ELSEVIER

Journal of Molecular Graphics and Modelling 21 (2002) 3–18

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

The three-dimensional model of *Dictyostelium discoideum* racE based on the human rhoA-GDP crystal structure

Madhavi Agarwal [a,*], Donald J. Nelson [b], Denis A. Larochelle [a]

[a] *Department of Biology, Clark University, 950 Main Street, Worcester, MA 01610, USA*
[b] *Department of Chemistry, Clark University, 950 Main Street, Worcester, MA 01610, USA*

Abstract

The three-dimensional structure of racE was modeled using several homologous small G proteins, and the best model obtained using the human rhoA as modeling template is reported. The three-dimensional fold of the racE model is remarkably similar to the cellular form of human ras p21 crystal structure. Its secondary structure consists of six α-helices, six β-strands and three 3_{10} helices. The model retains its secondary structure after a 300 K, 300 ps molecular dynamics (MD) simulation. Important domains of the protein include its effector loop (residues 34–46), the insertion domain (residues 121–136), and the polybasic motif (between 210 and 220) not modeled in the current structure. The effector loop is inherently flexible and the structure docked with GDP exhibits the effector loop moving significantly closer to the nucleotide binding pocket, forming a tighter complex with the bound GDP. The mobility of the effector loop is conferred by a single residue 'hinge' point at residue ^{34}Asp, also allowing the Switch I region, immediately preceding the effector loop, to be equally mobile. In comparison, the Switch II region shows average mobility. The insertion domain is highly flexible, with the insertion taking the form of a helical domain, with several charged residues forming a complex charged interface over the entire insertion region. While the GDP moiety is loosely held in the active site, the metal cation is extensively co-ordinated. The critical residue ^{38}Thr exhibits high mobility, and is seen interacting directly with the metal ion at a distance of 2.64 Å, and indirectly via an intervening water molecule. ^{64}Gln, a key residue involved in GTP hydrolysis in ras, is seen facing the β-phosphate group and the metal ion. Certain residues (i.e. ^{51}Asn, ^{38}Thr and ^{65}Glu) exhibit unique characteristics and these residues, together with ^{158}Val, may play important roles in the maintenance of the protein's integrity and function. There is strong consensus of secondary structural elements between models generated using various templates, such as h-rac1, h-rhoA and h-cdc42 bound to RhoGDI, all sharing only 50–55% sequence identity with racE, which suggests that this model is in all probability an accurate prediction of the true tertiary structure of racE.

Keywords: Modeling; Molecular dynamics; racE; rhoA

1. Introduction

The crystal structure for ras and several other small GTP-binding proteins has been solved [1–5]. These proteins function as molecular switches in cells, turning on through the binding of GTP, and turning off through the hydrolysis of the bound GTP to GDP [6–9]. The small GTPases rho, rac, and CDC42, members of the rho family of ras-related GTP-binding proteins, control various cell morphological processes such as formation of stress fibers, focal adhesions, membrane ruffling, and filopodia formation by modulating the actin cytoskeleton [10]. Although closely related, these small GTPases display uniqueness in terms of their regulation of the actin cytoskeleton. For example, activation of rho in Swiss 3T3 cells leads to the formation of stress fibers whereas activation of CDC42 leads to the formation of filopodia [11]. This uniqueness of function must come about, in part, through the binding of specific effector proteins, both upstream and downstream in their respective pathways. The binding to specific effector proteins is likely to be mediated through differences in single amino acid residues or short stretches of amino acids between these various small GTPases.

Members of the rho family of ras-related GTP-binding proteins have also been implicated in the regulation of cytokinesis [12,13]. RacE, a unique member of this family, was found to be essential for cytokinesis in *Dictyostelium discoideum* [14]. Cells lacking racE protein are unable to undergo cytokinesis when grown in suspension culture and become large and multi-nucleate (mitosis continues unabated). On the other hand, these cells divide and are viable in stationary culture, dividing by an alternative mechanism referred to as traction-mediated cytofission [15]. A mutation

* Corresponding author. Tel.: +1-508-793-7664; fax: +1-508-793-8861.
E-mail address: magarwal@clarku.edu (M. Agarwal).

1093-3263/02/$ – see front matter
PII: S1093-3263(01)00137-1

in the conserved P-loop of racE (i.e. V20racE, with Val substituted for ^{20}Gly) results in reduced GTPase activity and is constitutively active, whereas the N25racE (Asn substituted for ^{25}Thr) mutation exhibits a decreased ability to exchange GTP for bound GDP and is therefore, constitutively inactive [16]. In expressing these mutant racE genes in racE null cells, it was found that only the constitutively active form, V20racE, was able to rescue the cytokinesis defect in racE null cells. This suggests that active racE is essential for proper cytokinesis but its conversion to its inactive form is not. Using a GFP-racE construct it was also determined that racE localizes uniformly to the cell cortex and is present throughout the cell cycle. Taken together, these results suggested that racE is not required for determining either the timing or the placement of the contractile ring during cytokinesis [16]. Additional studies have shown that racE is involved in maintaining proper cortical tension in cells and that this cortical tension is essential for cytokinesis to take place when cells are grown in suspension culture. RacE null cells grown in these conditions form normal contractile rings but are unable to complete cytokinesis and the furrow regresses [17,18].

There are several closely related rac proteins in *Dictyostelium* [19], yet none can compensate for the cytokinesis defect observed in racE null cells. We would like to know what unique domains are present in racE that is essential to its function in cytokinesis. This can be assessed, in part, by generating various chimeras between racE and a closely related rac protein (racC in this case) that does not share racE's function in cytokinesis. This approach has been used to address the importance of the unique racE C-terminal domain [16]. The substitution of this domain for the comparable region on racC was not sufficient to convert racC into the function equivalent of racE. Therefore, although this domain may be important, there must be additional domains that are important, and presumably indispensable, in racE's function in cytokinesis. Consequently, we were interested in transiting from a linear amino acid sequence to a three-dimensional structure that would offer us clues with regard to functional specificity. Studying such a racE structure, powered with the current knowledge of different mutations and chimeric proteins, and in support with resolved structures of other small GTPases would allow us to identify certain domains and possibly, certain residues that may play a role in racE's function in cytokinesis.

In this work, we report the three-dimensional model of racE based on the human rhoA-GDP crystal structure, used as the model-building template. We have subjected the model to a 300 K, 300 ps molecular dynamics simulation run to assess the stability of the generated structure. The nascent chain model has been manually docked with GDP·Mg^{2+} and analyzed. Various domains of the racE protein model are studied in detail, revealing certain key domains and residues that must play a significant role in cytokinesis or other functions of racE.

2. Materials and methods

2.1. Initial three-dimensional structure prediction

The starting input for all of the computational studies was the amino acid sequence (one letter code) of the racE protein [14] (Swiss Prot: Q23862). A three-dimensional modeling request was submitted to the Swiss Modeling Server [25–29] on the World Wide Web (http://www.expasy.ch/swiss-model/). The First Approach Mode was chosen and the template was specified as the human rhoA crystal structure (PDB code: 1FTN.pdb, Swiss Prot code: P06749). PROCHECK [52] was used to evaluate the quality of the racE three-dimensional structure returned from the Swiss model site.

2.2. Molecular dynamics simulation of the Swiss model racE structure

Molecular dynamics simulations were performed on an Indy R5, 000 Silicon Graphics workstation, using the Insight II (v. 97.2) software package from Molecular Simulation Incorporated (MSI) (San Diego, CA). The MSI program *fdiscover* was used for the molecular dynamics calculations, employing the second generation cff91 force field. The coordinates associated with the Swiss model racE predicted structure were imported into the Insight II window. The Insight II Biopolymer module was then used to add hydrogen atoms to the appropriate heavy atoms of racE, with the pH value set to 7.0, and to calculate partial charges on all atoms. In order the mimic aqueous solvent conditions, the modified racE molecule was soaked with a 10 Å shell of 2308 water molecules. After generating the "assembly" (i.e. the racE protein plus water shell), the water shell was subjected to energy minimization using steepest descents to a maximum derivative less than 1.0 kcal/(mol Å), followed by the conjugate gradient method until the maximum derivative was less than 0.1 kcal/(mol Å). Subsequently, the remainder of the complex (i.e. the racE protein and the shell of water molecules) was energy minimized using steepest descents and conjugate gradient methods until the maximum derivative for the entire assembly was less than 0.1 kcal/(mol Å).

The kinetic energy for the molecular dynamics simulation was provided by a thermal bath at a constant temperature of 300 K for 10,000 steps of 1.0 fs each. Non-bonded interactions were evaluated with a cut-off distance of 12.0 Å and a switch distance of 2.0 Å. Following thermal equilibration, the simulation was continued for ∼300 ps with the output of binary data occurring every 1 ps during the trajectory and stored as a history file. The program also was set to record graphic dynamic structure "snapshots" every 20 ps starting after the 10 ps thermal equilibration period. These dynamic structures were output to an archive file for future analysis by Insight II. The 300 binary structures stored in the history file were subsequently examined using the Analysis Module of Insight II.

2.3. Docking of GDP into the Swiss model racE structure

The molecular mechanics associated with the "docking" of GDP-Mg^{2+} into the Swiss model racE structure were also performed on an Indy R5000 Silicon Graphics workstation, using the MSI Insight II (v. 97.2) software package; however, the computations were performed within the *discover*3 module, again using the cff91 force field. As was done for the MD simulation, the coordinates associated with the Swiss model racE-predicted structure were imported into the *biopolymer* module of Insight II, hydrogen atoms were added where appropriate, and charges calculated for all atoms (racE protein plus GDP-Mg^{2+}). The GDP-Mg^{2+} coordinates from the human rhoA-GDP/Mg^{2+} structure were modified to add missing bonds and the ligand^{docked} manually into the nucleotide binding pocket of the racE Swiss model, using the rhoA-GDP-Mg^{2+} X-ray structure for guidance. Following docking, an energy minimization was initiated within the *discover*3 module of InsightII using the cff91 force field, and continued to a maximum derivative of <0.01 kcal/mol (final derivative $= 0.009$ kcal/(mol Å)). This automatic energy minimization of the racE-GDP-Mg^{2+} complex was performed in the presence of 1075 surrounding water molecules.

3. Results and discussion

3.1. Search for a suitable racE template

An exhaustive search for suitable templates for modeling was carried out using the BLAST [20], SAWTED [21] and JPRED [22] servers on the World Wide Web. Four criterion were used in the generation of the final model: (1) potential templates must have high sequence identity ($I \geq 50\%$), (2) templates must be X-ray crystal structures, as NMR structures used as modeling templates have been found to generate models that exhibit a higher root mean square deviation (RMSD) with their template than models generated using X-ray crystal structures [23,24], (3) templates must have significantly low *E* (expect value) and *P* (probability) values. *E* and *P* values create a significant threshold for reporting sequence match results (note: the lower the *E* and *P* values, the lower the probability that the sequence match is purely random), and (4) structures of templates satisfying the above three criterion must be determined in the absence of any effector proteins.

Results of the template search yielded two potential templates satisfying all of the criterion mentioned above—the human rac1 (1MH1.pdb) and the human rhoA (1FTN.pdb). Although the human rac1 protein had the higher sequence identity with racE (approximately 55%) (versus approximately 50% for human rhoA), the human rac1 crystal structure had a number of missing side-chains (for residues ^{32}Y, ^{33}I, ^{36}T, ^{123}K, ^{127}E, ^{128}K, ^{130}K, ^{131}E, ^{132}K, 133 K) and it was decided to use the more complete rhoA crystal structure as a template for predicting an initial three-dimensional structure of racE. The model begins at Thr numbered 5 in the model *pdb* file and ends at Glu-179 for a total of 175 residues, and accordingly referred to throughout this article. A pair-wise alignment for entire sequences of rhoA and racE as well as the region modeled as indicated by brackets is shown in Fig. 1.

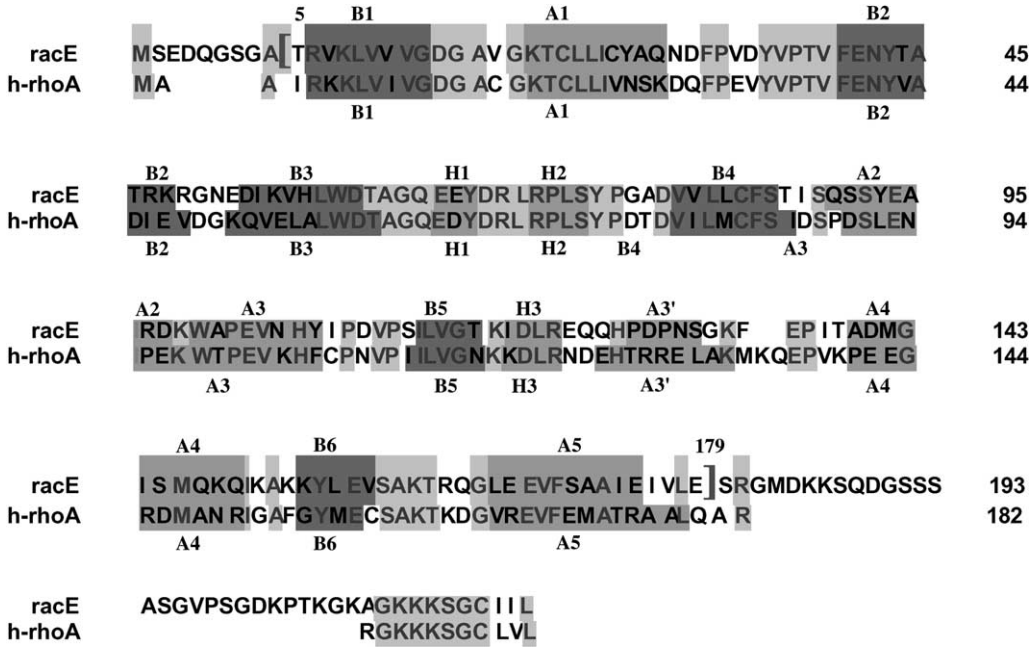
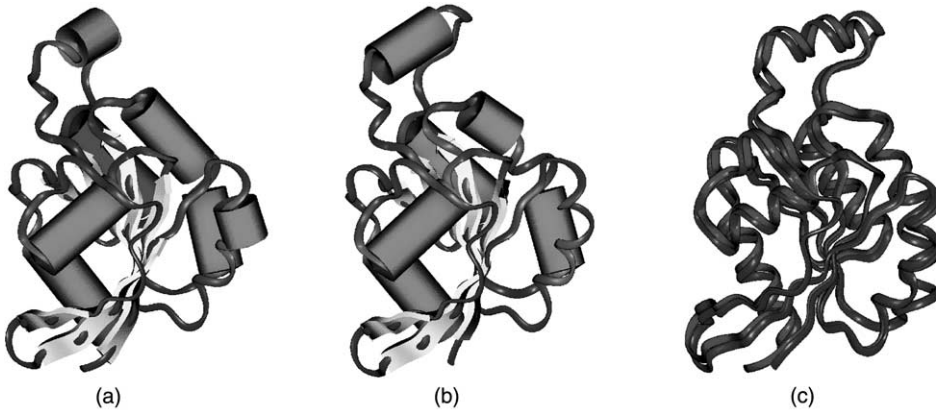


Fig. 1. Pair-wise amino acid sequence alignment of racE and human rhoA (h-rhoA) model template. racE shares 49.3% sequence identity with h-rhoA. The model begins at Thr5 and ends at Glu179, indicated by [ ] brackets. A1-A5, H1 and H2 refer to the α-helices and 3_{10} helices in the model, while B1-B6 refer to the β-strands and are shown in dark grey, the sequence identities between the two proteins is shown in light grey.

3.2. Swiss model three-dimensional structure prediction of racE

A three-dimensional modeling request was submitted to the Swiss model protein modeling server [25–29] on the World Wide Web (http://www.expasy.ch/swiss-model), using the 'first approach mode' with human rhoA specified as the template. Fig. 2(a) presents the Swiss model predicted structure of racE. For comparison, Fig. 2(b) shows the X-ray crystal structure of the human rhoA template. The superposition of the C_α traces of predicted racE and experimentally-determined rhoA structures are shown in Fig. 2(c). Examination of Fig. 2(a)–(c) reveals the great similarity between the two proteins. Fig. 2(d) presents the Ramachandran plot for the Swiss model predicted racE structure. The plot statistics indicated that 89.5% of the 175 residues are in the most favorable regions. Only two residues, Glu65 and Asn51 lie in a disallowed region. (Note: subsequent energy minimizations reduce the number of residues in disallowed regions to one, Asn51, and even this residue moves to a position very close to an allowed region (not shown).)

The predicted racE model in Fig. 2(a) reveals a common fold conserved in all members of the Ras super-family.



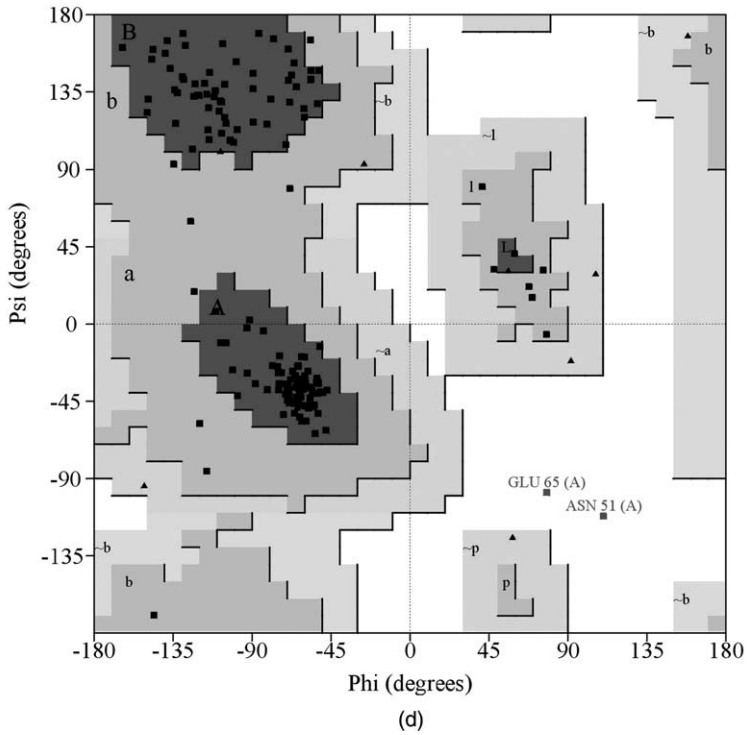


Fig. 2. The three-dimensional model of racE and its comparison with the human rhoA crystal structure. (a) Swiss model of racE; (b) the human rhoA crystal structure bound to GDP·Mg^{2+} (not shown) in the same orientation as (a). (c) Superimposition of racE upon rhoA. A superimposition of 350 C_{alpha} atoms (trace backbone) of racE upon h-rhoA, with root mean square deviation (rmsd) of 3.723. The two backbones are shown in ribbon form. (d) Ramachandran plot of Swiss model racE. 89.5% of all 175 residues are found in most favorable regions. Asn 51 and Glu65 are seen to be in disallowed regions. These residues moved to the borderline of an allowed region after energy minimization of the Swiss model (not shown).

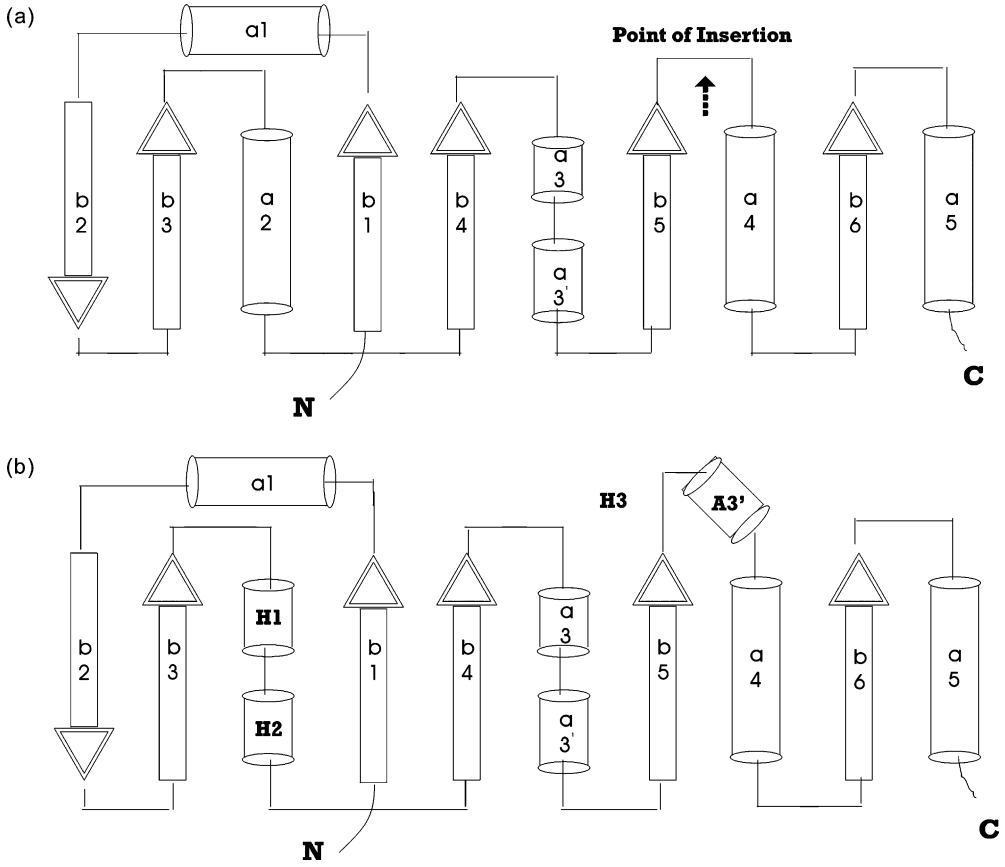


Fig. 3. Overall topology of the Swiss model of racE in comparison with the p21 h-ras crystal structure [44]. The two protein structures share the same topology of secondary structural elements (a, b). The α-helices are labeled α1–α5, and the β-strands as β1–β6. H1, H2, H3 are 3_{10} helices, while A3′ in race (b) refers to the insertion helix. The point of insertion in the h-ras structure is indicated by an arrow (a), where all Rho sub-family proteins exhibit a 11–13 residue insertion, also seen in the racE model (b).

It shares the exact secondary structure topology with ras (Fig. 3(a)), which includes a central β-sheet consisting of six β-strands, five being parallel (β1, β3, β4, β5, and β6) and one (β2) anti-parallel and five conserved α-helices. A major difference between the racE model and the ras structure is the presence of an *insertion domain*, comprising 11 residues (^{125}Gln–^{135}Phe) (Fig. 3(b)). This domain is a special feature shared by all Rho proteins and has been implicated in effector binding and downstream signaling events [30,31]. The human rhoA and the human rac1 proteins each contain a 13-residue insertion domain, compared with racE's 11-residues, and ras exhibits a short loop for this region. The insertion domain of racE is preceded by the G4 loop, which follows the β5 strand, and has a tendency to exhibit two helices (^{119}IDLR, a 3_{10} helix, preceding the insertion, and ^{126}PDPNS) designated as H2 and A3′, followed by a 7-residue-long loop.

Another relevant comparison between the predicted racE and rhoA structures relates to the functionally-significant Switch II region, comprising residues 62–79 in racE. This region shows the exact helical conformation as seen in the rhoA (and rac1) proteins. The sequences in both rhoA and racE are seen as short 3_{10} helices (in racE: ^{65}EEY and ^{71}RPLS, separated by the short loop, ^{68}DRL).

The RMSD for the 350 superimposed C_α atoms in racE and rhoA (Fig. 2(c)) is equal to 3.726 Å. Many hundreds of sequences with known structures were modeled during the three-dimensional Crunch Project [32–34] using the Swiss model ProMod II modeling software using the 'first approach mode'. 78 and 86% of the sequences that shared 40–49% and 50–59% identities with their templates generated a model that had a RMSD of less than 4 and 5 Å from their control templates, respectively (http://www.expasy.ch/swissmod/SWISS-MODEL, under reliability of models). The racE shares 49.3% sequence identity with its modeling template, the h-rhoA. The resulting racE model RMSD of 3.726 from its h-rhoA template is within the predicted limits of the Swiss model study. Regions measurably different in the two proteins that may contribute to this deviation are certain loop regions (i.e. the loops between β2 and β3, A3′ and A4) and the insertion helix (A3′) (Fig. 3(b)). These are regions between the two proteins with entirely different sequences. The racE model has a β hairpin motif joining two anti-parallel β-strands,

β2 and β3 separated by a short hair-pin loop. This loop is conserved in cH-ras, rac1 and rhoA proteins, and is two residues long (^{49}DG). The corresponding loop in the predicted racE model is four residues long (^{49}RGNE). This loop is a classic Type II reverse turn with glycine at the second position in the loop in all proteins. Except for Gly, this loop shares neither sequence identity nor length with its template. In addition, this region in the human rhoA-GTP structure showed high mobility, since the two β strands are followed by the relatively mobile Switch I and II sequences.

3.3. Docking of GDP into the predicted racE structure

Although the main objectives of this study were (1) to predict the three-dimensional structure of racE in the absence of any bound ligand (using a known crystal structure of a homologous protein as a template) and (2) to assess the stability of this structure in a molecular dynamics simulation, we also decided to see if the GDP cofactor could be docked into a potential binding cavity which appeared to be present in the predicted structure.

The coordinates for the GDP/Mg^{2+} complex were obtained from the protein data bank file of the human rhoA-GDP/Mg^{2+} structure (1FTN.pdb). The GDP structure was then modified within the Insight II interface to allow for correct bond order and missing (hydrogen) atoms. Potentials were calculated for all atoms. (Calcium parameters were used for the magnesium ion to maintain compatibility with the cff91 force field. As all metals are treated as non-directional point charges in this force field, this is a reasonable approximation.) Manual docking was then performed to fit the GDP into the apparent racE binding cavity, using the rhoA-GDP complex as a guide. Following reasonably successful docking, the resulting initial complex

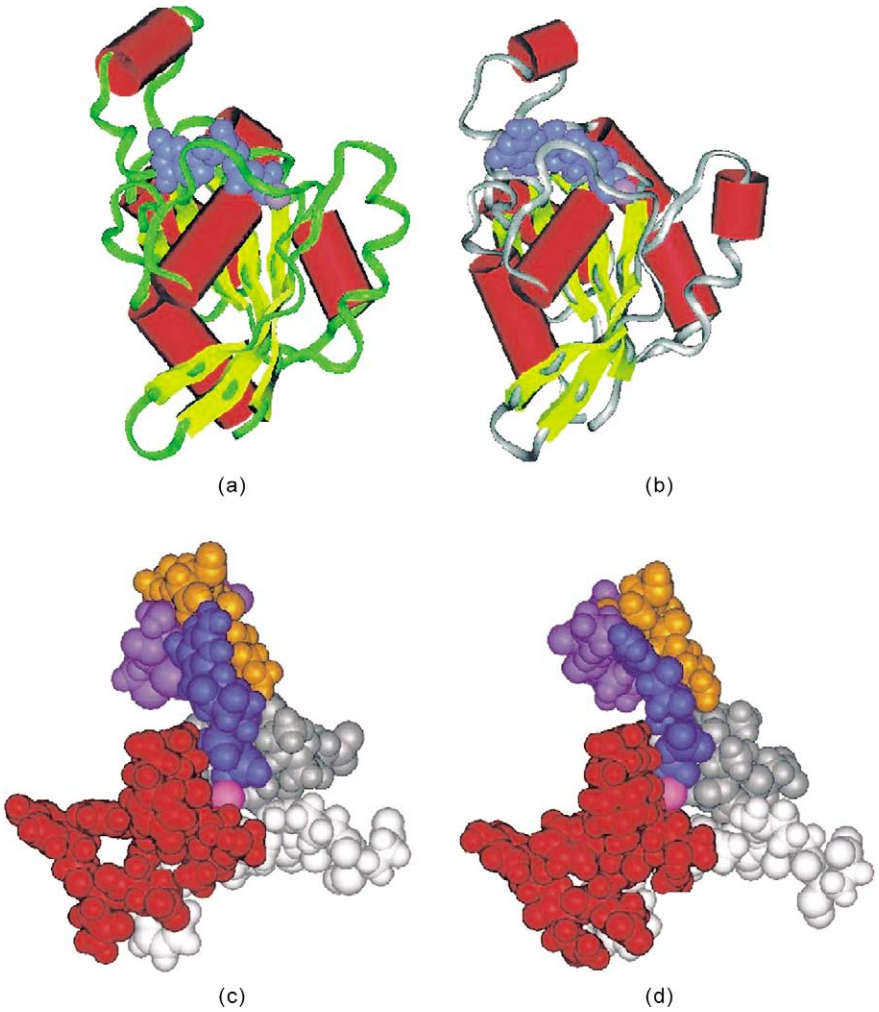


Fig. 4. Final energy minimized GDP·Mg^{2+} docked Swiss model racE and RhoA·GDP·Mg^{2+} crystal structure and their nucleotide binding pockets. The position and outlay of the nucleotide binding pocket is shown in RhoA·GDP·Mg^{2+} crystal structure (a) and the manually docked, energy minimized (maximum derivative = 0.005613 kcal/(mol/Å)) racE model (b). (c,d) Nucleotide binding pockets of rhoA structure and racE model are shown in space-filling rendering, respectively. The nucleotide binding pocket is made by 5 conserved loops (G1–G5). RacE: G1 (^{13}GDGAVGKT^{20}) is shown in grey, G2 (^{35}YVPTVFENY^{43}) in red, G3 (^{59}WDTAGQE^{65}) in white, G4 (^{118}TKID^{121}) in orange, G5 (^{158}VSAK^{161}) in purple, GDP in blue, and Mg^{2+} in pink. Corresponding G1–G5 loops in the h-rhoA structure is shown in similar colors.

was solvated with 1075 water molecules and subjected to an energy minimization using the *discover*3 module within the *Insight II* interface. Minimization was performed, using the conjugate gradient method, until a maximum derivative of 0.069 kcal/mol was attained. The resulting structure from the racE:GDP minimization is shown in Fig. 4(a). Fig. 4(b) presents the structural cartoon for the rhoA: GDP complex (from the X-ray crystal structure) in approximately the same orientation as the racE:GDP complex (Fig. 4(a)). The five "G-loops", which make close contact with the bound ligand (also illustrated) in both proteins, are also shown (Fig. 4(c) and (d)). Comparison of the nucleotide binding pockets in the predicted racE structure (Fig. 4(c)) and the experimentally determined rhoA structure (Fig. 4(d)) reveals great similarity, lending further credibility to the Swiss model predicted structure. In both proteins, the GDP is deeply buried in the binding pocket, although there are slight differences in both the arrangement of side chain groups coming off of the various G-loops as well as in the conformation of the bound GDP (Note: The racE loops (G-1 to G-5) shown in Fig. 4(c) appear to be larger than the rhoA loops due to the presence of hydrogen atoms on the racE predicted structure, but not on the rhoA X-ray crystal structure). Where appropriate, structural features of the energy-minimized racE-GDP complex will be related to relevant results obtained from the molecular dynamic simulation of the nascent structure to be described in the following section.

3.4. Molecular dynamics simulation of the energy-minimized racE structure

3.4.1. General considerations

A 300 ps, 300 K molecular dynamics (MD) simulation was carried out for the solvated, energy-minimized racE structure in the absence of any bound ligand. The main objective was to assess the structural stability of the model generated by the Swiss model procedure. The program *fdiscover*, with the cff91 force field, was employed for the MD simulation. The starting structure for the run was generated by an energy minimization (maximum derivative <0.001 kcal/mol) of the Swiss model structure solvated with a layer of 2308 water molecules.

3.4.2. Stability of the secondary structure during the MD simulation

Fig. 5(a) presents a plot of total energy versus time of the MD simulation. The profile shows an energy plateau after only about 50 ps, indicating that equilibrium has been attained in the simulation. In Fig. 5(b), 13 dynamic structures obtained every 20 ps after a 30 ps equilibration period during the course of the MD simulation are superimposed. In these images, only the trace of the main chain is shown. Note that the ribbon traces of the main chains shown are tightly clustered in most parts of the protein, indicating overall stability of the secondary structure. (For comparison, the

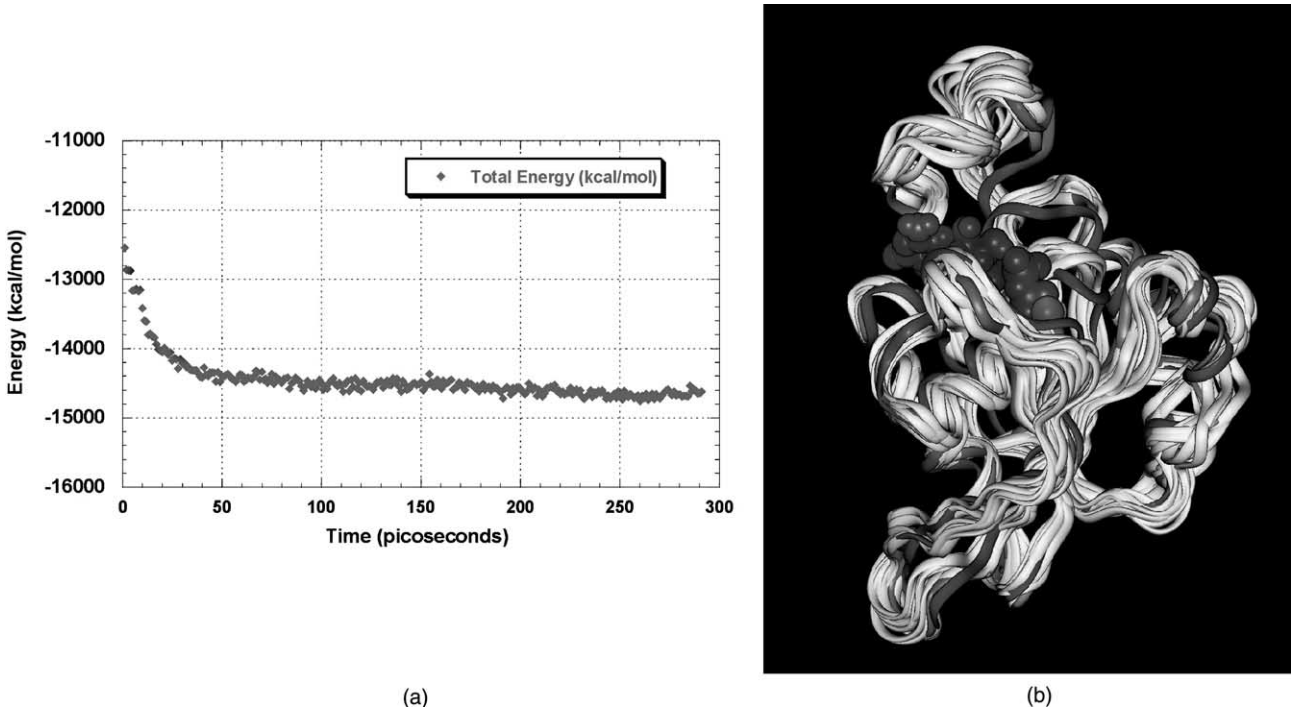


Fig. 5. Stability of overall structure during molecular dynamic (MD) simulations of the energy-minimized nascent racE Swiss model. (a) Total energy diagram indicating that the total energy (total potential + total kinetic) decreased continuously and reached equilibrium early during the 300 psec MD run. (b) Stability of the secondary structure. C_α trace (shown in ribbon form) superimposition of all 13 archive frames upon the starting structure (energy minimized) nascent Swiss model (light grey) and the GDP·Mg^{2+} docked structure (dark grey) shows tight clustering of most regions indicating stability.

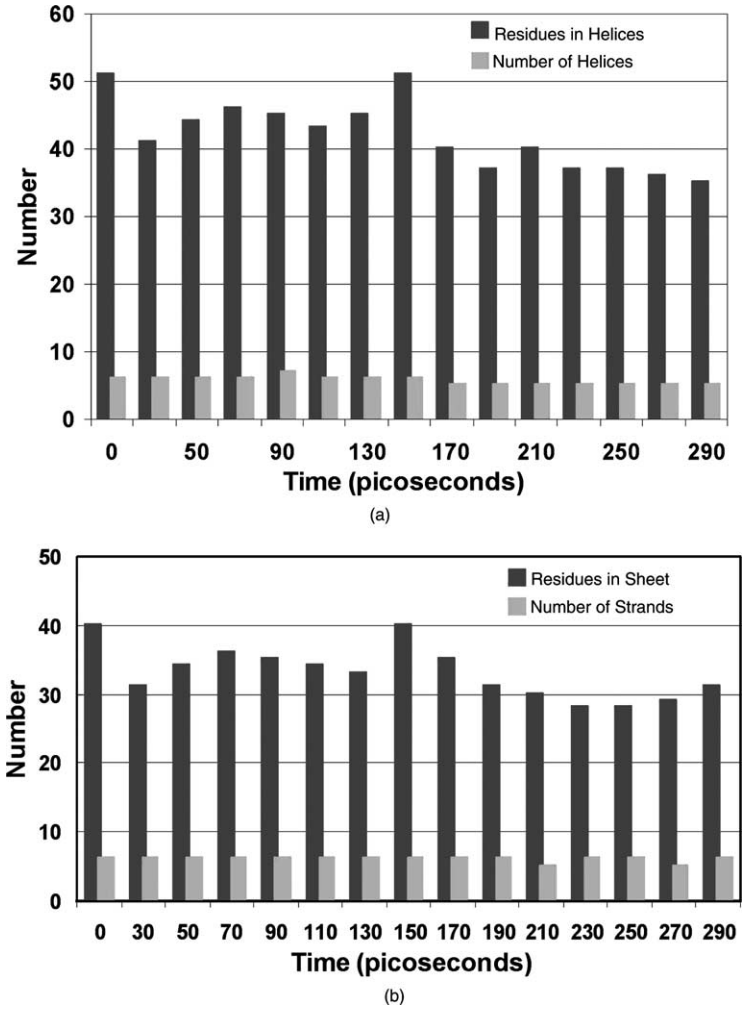


Fig. 6. Secondary structure stability during molecular dynamic simulations. (a) Change in total number of α-helices and residues in helices during MD simulations. The number of helices is mostly six, while the total number of residues in helices are averaged at 40. (b) Change in total number of β-strands and residues in β-sheet during MD simulations. The number of β-strands is stable throughout, fluctuating at the end between 5 and 6, while the total number of residues averages at 30.

superimposed trace of the main chain associated with the energy minimized nascent racE protein in the presence of bound GDP (in red) is also shown.)

Fig. 6(a) presents a plot of the number of helices (both α-helices and 3_{10} helices) and the number of helical residues in racE during the MD simulation. Fig. 6(b) presents the corresponding plot for the β-sheet structure in the protein. Overall, there are no dramatic changes in the secondary structure content during the MD simulation. There is minor fluctuation in both the number of helices and the number of β-strands, and a slight decrease in the number of residues in both types of secondary structure during the time trajectory. The five α-helices and six β-strands that form the basic fold of all ras super family proteins were found to be generally conserved throughout the MD run. The slight helical instability that was detected during the MD simulation was associated primarily with the 3_{10} helical and the insertion helix regions (i.e. the two 3_{10} helices in the Switch II region, the IDLR 3_{10} helix and the PDPNS α-helix in the insertion region). It is not unexpected to find increased dynamic behavior for residues in these regions upon input of kinetic energy during these MD simulations, since it allows them to sample other conformations in an attempt to break free of the tight 3-residue turn that is associated with the 3_{10} helix. The differing dynamic structures observed during the MD simulation could be the logical result of these stretches being part of the more flexible switch/insertion domain regions of the protein.

The racE β-strand composition during the MD simulation was slightly more stable than the α-helix composition, as the six-strand structure was retained for almost all of the MD run. (During the last half of the MD run, the sixth β-strand (^{154}KYLEV) exhibited fluctuating behavior. This strand overlaps the N-terminal portion of the G5 loop (KKA**KOOE** (**A**/**C**/**S**/**T**) SAK), ranging from 2 to 5 residues in length during the MD simulation.) An interesting observation is

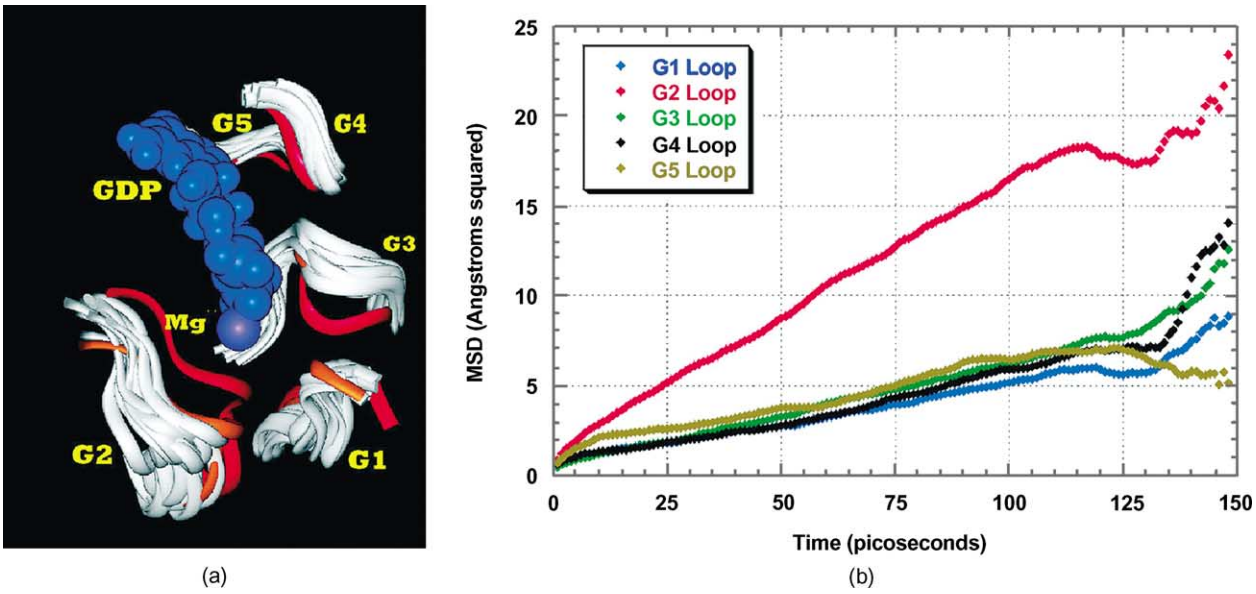


Fig. 7. Dynamic behavior of G-loops during molecular dynamic simulations. (a) All archive frames superimposed as in Fig. 5(b) exhibiting loops G1–G5. All G1–G5 loops move closer to the GDP·Mg^{2+} moiety, with the maximum movement shown by the G2 (effector) loop. (b) Mean square displacement (MSD) values plotted against time indicates the extreme mobility of the G2 loop in red, while all other G loops exhibit on average, movements of up to 5 Å. The G loops of the starting structure are shown in orange, while the docked structure G loops are colored red. All other archive frames G loops are seen in white, GDP in blue, and Mg^{2+} in pink.

that racE does not contain any of the four A/C/S/T residues commonly found preceding the conserved SAK sequence. Instead, it has a bulky, hydrophobic Val at this position. The significance of this occurrence remains to be elucidated.

3.4.3. *The "G-loops" of RacE*

Fig. 7(a) and (b) present the superpositioning of the five G-loop regions of racE and the mean-square-displacement (MSD) traces for these loops during the MD simulation, respectively.

Four of the five G loops (G-1, G-3, G-4 and G-5) present relatively shallow MSD profiles, indicative of relatively low motion. In contrast, the G-2 (or "effector") loop presents a steep MSD profile, indicative of residues in this loop undergoing significant motion during the MD run relative to residues in the other G loops. As shown in Fig. 7(a), the G2 loop spans residue numbers 34 to 43/46 and forms a long curved loop around the bound nucleotide, protecting the cofactor from exposure to the solvent. The G2 loop serves as a "door" to open or close the nucleotide binding pocket.

The effector loop is one of the most well studied domains in the rho family proteins, and plays an important role in signaling events. Interactions between ras and its effector, raf, are known to be mediated by the effector loop in ras [35]. Mutations in the effector region of rac1 [36–38] resulted in significantly reduced activation of PAK and the NADPH oxidase complex in human neutrophil cells and in conjunction with rac/Cdc42 chimeras [39] highlighted its role in mediating interactions in the NADPH oxidase complex. Specific mutations in the effector region also inhibited rac binding to $p67^{phox}$,[27]. The N-terminus (residues 30–32) of the effector region in rac1 is involved in the chick Sem3A-induced growth cone collapse studied in COS cells [40]. As mentioned above, the effector loop in the racE model is remarkably flexible, with an MSD of ∼20 $Å^2$ at the 150 ps mark in the MD run (Fig. 7(b)). This dynamic behavior of the G-2 loop is implicated to allow decreased surface affinity with the effector upon GTP hydrolysis allowing the effector molecule to dissociate [41]. The effector loop shows poor electron density in the rac1-GMPPNP crystal structure, possibly due to disorder or mobility. The disordered side-chains and the flexibility of the effector loop is supported by the observation that upon GDP docking, the effector loop displaces itself to move significantly closer to the GDP moiety, especially near the metal ion binding site, allowing the backbone to attain proximity to the α-, β-phosphate groups and the metal ion. Such displacement upon GDP binding (which is equivalent to the post GTP hydrolysis state) leads to a 'loose' complex, with decreased affinity for its effector, leading to subsequent dissociation. It also generates a tighter complex with GDP that, in effect, serves to maintain the protein in an inactive state. In this state, the protein can respond to an activating signal, releasing the tense GDP complex and forming a less tightly bound active, GTP complex allowing effectors such as GAP and GEF molecules to interact. The effector loop is therefore an important domain in racE's effector interactions, and its behavior during energy minimization and MD suggests that it plays an important role in *Dictyostelium* cells.

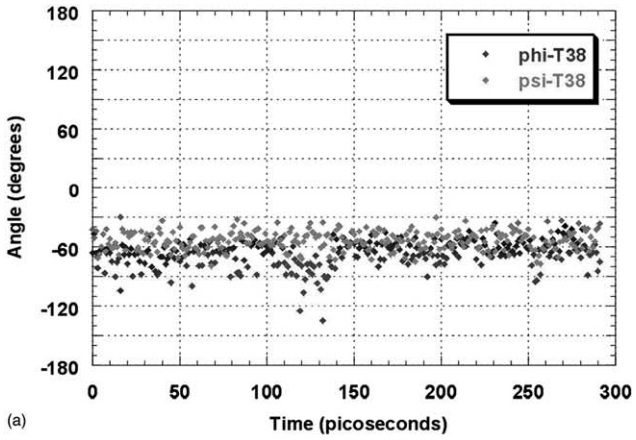
What is the source of the significant dynamic motion observed for the effector loop during the MD simulation (Fig. 7(b), red trace)? To address this question, the variation of dihedral angle values for residues 28–46 encompassing

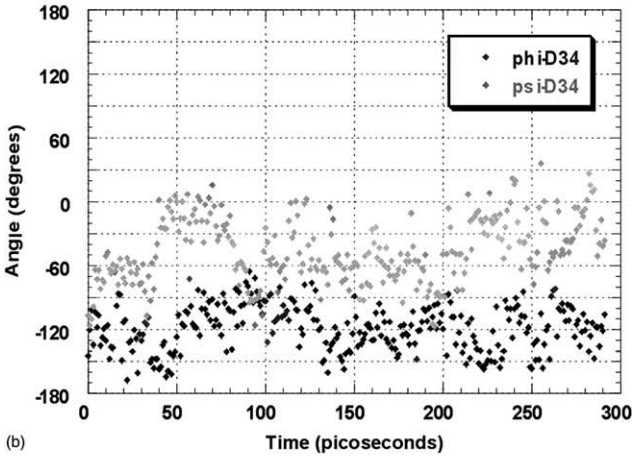
both Switch I (see Section 3.4.4) and effector regions were examined. While ^{28}Gln, present at the start of the Switch I, shows some variability, ^{34}Asp, the first residue in the effector loop, presents extreme variability (Fig. 8(a)). Thus, ^{34}Asp can conveniently serve as a "hinge" point for Switch I and, more importantly, effector loop motion. The MD simulation indicates that the entire effector loop moves as a single entity, either towards or away from the nucleotide binding pocket. In contrast, the rest of the residues in the 29–46 stretch (i.e. residues other than ^{34}Asp) show minimal dihedral angle variation. The phi and psi dihedral trajectories for ^{38}Thr, a functionally important residue in racE, are representative of these other residues (Fig. 8(b)). Fig. 8(c) shows the positions of ^{34}Asp and ^{38}Thr in the racE structure.

3.4.4. The switch regions of RacE

The switch regions are two domains that change conformation in small GTPases when the protein *switches* between the GTP- and GDP-bound states [42]. The Switch I region (^{29}NDFPV**DYVPTV**^{39}) of racE, as in h-ras and rac1, *overlaps* the effector loop 34DYVPTVFENYTAT46 (overlap shown in bold). The switch regions are well conserved in the rho sub-family, but are different from the ras-subfamily of GTPases. Changes in the amino acid sequence of the switch regions in Rho homologs often result in small changes in charge in these regions. In some cases, loss of charge *or* charge reversal occurs, while in other cases substitutions occur which leave charge unchanged (as in a D → E substitution). These alterations tend to be localized at the beginning of the Switch I (^{30}D, in racE) and effector loop regions (^{33}VD, in racE), with additional substitutions occurring at the end of the effector loop (^{44}TAT, in racE). Small changes in the charge distribution over the Switch I and effector regions *may* be important in discriminating between different members of the Rho family, allowing for several protein-specific effector interactions. Residues involved in binding and hydrolyzing GDP/GTP seem to be conserved between all of the ras and rho subfamily members. Specific implicated residues include a threonine (^{38}Thr) in Switch I and a glutamine (^{64}Gln) in Switch II.

The MD simulation revealed that the Switch II region (^{63}Ala to ^{79}Asp) shows movement of its backbone, although the overall fold of the region throughout the MD simulation is similar to that observed in rac1 and rhoA proteins. Protein backbone movement for the Switches I and II regions of racE during the MD simulation indicates greater stability (i.e. less movement) for the Switch II region as compared to the Switch I region (see Fig. 9). This apparent stability of the Switch II region is also observed in rhoA·GTP, rac1·GTP, cdc42-GAP and heterotrimeric G proteins [41]. Switch II residues 63–65 are disordered in almost all ras structures resolved to date, and the rhoA-GDP structure [1,43,44]. However, in the racE model, there is no extreme backbone mobility or flexibility observed, although the region is on the average slightly more flexible than other domains in the protein. Conserved residues at the





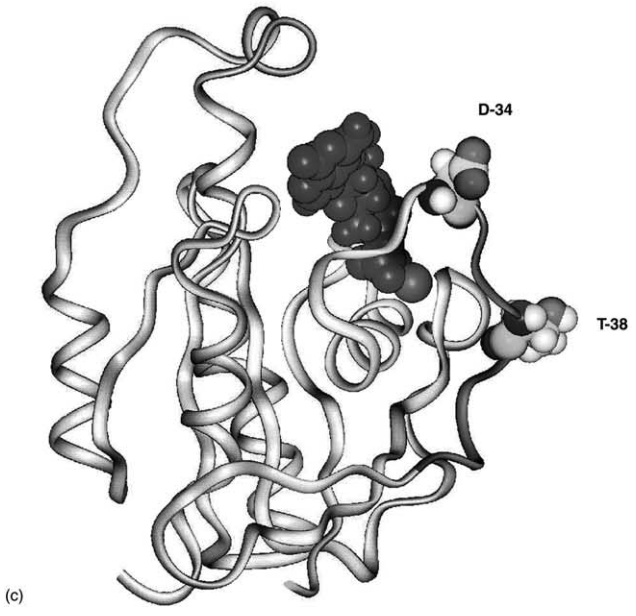


Fig. 8. Dihedral angle distribution of certain residues during molecular dynamic simulations. (a) ^{38}Thr is seen to be the most stable residue in the effector loop during MD simulations, while (b) ^{34}Asp exhibits the most variation in dihedral angles of the entire G2 loop ranging from residues 34 to 43. (c) ^{43}Asp, ^{38}T, and GDP·Mg^{2+} are shown in space-filling rending on the trace of the main chain.

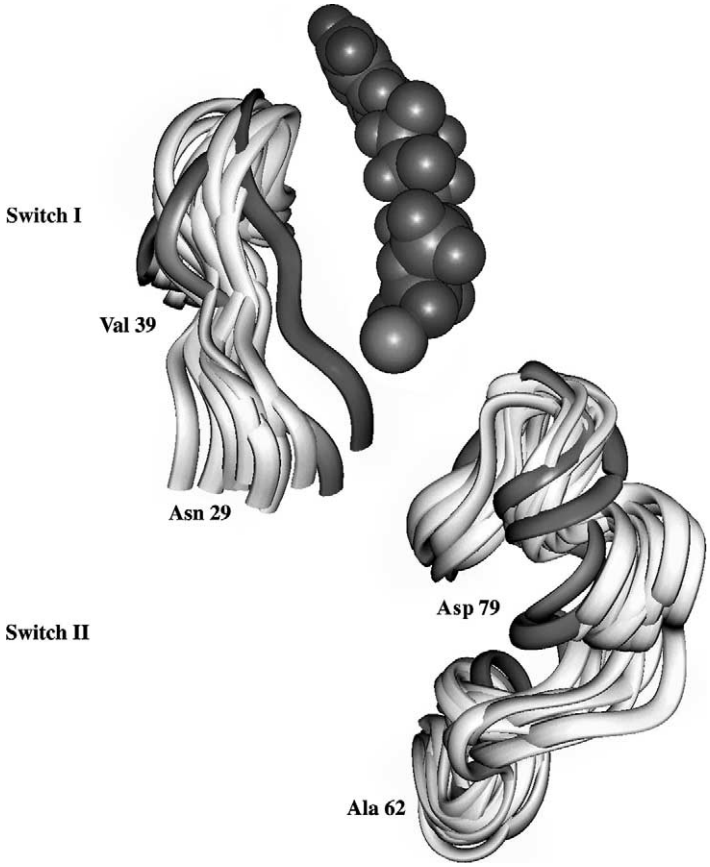


Fig. 9. Switch regions of racE during molecular dynamic simulations. Switch I region shows large movement toward the GDP·Mg^{2+} moiety, while the Switch II region does not. The Switch I (29–39) and Switch II (62–79) regions are shown as all archive structures in white, starting structure in light grey and the docked structure in dark grey. The GDP·Mg^{2+} moiety is in space-filling rending.

beginning of the Switch I and at the end of Switch II are well ordered, whereas the rest of the two domains exhibit slightly disordered side-chains during the MD simulation.

Interesting structural features commonly found on the C-terminal flank of the Switch II region are a pair of single turn 3_{10} helices. The two rho homologs rhoA and rac1 exhibit these two short 3_{10} helices. The first 3_{10} helix has the sequence EDY in both proteins, the second contains the sequence RPLS in rac1, and RPL in rhoA, separated by a short conserved loop, DRL. The racE model structure contains similar 3_{10} helices, with ^{65}EEY and ^{71}RPLS separated by the DRL loop. Addition of the GDP moiety into the nucleotide-binding pocket of the model caused the backbone in this domain to move closer to the GDP, although the distance remained around 4–7 Å between the β-phosphate group and the backbone atoms of the N-terminal Switch II region. The substitution of GTP for GDP in the model would cover an additional 2.5 Å with the triphosphate moiety to bring the backbone as well as the side-chain groups within coordination distance of the guanine nucleotide phosphate groups and the coordinated metal ion. The ^{64}Gln side chain exhibits various conformations within the nucleotide binding pocket ranging up to 5 Å, but still faces the β-phosphate oxygen throughout the MD simulations.

3.4.5. Insertion domain

An insertion of approximately 13 residues between strand β5 and helix α4, which forms a *helical domain*, is a characteristic feature present in all Rho family homologs, but *not* in ras. The insertion domain is known to contain several charged residues and (as observed in rac1 and rhoA) is solvent exposed [41]. The results of recent site-directed mutagenesis and deletion experiments in rac1 [30] have suggested a role for this domain in effector binding and enzyme activation processes. The insertion domain in racE extends from residue 120 to residue 138. The insertion itself is 11 residues long with the sequence ^{126}QHPDPNSGKFE. This domain has a similar conformation to rac1 and rhoA, and its secondary structure consists of two helices. The ^{120}XDLRX′ sequence, (X = I in racE; K in rhoA and h-ras; L in rac1, and X′ = E in racE; N in rhoA; D in rac1and A in h-ras) which forms a short 3_{10} helix in rhoA and racE, and an α-helix for the same sequence in rac1. A short loop of two residues follows in rac1 (^{121}DD) and rhoA (^{124}ND) while this loop is four residues long in racE (^{124}EQQH). The insertion contains a α-helix in racE, which is five residues long (^{128}PDPNS). The corresponding region in rac1 and rhoA exhibit longer helices (^{123}KDTIEKLK in rac1) and

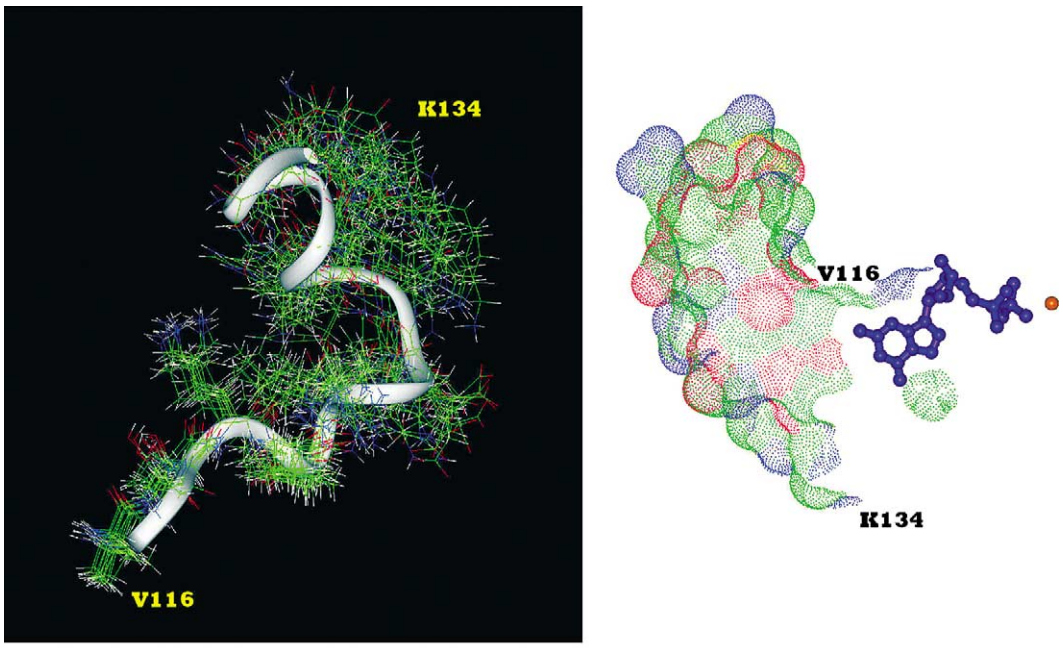


Fig. 10. Insertion domain in racE is dynamic, flexible and has a complex charge distribution. (a) All 13 archive structures (in wire frame) are superimposed on the starting structure (white ribbon). Most residues in the insertion domain (^{116}V–K^{134}) are dynamic, scattered in their distribution leading to an unstable insertion domain during MD simulations (atoms: C—green, O—red, N—blue, H—white). (b) The insertion domain exhibits a complex charged surface that is solvent accessible (red-negative, blue-positive, green-neutral/hydrophobic). GDP is in purple, Mg^{2+} in orange. The beginning and the end of the domain are indicated by ^{116}V and ^{134}K.

(^{126}EHTRRELAK in rhoA). This region in ras consists of a short loop.

Fig. 10(a) presents the superpositions of the insertion domains of the dynamic structures recorded during the racE MD simulation. (The white ribbon gives the trace of the insertion domain main chain of the first archive structure, recorded after 30 ps of dynamics.) The MD results indicate that the residues in the insertion domain of racE adopt different conformations during the simulation and the domain is quite flexible, with an MSD of $\sim$15 Å^2 at 150 ps (not shown). These results are consistent with X-ray crystal data obtained for rac1, where atoms of the insertion domain are observed to be quite mobile, exhibiting average B factors $\sim$10 Å^2 higher than that measured for atoms in the rest of the protein [2]. The first residue in the ^{120}IDLR helix in the insertion domain is an Ile in racE. Long side chain residues at this position without branching at the β carbon side chain, such as Lys in rhoA or Leu in rac1, tend to be compatible with the helical conformation. But racE contains an Ile residue, which may be too bulky to optimally fit into the small diameter 3_{10} helical conformation, causing instability (i.e. increased motion) in this region during the MD simulation, with an overall MSD value for the region greater than 20 Å^2 at the 150 ps mark (Fig. 10(b)). It also is of interest to note that the residues in the loop before (118–126) and after (132–138) the ^{127}PDPNS helix in this domain exhibit considerable variability in the dihedral angle values over the MD trajectory. Of these only ^{122}L and ^{135}F are hydrophobic residues, while most of the rest are charged/polar. The insertion domain displays a “spreading” of its charged surface, with positive and negative patches forming a complex pattern over the entire insertion domain, with side-chains that are fully solvent accessible (Fig. 10(b)). Another source of instability in this region relates to the presence of the two proline residues in the ^{126}PDPNS helical segment. While a proline residue fits well as part of the first-turn of the helix, it is generally known to ‘kink’ the helix if present in the middle of it. In addition, Ser is known to be a poor helix conformer. Therefore, it is not surprising that the ^{127}PDPNS helix, containing proline residues in the first and third positions and serine at the fifth position, is also unstable during MD.

3.4.6. The nucleotide binding pocket

Side chains of residues of the highly conserved G-1, G-3, G-4 and G-5 loops and the less conserved G-2 loop line the nucleotide binding pocket. The guanine ring and ribose sugar atoms interact most strongly with residues of the G-4 and G-5 loops, while the α- and β-phosphate groups interact primarily with G-1 and G-2 loop residues. Our model further indicates that residues in the G-3 loop are, for the most part, too far away to interact strongly with any atoms of the bound GDP. This is not unexpected, since the G-3 loop contains ^{64}Q, which is the so-called γ-phosphate sensor, relevant in the GTP-bound state.

A number of specific interactions between protein residues and the GDP ligand are revealed in the model. The side chain amino groups of ^{119}Lys and ^{161}Lys interact with the endocyclic oxygen, the C^5 oxygen as well as the oxygen (O^6) connected to the guanine ring, respectively. The backbone oxygen of the conserved ^{121}Asp is in proximity to the N^7 atom of the guanine nucleotide moiety. Backbone carbonyl oxygen and amide hydrogen atoms of the G-1 loop residues ^{17}Val, ^{18}Gly, ^{19}Lys, and ^{20}Thr are involved in extensive interaction with oxygen atoms of the α- and β-phosphate groups and the metal cation. The metal cation interacts strongly with the backbone carbonyl oxygen atoms of ^{20}Thr, ^{37}Pro, and ^{38}Thr. A water molecule is seen to bridge ^{38}Thr to the metal ion. The metal also interacts with two other water molecules, which are in close proximity to the various phosphate oxygen atoms (not shown), while the third water molecule is found directly between the side-chain of ^{64}Gln and the metal ion. Our current model indicates that a γ-phosphate group (of GTP) would be in close proximity to the ^{64}Gln in the G-3 loop, facilitating a strong interaction between ^{64}Gln and the bound ligand. The backbone and side-chain oxygen atoms of ^{60}Asp, and the backbone carbonyl oxygen atom of ^{61}Thr would also be quite close to the γ-phosphate group and the metal ion, leading to the formation of a tightly bound complex.

3.4.7. Some important residues in RacE

In ras, rac1 and rhoA proteins, the G-2 or the effector loop contains the critical residue, ^{38}Thr, which is known to coordinate Mg^{2+} ion with its *side-chain* oxygen when these proteins are bound to GTP [2,3,44]. On the other hand, the GDP bound structure of the human rhoA shows the *backbone carbonyl oxygen* coordinating with Mg^{2+}, instead of the side-chain oxygen. In our racE-GDP model, ^{38}Thr utilizes its backbone carbonyl oxygen for metal ion coordination while its side-chain remains solvent-accessible. The side-chain was seen facing the nucleotide-binding pocket when modeled using the human rac1-GMPPNP structure, confirming this residue's unique role in metal ion coordination in both the GDP- and GTP-bound states [6,44]. Although ^{38}Thr exhibits significant mobility, as reflected by its MSD value of ∼20 Å^2 at the 150 ps mark, values for the ϕ and ψ dihedral angles remain remarkably stable during MD trajectory, serving to rigidly maintain the orientation of its backbone carbonyl function which faces the nucleotide-binding pocket (Fig. 8(a)). The high mobility of the effector loop (including the observed displacement of the ^{38}Thr residue) is a consequence, in part at least, of the ϕ–ψ angle rotations associated with the ^{34}Asp residue (Fig. 8(b)). Rotations about these main chain dihedrals would result in effector loop movement, either closer to or away from the nucleotide-binding pocket during GTP/GDP exchange. The stability of the ^{38}Thr dihedrals suggests that there is no occurrence of a flipping of the ^{38}Thr side-chain into the nucleotide-binding pocket during MD simulation. Additional evidence is provided by the observation that the ^{38}Thr backbone carbonyl function is oriented towards the nucleotide-binding pocket in all 13 archive structures collected during the MD run.

It has been difficult to precisely establish the orientation of the residues following the conserved ^{59}WDTAG sequence and part of the Switch II (i.e. ^{64}QE E/D) in the crystal structures of active ras, rac1-GTP, rhoA-GDP and in the NMR solution structure of ras-GDP [2,5,43,44]. In all six different racE models analyzed, the backbone takes a sharp, abrupt 180° turn at ^{65}Glu pushing its side-chain outward into the solvent, lying across the width of the nucleotide-binding pocket. ^{66}Glu faces away from the pocket and is also solvent exposed, exhibiting different conformations.

The Switch II region is implicated as one of the regions used by GAP proteins to make contact in ras, rho and cdc42 protein GAP complexes [45]. Arg fingers are hypothesized to enhance the GTPase activity by interacting with two of the γ phosphate oxygen as well as the side-chain amino group of conserved Gln residue [46]. This residue is essential for GTP hydrolysis and, except in Rab, is conserved in the ras superfamily of GTPases [47]. Its substitution by any other amino acid side-chain, with the exception of Glu, hinders intrinsic GTP hydrolysis [35,48,49] and is oncogenic. In our study, the side chain of ^{64}Gln in racE exhibited various conformations, with its side-chain amino hydrogen between 4.3 and 7.0 Å away from the β-phosphate oxygen and the metal ion binding site. The presence of a water molecule and the γ-phosphate group could gap the distance by 2.5 Å, and the water molecule would be able to mediate an indirect interaction between the side chain amino groups and the metal ion. A water molecule is known to mediate ^{64}Gln side-chain attack leading to GTP hydrolysis and is seen present in the rhoA-GTP and rac1-GTP crystal structures. Amide hydrogen of ^{64}Gln and its side-chain oxygen atom are also seen to be within hydrogen bonding distance. ^{63}Glu, seen in h-ras (also present in racE), has been implicated in stabilizing the transition state of GTP hydrolysis by hydrogen-bonding to Gln61 [44]. In our model, the backbone amide group of ^{66}Glu is in close proximity (2.35 Å) to the ^{64}Gln backbone oxygen and this distance is more or less maintained (i.e. between 1.82 and 3.6 Å) throughout the MD simulation. The backbone and the side-chain of ^{66}Glu are, on average, flexible enough to allow their participation in GTP hydrolysis. ^{158}Val present before the G-5 loop (SAK) is not seen in any other GTPase, most of which contain either A/C/S/T residues only. There are no significant interactions of this residue side-chain with the bound GDP, and its significance at present is not understood.

3.4.8. Validation of the model and general conclusions

Our model analysis seeks credence from the widespread observation that while endowed with distinct cellular functions, all proteins of the Ras superfamily share the same overall three-dimensional architecture [6]. It has also been demonstrated experimentally that their role as diverse effectors in signaling pathways is due, not to gross structural

changes in the three-dimensional conformation, but rather, to local subtle changes occurring within a sub-set of domains in the proteins [50].

Modeling of a protein homologous to certain proteins of the Ras super family would generate a three-dimensional model that could be treated as a reasonable starting point for further analyses. Persistence of the initial predicted three-dimensional fold when the model is subjected to molecular dynamic simulations (at 300 K) should provide an assessment of the stability, and thus reliability of the model structure. On the other hand, since mechanics of folding into a particular conformation are inherent in the amino acid sequence of a protein, using homologous yet distinct templates for the purpose of modeling should in essence, produce a model with the same three-dimensional architecture with similar secondary structural elements. The crystal structures of human rac1 and human rhoA, the NMR structure of isolated cdc42 and crystal structures of cdc42 in complex with several effectors, have all been used to serve as templates for racE modeling. Among the templates, rhoA and rac1, rhoA and cdc42 share 58 and 53% sequence identity, respectively, while rac1 and cdc42 are almost 70% identical. The human rac1, human cdc42 and the human rhoA proteins share 55, 51 and 50% homology with racE. Since all three templates share similar sequence homology to racE, models generated using them yield the same three-dimensional conformation and secondary structure. The most similar result to the human rac1 and human rhoA based models is obtained from the cdc42 bound to rhoGDI based structure, secondary structures of all the three models are shown in Table 1. As mentioned earlier, the other cdc42 based structures differed in the insertion domain regions and the lengths of the secondary structural elements. All six of the α-helices and three 3_{10} helices predicted are observed in the three models. In addition, all models have remarkable consensus for the secondary structure regions. Minor differences are present in mostly the end of H5 and H6, and the beginning of two 3_{10} helices. The β-strands too, are similar with substantial difference present only in the length of S1. Differences in all other β strands are small, with a single residue difference. Thus, the current model of racE in all probability reflects the true tertiary structure of racE for approximately 1–190 residues. While most ras family or specifically the rho family homologues are 185–195 residues long, racE is unique with a length of 223 amino acids. Homology of racE with other proteins is for residues 1–190, and the C-terminal polybasic motif and CAAX box. Residues preceding the polybasic motif that is not present in the model are exclusive to racE and are predicted to exist as a long extended region. However, any secondary structural elements possibly existing in this region cannot be excluded.

Table 1
Consensus of secondary structural elements in racE models derived from various templates

	Model 1	Model 2	Model 3
Helices			
A1	KTCLLICYAQ	KTCLLICYAQ	KTCLLICYAQ
H1	-EYD	EEY-	EEY-
H2	RPLS	RPLS	RPLS
A2	QSSYEAIRD	QSSYEAIRD	QSSYEAIRD
A2′	WAPEVNHY	WAPEVNHY	WAPEVNHY
H3	IDL-	IDLR	IDLR
A3	PDPNS	PDPNS	PDPN-
A4	ADMGISMQKQ	ADMGISMQKQ	ADMGISMQKQ
A5	LEEVFSAAIE--	--EVFSAAIEIV	--EVFSAAIE--
β-Sheet			
β1	-----VKLVVVG	SGATRVKLVVVG	-----VKLVVVG
β2	-YTATRK	NYTATRK	-YTATRK
β3	DIKVHLWD	DIKVHLWD	DIKVHLW-
β4	VVLLCFS	VVLLCFS	VVLLCFS
β5	SILVGT	SILVGT	SILVGT
β6	KYLEV	KYLEV	KYLEV

Model 1 uses human rhoA-GDP crystal structure (1FTN.pdb) as template; Model 2 the human rac1-GMPPNP crystal structure (1MH1.pdb) and Cdc42 (bound to rhoGDI) was used as template to build Model 3.

4. Summary

Modeling the small GTP binding protein racE from *D. discoideum* using five templates generated racE models that are remarkably similar. The modeling procedure used various currently available resolved structures as templates such as the human rac1, human cdc42 and the human rhoA proteins, all of which share between 50 and 55% sequence identity with racE. Domains different among the various models included the first 3_{10} helix of the Switch II region, the 3_{10} helix of the insertion domain and the length of the insertion helix. Maintaining the criteria specified earlier in the paper and the crystal structures of several Rho family members, the human rhoA-based racE model was found to be the most accurate. The secondary structure of racE has seven α-helices, six β-strands and three 3_{10} helices. The Switch I and effector loop regions are mobile, whereas the Switch II region demonstrated average mobility during MD. A single residue "hinge" point, at ^{34}Asp, which allows the entire loop to be mobile as a single unit, mediates the specific mobility of the entire effector loop rather than several flexible portions within the effector loop. This may also be true of the Switch I region, which lies at the other side of the hinge point residue. The effector loop exhibits maximum displacement upon docking of the GDP ligand in an effort to establish optimal hydrogen bonding distances with the bound ligand, while G1 established the most number of interactions with the oxygen atoms of α- and β-phosphate groups. The insertion domain as a whole is highly mobile and flexible, with several positively and negatively charged residues forming a complex charged surface over the insertion domain. Certain residues (i.e. ^{38}Thr, ^{65}Glu and ^{51}Asn) exhibit remarkable mobility, ^{158}Val, ^{66}Glu, ^{34}Asp have a

unique presence in racE, as opposed to other Rho family homologues and it is therefore important to understand the involvement, if any, of all these residues in racE's function in *Dictyostelium* cells. The polybasic motif present in the C-terminal region of racE is another region implicated in effector functions from previous studies, and already observed in the human rac1 protein [51]. This motif and its preceding thirty residues with its yet unknown secondary structure have not been modeled yet, and still remains an enigmatic domain, important, in addition to other domains mentioned above, in designing future experiments to delineate their role in racE' signaling functions in cytokinesis.

Acknowledgements

We would like to express our thanks to the Chemistry Department of Clark University for providing the facilities in which this work was conducted. This work was funded in part by a NIH AREA award (1R15GM58658-01) to DAL.

References

[1] A.M. DeVos, L. Tong, M.V. Milburn, P.M. Matias, J. Jancarik, S. Noguchi, S. Nishimura, K. Miura, E. Ohtsuka, S. Kim, Three-dimensional structure of an oncogenic protein: catalytic domain of human c-H-ras p21, Science 239 (1988) 888–893.

[2] M. Hirshberg, R.W. Stockley, G. Dodson, M.R. Webb, The crystal structure of human rac1, a member of the rho-family complexed with a GTP analogue, Nat. Struct. Biol. 4 (1997) 147–152.

[3] K. Ihara, S. Muragachi, M. Kato, T. Shimizu, M. Shirakawa, S. Kuroda, K. Kaibuchi, T. Hakoshima, Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue, J. Biol. Chem. 273 (1998) 9656–9666.

[4] K. Rittinger, P.A. Walker, J.F. Eccleston, K. Nurmahomed, D. Owen, E. Laue, S.J. Gamblin, S.J. Smerdon, Crystal structure of a small G protein in complex with the GTPase-activating protein rhoGAP, Nature 388 (1997) 693–697.

[5] Y. Wei, Y. Zhang, U. Derewanda, X. Liu, W. Minor, R.K. Nakamoto, A.V. Somlyo, A.P. Somlyo, Z.S. Derewenda, Crystal structure of RhoA-GDP and its functional implications, Nat. Struct. Biol. 4 (1997) 699–703.

[6] H.R. Bourne, D.A. Sanders, F. McCormick, The GTPase superfamily: conserved structure and molecular mechanism, Nature 349 (1991) 117–127.

[7] A.G. Gilman, G proteins: transducers of receptor-generated signals, Annu. Rev. Biochem. 56 (1987) 615–649.

[8] Y. Kaziro, The role of guanosine 5′-triphosphate in polypeptide chain elongation, Biochim. Biophys. Acta 505 (1978) 95–127.

[9] L. Stryer, H.R. Bourne, G proteins: a family of signal transducers, Annu. Rev. Cell. Biol. 2 (1986) 391–419.

[10] A. Hall, Small GTP-binding proteins and the regulation of the action cytoskeleton, Annu. Rev. Cell Biol. 10 (1994) 31–54.

[11] A.J. Ridley, A. Hall, The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors, Cell 79 (1992) 389–399.

[12] K. Kishi, T. Sasaki, S. Kuroda, T. Itoh, Y. Takai, Regulation of cytoplasmic division of *Xenopus* embryo by rho p21 and its inhibitory GDP/GTP exchange protein (rho GDI), J. Cell. Biol. 120 (1993) 1187–1195.

[13] I. Mabuchi, Y. Hamaguchi, H. Fujimoto, N. Morii, M. Mishima, S. Narumiya, A rho-like protein is involved in the organisation of the contractile ring in dividing sand dollar eggs, Zygote 1 (1993) 325–331.

[14] D.A. Larochelle, K.K. Vithalani, A. De Lozanne, A novel member of the rho family of small GTP-binding proteins is specifically required for cytokinesis, J. Cell Biol. 133 (1996) 1321–1329.

[15] Y. Fukui, A. De Lozanne, J.A. Spudich, Structure and function of the cytoskeleton of a *Dictyostelium myosin*-defective mutant, J. Cell Biol. 110 (1990) 367–378.

[16] D.A. Larochelle, K.K. Vithalani, A. De Lozanne, Role of *Dictyostelium* racE in cytokinesis: mutational analysis and localization studies by use of green fluorescent protein, Mol. Biol. Cell 8 (1997) 935–944.

[17] N. Gerald, J. Dai, H.P. Ting-Beall, A. De Lozanne, A role for *Dictyostelium* racE in cortical tension and cleavage furrow progression, J. Cell Biol. 141 (1998) 483–492.

[18] D.A. Larochelle, N. Gerald, A. De Lozanne, Molecular analysis of racE function in *Dictyostelium*, Microsc. Res. Tech. 49 (2000) 145–151.

[19] J. Bush, K. Franek, J. Cardelli, Cloning and characterization of seven novel *Dictyostelium discoideum* rac-related genes belonging to the rho family of GTPases, Gene 136 (1/2) (1993) 61–68.

[20] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.

[21] R.M. MacCallum, L.A. Kelley, M.J.E. Sternberg, SAWTED: structure assignment with text description-enhanced detection of remote homologues with automated SWISS-PROT annotation comparisons, Bioinformatics 16 (2000) 125–129.

[22] J.A. Cuff, M.E. Clamp, A.S. Siddiqui, M. Finlay, G.J. Barton, Jpred: a consensus secondary structure prediction server, Bioinformatics 14 (1998) 892–893.

[23] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, Electrophoresis 18 (1997) 2714–2723.

[24] R.W. Harrison, D. Chatterjee, I.T. Weber, Analysis of six protein structures predicted by comparative modeling techniques, Proteins 23 (1995) 463–471.

[25] M.C. Peitsch, N. Guex, Large-scale comparative protein modelling, in: M.R. Wilkins, K.L. Williams, R.O. Appel, D.F. Hochstrasser (Eds.), Proteome Research: New Frontiers in Functional Genomics, Springer, Berlin, 1997, pp. 177–186.

[26] M.C. Peitsch, Protein modelling by E-Mail, BiolTechnology 13 (1995) 658–660.

[27] M.C. Peitsch, ProMod: automated knowledge-based protein modelling tool, PDB Quart. Newslett. 72 (1995) 4.

[28] M.C. Peitsch, ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling, Biochem. Soc. Trans. 24 (1996) 274–279.

[29] M.C. Peitsch, Large scale protein modelling and model repository, in: T. Gaasterland, P. Karp, K. Karplus, C. Ouzounis, C. Sander, A. Valencia (Eds.), Proceedings of the Fifth International Conference on Intelligent Systems for Molecular Biology, Vol. 5, AAAI Press, 1997, pp. 234–236.

[30] J.L. Freeman, A. Abo, J.D. Lambeth, Rac insert region is a novel effector region that is implicated in the activation of NADPH oxidase, but not PAK65, J. Biol. Chem. 271 (1996) 19794–19801.

[31] C.H. Kwong, H.L. Malech, D. Rotrosen, T.L. Leto, Regulation of the human NADPH oxidase by rho-related G-Proteins, Biochemistry 32 (1993) 5711–5717.

[32] N. Guex, M.C. Peitsch, Molecular modelling of proteins, Immunol. News 6 (1999) 132–134.

[33] N. Guex, A. Diemand, M.C. Peitsch, Protein modelling for all, TIBS 24 (1999) 364–367.

[34] C. Holden, High protein-crunch, Science 280 (1998) 1353.

[35] M.S. Marshall, The effector interactions of p21ras, Trends Biochem. Sci. 18 (1993) 250–254.
[36] D. Diekmann, A. Abo, C. Johnston, A.W. Segal, A. Hall, Interaction of Rac with p67phox and regulation of phagocytic NADPH oxidase activity, Science 265 (1994) 531–533.
[37] J.L. Freeman, M.L. Kreck, D.J. Uhlinger, J.D. Lambeth, Ras effector-homologue region on Rac regulates protein associations in the neutrophil respiratory burst oxidase complex, Biochemistry 33 (1994) 13431–13435.
[38] X. Xu, D.C. Barry, J. Settleman, M.A. Schwartz, G.M. Bokoch, Differing structural requirements for GTPase-activating protein responsiveness and NADPH oxidase activation by Rac, J. Biol. Chem. 269 (1994) 23569–23574.
[39] C.H. Kwong, A.G. Adams, T.L. Leto, Characterization of the effector-specifying domain of Rac involved in NADPH oxidase activation, J. Biol. Chem. 270 (1995) 19868–19872.
[40] I. Vastrik, B.J. Eickholt, F.S. Walsh, A. Ridley, P. Doherty, Sema3A-induced growth-cone collapse is mediated by Rac1 amino acids 17–32, Curr. Biol. 9 (1999) 991–998.
[41] S.R. Sprang, G proteins, effectors and GAPs: structure and mechanism, Curr. Biol. 7 (1997) 849–856.
[42] M.W. Milburn, L. Tong, A.M. deVos, A. Brnger, Z. Yamaizumi, S. Nishimura, S.-H. Kim, Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins, Science 247 (1990) 939–945.
[43] P.J. Kraulis, P.J. Domaille, S.L. Campbell-Burk, T. Van Aken, E.D. Laue, Solution structure and dynamics of ras p21.GDP determined by heteronuclear three- and four-dimensional NMR spectroscopy, Biochemistry 33 (1994) 3515–3531.
[44] E.F. Pai, W. Kabsch, U. Krengel, K.C. Holmes, J. John, A. Wittinghofer, Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation, Nature 341 (1989) 209–214.
[45] S.J. Gamblin, S.J. Smerdon, GTPase-activating proteins and their complexes, Curr. Opin. Struc. Biol. 8 (1998) 195–201.
[46] K. Scheffzek, M.R. Ahmadian, A. Wittinghofer, GTPase-activating proteins: helping hands to complement an active site, TIBS 10 (1998) 257–262.
[47] K. Scheffzek, M.R. Ahmadian, W. Kabsch, L. Wiesmuller, A. Lautwein, F. Schmitz, A. Wittinghofer, The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants, Science 277 (1997) 333–338.
[48] C.J. Der, T. Finkel, G.M. Cooper, Biological and biochemical properties of human RasH genes mutated at codon 61, Cell 44 (1986) 167–176.
[49] P. Polakis, F. McCormick, Structural requirements for the interaction of p21ras with GAP, exchange factors, and its biological effector target, J. Biol. Chem. 268 (1993) 9157–9160.
[50] A. Toporik, Y. Gorzalczany, M. Hirshberg, E. Pick, O. Lotan, Mutational analysis of novel effector domains in Rac1 involved in the activation of nicotinamide adenine dinucleotide phosphate (reduced) oxidase, Biochemistry 37 (1998) 7147–7156.
[51] G. Joseph, E. Pick, Peptide walking is a novel method for mapping functional domains of proteins, J. Biol. Chem. 270 (1995) 29079–29082.
[52] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283–291.