no pairwise C–alpha distance from the two subunits should be less than 3 Angstrom) and close contact rule (the minimum distance (d(1–6)) between the close–contact residues and the corresponding subunits should be less than 12 Angstrom while the distance between the two residues that form the ion–pair (d(0)) should be less than 10 Angstrom) as constraints, the initial screening shows that the grid sampling produces no satisfying dimer structure from initial structures 1 and 2 (see Fig. 3). The same screening produces 2642 and 10631 dimer conformations from initial structures 3 and 4 respectively.

To further delineate these structures, we define three empirical parameters (Pvdw, Pc1, Pc2). Pvdw is a measure of the 6–12 vdw energy with parameters corresponding to vdw radii of 4.0 Angstrom and a depth of 0.12. Pc1 is a constraint associated with the residues in close contact upon the complex formation. Pc1 can be calculated according to:

$$Pc1 = \sum_i Q_i ,$$

where

$$Q_i = C_i * (d(i) - d0)^2 \quad \text{if } d(i) > 8.0,$$

$$= 0, \quad \text{if } d(i) <= 8.0 .$$

C_i is a constant arbitrary chosen to be 100 Kcal/mol/Angstrom^2 for the ion pair and 10 Kcal/mol/Angstrom^2 for the six residues while d0 is 8 Angstrom. Pc2 is a parameter to describe the goodness of fit between the structure and the neutron model. Pc2 is defined to be the number of residues in the structure that are external to the neutron model.

A second screening is performed with the limits of the three parameters set at 1,000, 200., and 200. respectively. Only 7 dimer structures around the initial structure 4 are selected from the second screening. Out of these 7 dimer structures, the one has the best overall fit is chosen to be the structure for the complex and shown in Fig. 4a. The atomic model of the heterodimer is constructed based on this chosen structure. The atomic model is subjected to energy minimization using AMBER. The energy minimized structure is shown in Fig. 4b.

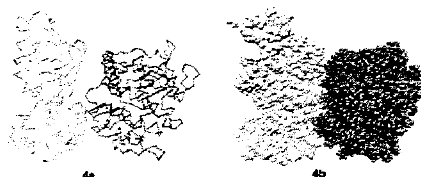


Figure 4. Structure of the R/C heterodimer derived using a grid–sampling method. Structure of the dimer with only the C–alpha atoms is shown in 4a and the energy–minimized all–atom structure is shown in 4b.

Summary: SASMODEL (Zhao et al., 1998) was used with the the x–ray scattering data of the R/C dimeric complex to generate a refined model structure (the dotted structure shown in Fig. 5) that reveals more information with regard to the shape of the complex. The energy minimized structure (the stick model) when docked onto this refined scattering–derived model shows a good agreement as depicted in Fig. 5.

Figure 5. The energy–minimized structure of the R/C complex (the stick structure) is superimposed with the refined scattering–derived model (the dotted structure).

To obtain the detailed R–C interaction, the structure of the heterodimer was allowed to relax at room temperature (300 K). This procedure is accomplished by subjecting the complex to a 100 pico second run of molecular dynamics using AMBER. During the simulation, C–alpha atoms on both subunits, except those located on the proximity of the interface were constraint to their original positions. This procedure allows for those residues at the interface to relax and sample more of the conformational space while maintaining relative orientation of the two subunits. Structures at every 10 pico second during the simulation were selected and energy–minimized. The resulting 10 energy–minimized structures are shown in Fig. 6.

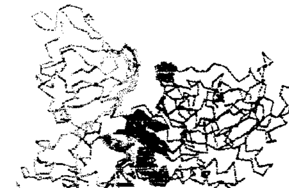


Figure 6. Ten structures of the dimeric complex derived using an annealing procedure involving MD simulation and energy–minimization.

Based on these energy–minimized structures, hydrogen–bonds (H–bond) between the two subunits of the R–C complex were identified. Those H–bonds between the subunits existing in the simulation at least 80% of the time are tabulated in Table 1. The boxed residues are those identified to be important in the formation of the dimeric complex. Those residues involved in the H–bond using the main–chain carbonyl oxygen are indicated with (O) in the table. Lys–23 of the C–subunit forms a H–bond alternating between Glu–299 and Ser–300 of the R–subunit while Arg–194 of the C–subunit forms a H–bond alternating between Asn–139 and Leu–140 (O) of the R–subunt.

One of the methods for evaluating binding strength between two subunits in a complex is to calculate the contact surface area. In general, the larger the contact surface area, the stronger the binding. The contact surface area is defined to be the difference between the total surface areas of the two individual subunits and the surface area of the complex. The surface areas were calculated usning NACCESS (Hubbard and Thornton, 1993). The interface surface area of the R/C complex is 2010 square Angstrom, a value compatible to the surface area of a typical protein complex (Conte et al., 1999). When the R–subunit is divided into two structurally distinct domains (A: residues 114–250; B: residues 251–383), the interface surface areas between the domain and the C–subunit are 1584 and 423 square Angstrom respectively for A and B. This result of A–domain being the C–subunit high affinity binding domain is consistent with the findings based on a binding study by Huang and Taylor (1998). In their study, residues 94–169 and 236–244 were identified as primary and secondary sufaces for C–subunit binding. Our model structure of the complex (Fig. 7) shows both residues 114–169 (depicted in yellow) and residues 236–244 are interacting with the C–subunit (shown in cyan).

H–bonds between R/C subunits

R	C	
E-143	K213	100%
R-244(O)	K-217	100%
K-247	[illegible]	100%
D-141	[illegible]	100%
N-139 L-140(O)	R-194	100%
E-299 S-300	K-23	100%
D-301	[illegible]	90%
[illegible]	[illegible]	90%
[illegible]	[illegible]	80%
[illegible]	[illegible]	80%

Figure 7. The interaction of A and B domain of the R–subunit (molecule on the right) with the C–subunit (molecule on the left).

STRUCTURAL MODELING OF THE CATALYTIC SUBUNIT–REGULATORY SUBUNIT DIMERIC COMPLEX OF THE CAMP–DEPENDENT PROTEIN KINASE.

C.S. Tung(1), S.C. Gallagher(2), D.A. Walsh(3), J. Trewhella(2). (1)Theoretical Division, LANL
(2) Bioscience, LANL, Los Alamos, NM 87545, (3)Dept of Biological Chemistry, UC Davis, CA 95616.

Abstract: The cAMP−dependent protein kinase (PKA) is a multifunctional kinase that serves as a prototype for under−standing second messenger signaling and protein phosphorylation. In the absence of a cAMP signal, PKA exists as a dimer of dimers, consisting of two regulatory (R) and two catalytic (C) subunits. Based on experimentally derived data (i.e., crystal structures of the R and C subunits, mutagenesis data identifying points of subunit−subunit contacts), the neutron scattering derived model for the heterodimer (Zhao et al., 1998) and using a set of computational approaches (homology modeling, Monte Carlo simulation), we have developed a high−resolution model of the RIIα−Cα dimer. The nature of the subunit−subunit interface was studied. Our model reveals an averaged size dimer interface (2100 Angstrom^2) that is distant from the pseudo−substrate binding site on the C subunit. The additional contacts made by the pseudosubstrate increases the stability of the dimeric complex. Based on a set of R−C dimer structures derived using a simulated annealing approach, specific interactions (hydrogen bonds) between the two subunits were identified.

Homology modeling the catalytic subunit. The previous neutron scattering experiments that provided the structures of the PKA heterodimer and holoenzyme were done using bovine Cα subunit. As the template for homology modeling of the structure of this C–subunit form we chose the crystal structure (1CDK in the Protein Data Bank, PDB) of the porcine Cα subunit of PKA. The sequence of the target protein has been determined (accession number X67154, GenBank). The sequence identity between the target bovine Cα–subunit and the porcine Cα–subunit template is extremely high with no insertions or deletions and with the two proteins differing by only by a single conservative amino acid change (Fig. 1, upper panel). Modeling the target structure based on the template structure was therefore simple and straightforward. The porcine crystal structure is for the C–subunit PKI–peptide complex with C in its "closed" conformation. Since the binding of the R–subunit pseudosubstrate domain to the C–subunit also results in a closed conformation, this model is the best choice for constructing the R–C heterodimer. The homology modeled structure of the target C–subunit, as well as the structure of the template, are shown in the lower panel of Fig. 1.

```
               10        20        30        40        50        60
1cdk   GNAAAAKKGSEQESVKEFLAKAKEDFLKKWENPAQNTAHLDQFERIKTLGTGSFGRVMLV
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
x67154 GNAAAAKKGSEQESVKEFLAKAKEDFLKKWENPAQNTAHLDQFERIKTLGTGSFGRVMLV
               10        20        30        40        50        60

               70        80        90       100       110       120
1cdk   KHKETGNHFAMKILDKQKVVKLKQIEHTLNEKRILQAVNFPFLVKLEYSFKDNSNLYMVM
       :: :::::.::::::::::::::::::::::::::::::::::::::.::::::::::::
x67154 KHMETGNHYAMKILDKQKVVKLKQIEHTLNEKRILQAVNFPFLVKLEFSFKDNSNLYMVM
               70        80        90       100       110       120

              130       140       150       160       170       180
1cdk   EYVPGGEMFSHLRRIGRFSEPHARFYAAQIVLTFEYLHSLDLIYRDLKPENLLIDQQGYI
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
x67154 EYVPGGEMFSHLRRIGRFSEPHARFYAAQIVLTFEYLHSLDLIYRDLKPENLLIDQQGYI
              130       140       150       160       170       180

              190       200       210       220       230       240
1cdk   QVTDFGFAKRVKGRTWTLCGTPEYLAPEIILSKGYNKAVDWWALGVLIYEMAAGYPPFFA
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
x67154 QVTDFGFAKRVKGRTWTLCGTPEYLAPEIILSKGYNKAVDWWALGVLIYEMAAGYPPFFA
              190       200       210       220       230       240

              250       260       270       280       290       300
1cdk   DQPIQIYEKIVSGKVRFPSHFSSDLKDLLRNLLQVDLTKRFGNLKDGVNDIKNHKWFATT
       ::::::::::::::::::::::::::::::::::::::::::::.:::::::::::::::
x67154 DQPIQIYEKIVSGKVRFPSHFSSDLKDLLRNLLQVDLTKRFGNLKNGVNDIKNHKWFATT
              250       260       270       280       290       300

              310       320       330       340       350
1cdk   DWIAIYQRKVEAPFIPKFKGPGDTSNFDDYEEEEIRVSINEKCGKEFSEF
       ::::::::::::::::::::::::::::::::::::::::::::::::::
x67154 DWIAIYQRKVEAPFIPKFKGPGDTSNFDDYEEEEIRVSINEKCGKEFSEF
              310       320       330       340       350
```

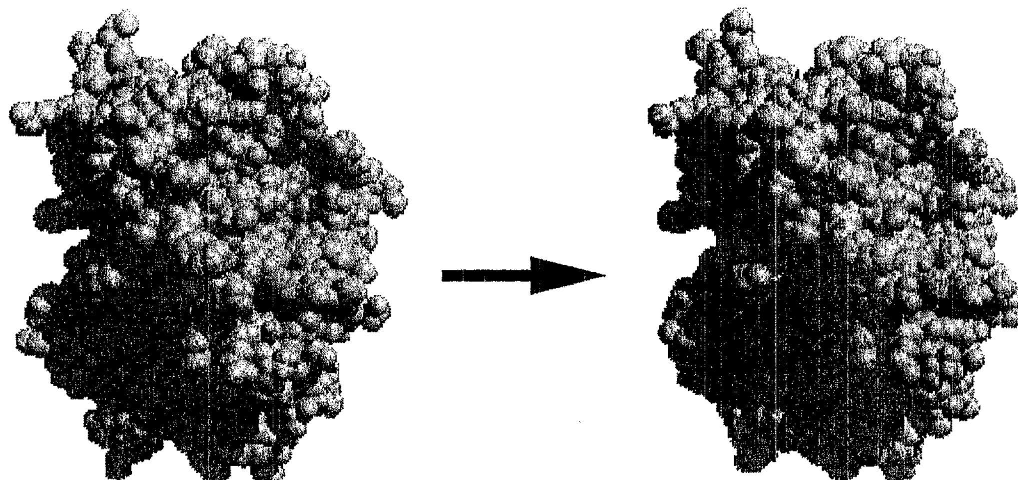


Figure 1. The sequence of the bovine C–subunit (the target) of the cAMP–dependent protein kinase is aligned with that of the porcine C–subunit (the template). The homology modeled structure of the C–subunit (the molecule on the right) and the crystal structure of the C–subunit (the molecule on the left) are shown in the lower part of the figure.

RIα subunit of the cAMP–dependent protein kinase (1RGS in PDB), which is the only 3D structure so far resolved for any PKA regulatory subunit, was used as the template. The homologous murine Δ(1–91)RIIa subunit, as used for previous neutron scattering experiments of PKA heterodimer and holoenzyme, was the target sequence. Its sequence has been reported (accession number J02935 in GenBank). This form of RII contains the pseudo–substrate and cAMP– binding domains but not the dimerization domain and thus in combination with C–subunit produces an R–C heterodimer. The target and the template sequences are 43% identical with an additional 22% conservative residue substitution. Three insertions provide a total of eleven additional residues. The alignment of the template and target sequences is shown in Fig. 2, upper panel. The homology modeled structure of the target molecule, as well the structure of the template molecule, are shown at the lower panel of Fig. 2 (right and left images). The three insertions in the target molecule are shown on the plot in dark gray. The low resolution neutron scattering–resolved structure of the heterodimer, coupled to mutagenesis studies, show that none of the target R–subunit insertions are involved in the binding of the C–subunit. The C–subunit binding surface on the R–subunit is located on the lower back side of the molecule (as depicted in Fig. 2); further conformation of this binding site is provided by this current study. Our past x–ray scattering data have shown that there is no prominent conformational difference between R in the presence and absence of cAMP, indicating that this is a suitable template structure to determine the target RII–subunit structure; upon which to then construct the R–C heterodimer.

```
               10        20        30        40        50        60
1rgs   RKVIPKDYKTMAALAKAIEKNVLFSHLDDNERSDIFDAMFPVSFIAGETVIQQGDEGDNF
       : : ::  .    : .: .  .::..::... :...::::     . : ::.:::.::::
j02935 RVVHPKTDEQRCRLQEACKDILLFKNLDQEQLSQVLDAMFEKIVKTDEHVIDQGDDGDNF
               10        20        30        40        50        60

               70         80           90       100       110
1rgs   YVIDQGEMDVYVNNEWAT-SVGEG---GSFGELALIYGTPRAATVKAKTNVKLWGIDRDS
       :::..: .:. :...  : :::.    ::::::::.:.::::::. : .. .:::.:: .
j02935 YVIERGTYDILVTKDNQTRSVGQYDNRGSFGELALMYNTPRAATIIATSEGSLWGLDRVT
               70        80        90       100       110       120

        120       130       140       150       160       170
1rgs   YRRILMGSTLRKRKMYEEFLSKVSILESLDKWERLTVADALEPVQFEDGQKIVVQGEPGD
       .:::.. .. .::::.: :. .: ...::.  ::. ..:..    ..::..:..::: .:
j02935 FRRIIVKNNAKKRKMFESFIESVPLFKSLEMSERMKIVDVIGEKIYKDGERIIAQGEKAD
              130       140       150       160       170       180

        180       190             200       210       220       230
1rgs   EFFIILEGSAAVLQR------RSENEEFVEVGRLGPSDYFGEIALLMNRPRAATVVARGP
        :.::  : ...: :      ..  ...  ::...   ..::::.::. :.::::.. . :
j02935 SFYIIESGEVSILIRSKTKSNKNGGNQEVEIAHCHKGQYFGELALVTNKPRAASAYGVGD
              190       200       210       220       230       240

              240       250       260
1rgs   LKCVKLDRPRFERVLGPCSDILKRNIQQYN-SFVSL
       .::. .:   :::.:::: ::.::::..:. ..:..
j02935 VKCLVMDVQAFERLLGPCMDIMKRNISHYEEQLVKM
              250       260       270
```

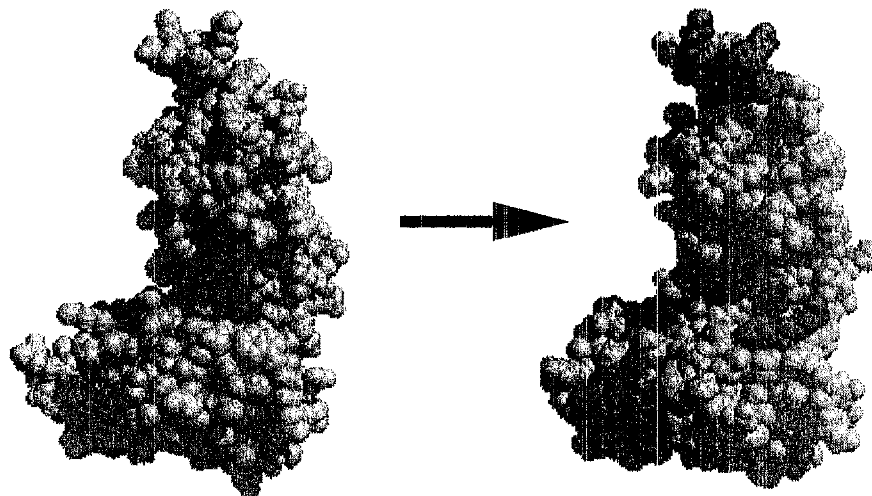


Figure 2. The sequence of the murine regulatory–subunit (the target) of the cAMP–dependent protein kinase is aligned with that of the bovine regulatory subunit (the template). The homology modeled structure of the R–subunit (molecule on the right) and the crystal structure of the R–subunit (molecule on the left) are shown in the lower part of the figure.

Modeling the heterodimer: Modeling the complex structure of the R/C dimer is essentially a docking problem. We have divided the docking procedure into two steps. The first being the coarse docking of the two subunits into the modeled complex derived from neutron scattering study (Zhao et al., 1998). The coarse docking involves the matching of centers–of–mass and principal–moment–axes between the two homology modeled subunits and those in the neutron model. Due to rotational symmetries, each of the subunits can be matched in four different configurations giving a total of 16 complex conformations. To reduce the number of possible complex conformations, we use the observed ion pair between Glu–143 of the R–subunit and Lys–213 of the C–subunit (Gibson et al., 1997) as a structural constraint. For each of the subunits, two configurations are having the residue that form the observed ion–pair located at the far side of the complex interaface, therefore must be discarded. This leaves 4 different conformations of the R/C complex as the potential solution (see fig. 3).

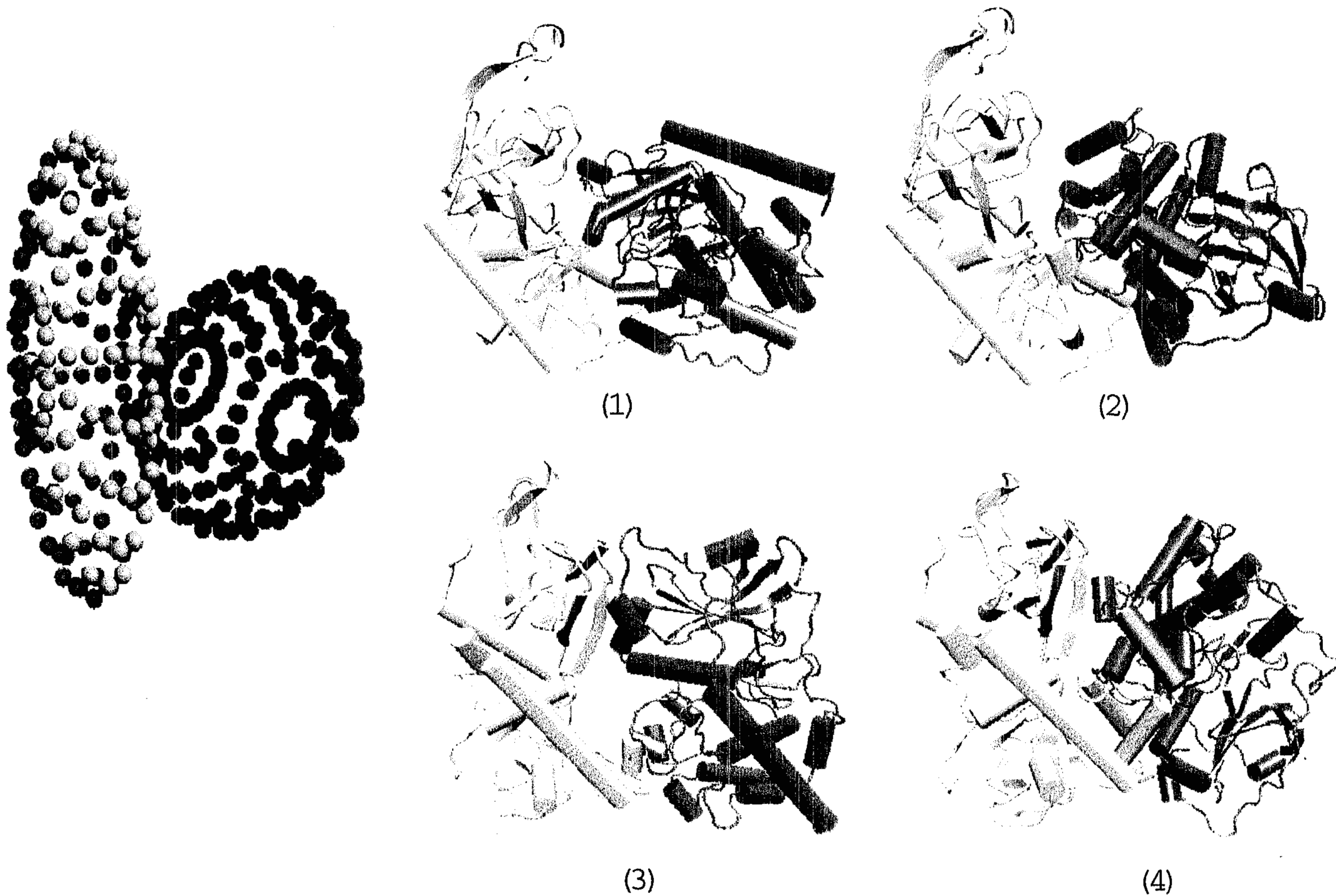


Figure 3. Four different R/C heterodimer structures derived from a coarse docking procedure. The neutron derived model of the heterodimer is shown at the left.

The second step of the docking involves a detailed search in the conformational space around the 4 heterodimer structures derived from the coarse docking procedure. Each of the subunit structure is sampled at a conformational grid associated with the six degrees of freedom (2 Angsrom grid for translational, 10 degree grid for rotational). Taking the size and shape into consideration, we will sample 3,600 configurations for the R subunit and 6,000 configurations for the C subunit, giving a total of almost 22 million conformations to be sampled for each of the four initial heterodimer structures. To speed up the calculation, we will use a reduced-coordinate representation that includes an extended atom (at C-alpha) per residue. Each of the subunits is treated as a rigid body in the docking procedure. In addition to the ion pair, six residues (Asp-141(R), Lys-247(R), Thr-195(C), Trp-196(C), Thr-197(C), Lys-217(C)) were identified to be important to the complex formation of the dimer, therefore should be physically close to the corresponding subunit. Using van der Waals exclusion (no pairwiase C-alpha distance from the two subunits should be less than 3 Angstrom) and close contact rule (the minimum distance (d(1-6)) between the close-contact residues and the corresponding subunits should be less than 12 Angstrom while the distance between the two residues that form the ion-pair (d(0)) should be less than 10 Angstrom) as constraints, the initial screening shows that the grid sampling produces no satisfying dimer structure from initial structures 1 and 2 (see Fig. 3). The same screening produces 2642 and 10631 dimer conformations from initial structures 3 and 4 respectively.

To further delineate these structures, we define three empirical parameters (Pvdw, Pc1, Pc2). Pvdw is a measure of the 6-12 vdw energy with parameters corresponding to vdw radii of 4.0 Angstrom and a depth of 0.12. Pc1 is a constraint associated with the residues in close contact upon the complex formation. Pc1 can be calculated according to:

$$Pc1 = \sum_i Q_i ,$$

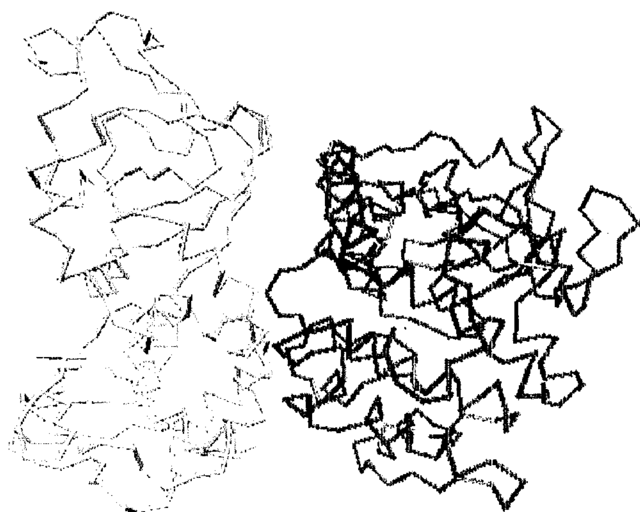
where

$$Q_i = C_i * (d(i) - d0)^2 \qquad \text{if } d(i) > 8.0,$$

$$= 0. \qquad \text{if } d(i) <= 8.0 ,$$

C_i is a constant arbitrary chosen to be 100 Kcal/molAngstrom^2 for the ion pair and 10 Kcal/molAngsrom^2 for the six residues while d0 is 8 Angstrom. Pc2 is a parameter to describe the goodness of fit between the structure and the neutron model. Pc2 is defined to be the number of residues in the structure that are external to the neutron model.

A second screening is performed with the limits of the three parameters set at 1,000, 200., and 200. respectively. Only 7 dimer structures around the initial structure 4 are selected from the second screening. Out of these 7 dimer structures, the one has the best overall fit is chosen to be the structure for the complex and shown in Fig. 4a. The atomic model of the heterodimer is constructed based on this chosen structure. The atomic model is subjected to energy minimization using AMBER. The energy minimized structure is shown in Fig. 4b.



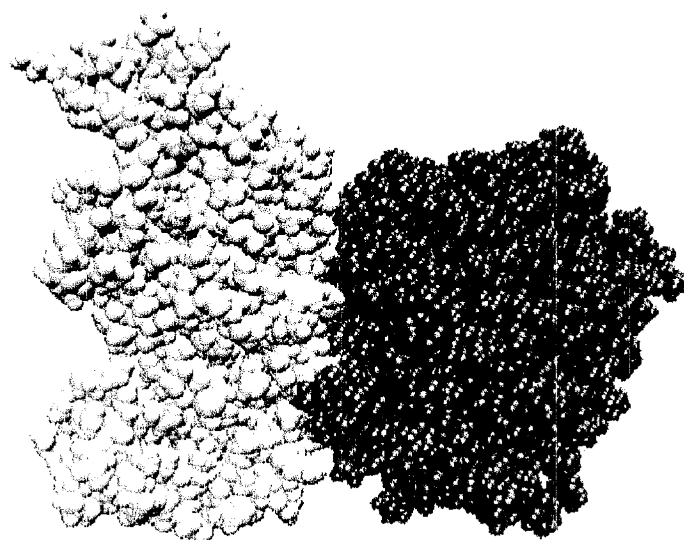
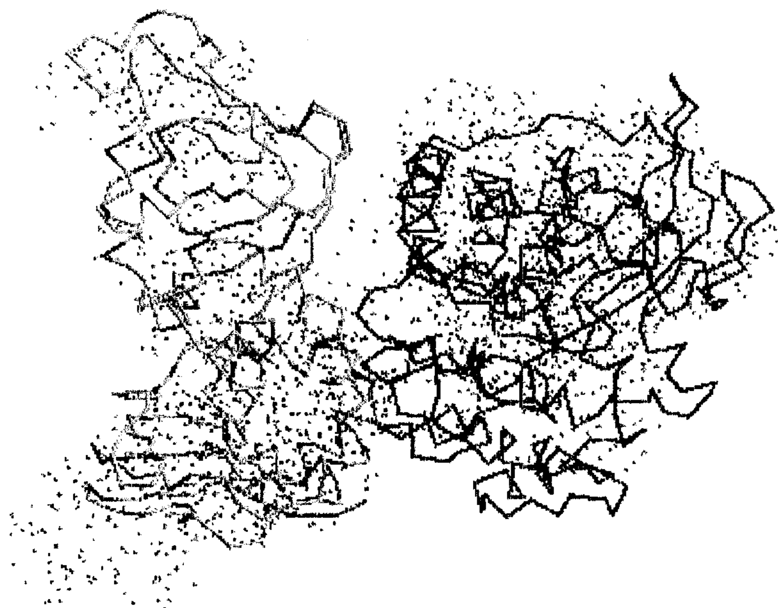


Figure 4. Structure of the R/C heterodimer derived using a grid–sampling method. Structure of the dimer with only the C–alpha atoms is shown in 4a and the energy–minimized all–atom structure is shown in 4b.

Summary: SASMODEL (Zhao et al., 1998) was used with the the x–ray scattering data of the R/C dimeric complex to generate a refined model structure (the dotted structure shown in Fig. 5) that reveals more information with regard to the shape of the complex. The energy minimized structure (the stick model) when docked onto this refined scattering–derived model shows a good agreement as depicted in Fig. 5.



Fogure 5. The energy–minimized structure of the R/C complex (the stick structure) is superimposed with the refiened scattering–derived model (the dotted structure).

To obtain the detailed R–C interaction, the structure of the heterodimer was allowed to relax at room temperature (300 K). This procedure is accomplished by subjecting the complex to a 100 pico second run of molecular dynamics using AMBER. During the simulation, C–alpha atoms on both subunits, except those located on the proximity of the interface were constraint to their original positions. This procedure allows for those residues at the interface to relax and sample more of the conformational space while maintaining relative orientation of the two subunits. Structures at every 10 pico second during the simulation were selected and energy–minimized. The resulting 10 energy–minimized structures are shown in Fig. 6.

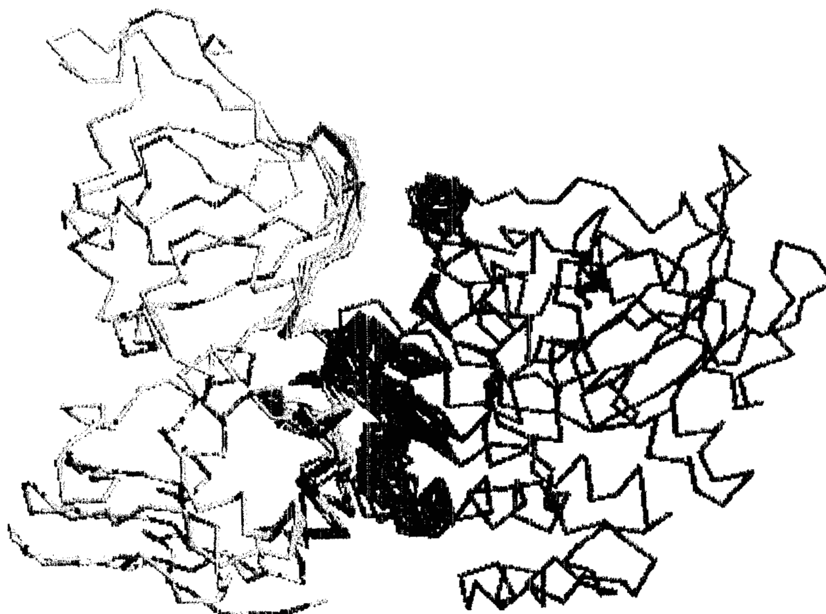


Figure 6. Ten structures of the dimeric complex derived using an annealing procedure involving MD simulation and energy–minimization.

Based on these energy–minimized structures, hydrogen–bonds (H–bond) between the two subunits of the R–C complex were identified. Those H–bonds between the subunits existing in the simulation at least 80% of the time are tabulated in Table I. The boxed residues are those identified to be important in the formation of the dimeric complex. Those residues involved in the H–bond using the main–chain carbonyl oxygen are indicated with (O) in the table. Lys–23 of the C–subunit forms a H–bond alternating between Glu–299 and Ser–300 of the R–subunit while Arg–194 of the C–subunit forms a H–bond alternating between Asn–139 and Leu–140 (O) of the R–subunt.

One of the methods for evaluating binding strength between two subunits in a complex is to calculate the contact surface area. In general, the larger the contact surface area, the stronger the binding. The contact surface area is defined to be the difference between the total surface areas of the two individual subunits and the surface area of the complex. The surface areas were calculated usning NACCESS (Hubbard and Thornton, 1993). The interface surface area of the R/C complex is 2010 square Angstrom, a value compatible to the surface area of a typical protein complex (Conte et al., 1999). When the R–subunit is divided into two structurally distinct domains (A: residues 114–250; B: residues 251–383), the interface surface areas between the domain and the C–subunit are 1584 and 423 square Angstrom respectively for A and B. This result of A–domain being the C–subunit high affinity binding domain is consistent with the findings based on a binding study by Huang and Taylor (1998). In their study, residues 94–169 and 236–244 were identified as primary and secondary sufaces for C–subunit binding. Our model structure of the complex (Fig. 7) shows both residues 114–169 (depicted in yellow) and residues 236–244 are interacting with the C–subunit (shown in cyan).

H-bonds between R/C subunits

R	C	
E-143	K213	100%
K-244(O)	K-217	100%
K-247	N-283	100%
D-141	S-212	100%
N-139 L-140(O)	R-194	100%
E-299 S-300	K-23	100%
D-281	K-28	90%
K-329	E-24	90%
K-244	E-208	80%
E-276	K-192	80%

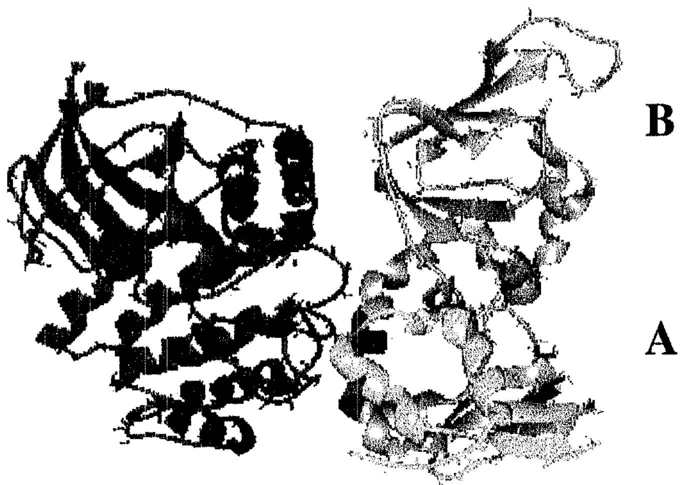


Figure 7. The interaction of A and B domain of the R–subunit (molecule on the right) with the C–subunit (molecule on the left).

Based on these energy--minimized structures, hydrogen–bonds (H–bond) between the two subunits of the R–C complex were identified. Those H–bonds between the subunits existing in the simulation at least 80% of the time are tabulated in Table I. The boxed residues are those identified to be important in the formation of the dimeric complex. Those residues involved in the H--bond using the main–chain carbonyl oxygen are indicated with (O) in the table. Lys–23 of the C–subunit forms a H–bond alternating between Glu–299 and Ser–300 of the R–subunit while Arg–194 of the C–subunit forms a H–bond alternating between Asn--139 and Leu–140 (O) of the R–subunt.

One of the methods for evaluating binding strength between two subunits in a complex is to calculate the contact surface area. In general, the larger the contact surface area, the stronger the binding. The contact surface area is defined to be the difference between the total surface areas of the two individual subunits and the surface area of the complex. The surface areas were calculated usning NACCESS (Hubbard and Thornton, 1993). The interface surface area of the R/C complex is 2010 square Angstrom, a value compatible to the surface area of a typical protein complex (Conte et al., 1999). When the R–subunit is divided into two structurally distinct domains (A: residues 114–250; B: residues 251–383), the interface surface areas between the domain and the C–subunit are 1584 and 423 square Angstrom respectively for A and B. This result of A–domain being the C–subunit high affinity binding domain is consistent with the findings based on a binding study by Huang and Taylor (1998). In their study, residues 94–169 and 236–244 were identified as primary and secondary sufaces for C–subunit binding. Our model structure of the complex (Fig. 7) shows both residues 114–169 (shown in yellow) and residues 236–244 (shown in magenta) are interacting with the C–subunit (shown in cyan).

H-bonds between R/C subunits

R	C	
E-143	K213	100%
K-244(O)	K-217	100%
K-247	N-283	100%
D-141	S-212	100%
N-139 L-140(O)	R-194	100%
E-299 S-300	K-23	100%
D-281	K-28	90%
K-329	E-24	90%
K-244	E-208	80%
E-276	K-192	80%

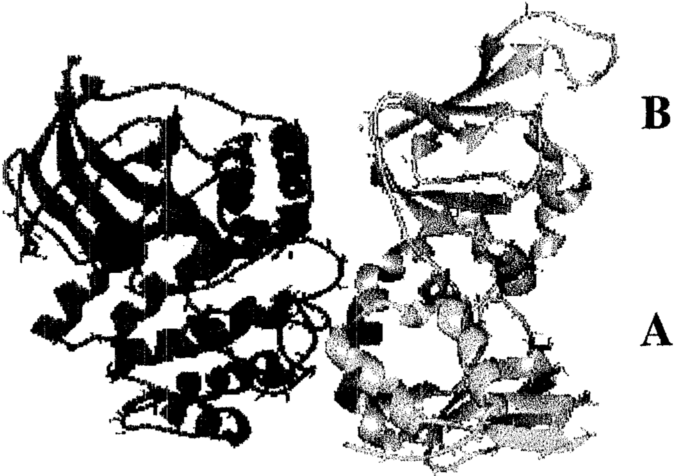


Figure 7. The interaction of A and B domain of the R–subunit (molecule on the right) with the C–subunit (molecule on the left).