Electrostatic analysis of the Ras active site

Abdul Kareem, KHAN

PhD Thesis UPF

Barcelona 2009

UNIVERSITAT POMPEU FABRA

Electrostatic analysis of the Ras active site

Abdul Kareem, KHAN

PhD Thesis UPF

Barcelona 2009

Jordi Villà i Freixa
(Director de tesi)
Unitat de Recerca en Informàtica Biomèdica
Universitat Pompeu Fabra
Barcelona - 08003
Catalunya - Spain

White page:

- (Reserved for the TDX edition)

Dedicated to;

Three personalities who have sharpened my life
my Mother
my Wife
&
my Daughter

University of Balochistan, Quetta that inculcated the zest for research

“Pashtunkhwa” The mountnious terrain that belong to Pashtuns..!

مر ا م ټا کل عقل غو ا ړ ې

خو ګټل لپو نتو ب غو ا ړ ې

عبد ا لکر یم **خا ن** کا کړ
د ر ک : با ر سلو نا - ا سپا نپا

ACKNOWLEDGEMENT

I would like to extend sincere thanks to my thesis supervisor Jordi Villà i Freixa, for allowing me this opportunity to work with him in a totally new discipline (to me). Overlooking my unintentional mistakes and shortcomings. And ofcourse for his tremendous help and support in completing this thesis. Really Thanking you

I am very thankful to my wife (Safia Khatri), who took every pain to make it possible to allow me study abroad and today write down this thesis. I am indepthly thankful to you.

To my mother who always cared for me starting from class one and till today.. Words are not enough to express my gratitude to you, I really love you mom.

To my little shining star Salwa Khan, in your company I always used to forget every pain and had only joy. Thank you for not putting me into trouble in those late hours when mom was working and away. Thanks to Allah for blessing us with his blessings.

I wish to acknowledge Bea and Atif Lodhi, the tremendous friends whose company always make us feeling home. Thanks for helps while frequent shifting of the house and all those favors and hot debates over Green Tea.

I am thankful to Haji Gul Muhammad and Masoom Yasinzai for always cooperating, assisting, and helping me.

I am thankful to my brothers (Shakoor Kakar, Rahim Kakar) who were always eager to help me out in any affair of life and my sisters for always making prayers for me.

I am thankful to my brothers-in-law, who have always extended their family support to us and visiting us abroad especially Azizullah A. Khatri, Khalil A. Khatri, Asadullah A. Khatri and Naeem Khan Nasar for doing amazing jobs for us. To my sister in-law Kalsoom Khatri and Mussarat Arif for tremendous support… Thanking you all.

I am really thankful to Adrian L. García de-Lomana and Michael A. Johnston for always being ready to help me out in lab, especially for those Postscript files that Adrian gave to me. Thanking all CBBL members for their kind behavior.

I wish to extend my thanks to Javed Khan, Ziaullah Khan, Fatima Maghribi.
I am thankful to prof. Amiram Goldblum, a selfless person for guiding me during my admission and Prof. Ferran Sanz. I take my hat off.

I am thankful to all those Teachers who taught me during PhD courses and made the foundations for my degree. Thanks to all.

Also, to all those who have helped and appreciated me in the decade (1999-2008) that had been spent in 'janoon'.

Institutions:

Thanks to Family the first and prime institute in my life.

Thanks to University Pompeu Fabra

Thanks to Higher Education Commission of Pakistan.

Thanks to IMIM

Thanks to Masoom Educational Foundation.

Thanks to EU commission for financing CosQos Grid project

SUMMARY

The electrostatic preorganization of the active site has been put forward as the general framework of action of enzymes. Thus, enzymes would position "strategic" residues in such a way to be prepared to catalyze reactions by interacting in a stronger way with the transition state, in this way decreasing the activation energy $\Delta g^{\ddagger}_{cat}$ for the catalytic process. It has been proposed that such electrostatic preorientation should be shown by analyzing the electrostatic stability of individual residues in the active site.

Ras protein is an essential signaling molecule and functions as a switch in the cell. The structural features of the Ras protein in its active state (ON state) are different than those in its inactive state (OFF state). In this thesis, an exhaustive analysis of the stability of residues in the active and inactive Ras active site is performed.

La preorganització electrostàtica del centre actiu s'ha postulat com el mecanisme genèric de l'acció dels enzims. Així, alguns residus "estratègics" es disposarien per catalitzar reaccions interaccionant en una forma més forta amb l'estat de transició, baixant d'aquesta manera el valor de l'energia dactivació $\Delta g^{\ddagger}_{cat}$. *S'ha proposat que aquesta preorientació electrostàtica s'hauria de poder mostrar analitzant l'estabilitat electrostàtica de residus individuals en el centre actiu.*

Ras es una proteïna essencial de senyalització i actúa com un interruptor cel.lular. Les característiques estructurals de Ras en el seu estat actiu (ON) són diferents de les que té a l'estat inactiu (OFF). En aquesta tesi es duu a terme una anàlisi exhaustiva de l'estabilitat dels residus del centre actiu de Ras en l'estat actiu i inactiu.

PREFACE

The electrostatic preorganization of the active site has been put forward as the general framework of the mechanisms of action of enzymes (Warshel 1978; Warshel *et al.,* 2001). Several decades ago, Warshel (Warshel 1978) developed a theory on the origin of the catalytic effect of enzymes based solely on electrostatic effects. He realized that the most proficient enzymes catalyze reactions in which there exists a charge separation between reactants (RS) and the transition state (TS). According to this view, enzymes would have evolved to stabilize the TS of the reaction being catalyzed and the different polarity of reactants and products is used by the enzyme to accelerate the rate of the reaction with respect to what occurs in solution. However, water is, in fact, a very polar environment itself, and not much seems to be gained by restricting those types of reactions to occur in a different environment. To understand this apparent paradox, one should consider the fact that in polar solvents, during the transformation from RS to the TS, about half of the energy gained from charge–dipole interaction, $\Delta\Delta G_{q\mu}$, is spent on changing the dipole–dipole interaction, $\Delta\Delta G_{\mu\mu}$, (Warshel 1991; Rouhi 2000; Villà and Warshel 2001). In proteins, however, the active site dipoles associated with polar groups, internal water molecules, and ionized residues are already partially oriented towards the transition state charge center (see Figure 1).

Thus, $\Delta\Delta G_{\mu\mu}$ is smaller in the enzyme than in water, and less free energy is spent on creating the oriented dipoles of the protein transition state. The free energy term $\Delta\Delta G_{\mu\mu}$ is basically the so-called “reorganization energy” felt by the solvent during the process of forming the transition state solute charges. The reorganization energy that water has to pay to “follow” the change in charge distribution along the reaction coordinate from reactants to the transition state has been already “prepaid” by the enzyme during the folding process. The preorganization hypothesis has been corroborated as the main contribution to catalysis in many systems by extensive both theoretical and experimental studies (see a partial list in Bonet *et al.,* 2006; Villà and Warshel 2001).

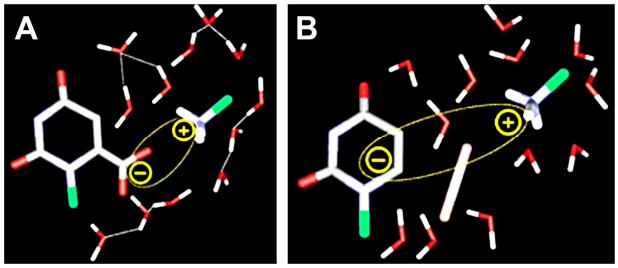


Figure 1: Water reorganization energy
During a charge separation reaction in water (a model for lysine assisted orotidine decarboxylation in the example) the water molecules need to reorganize in order to follow the polarity change from **a**) reactants state to **b**) transition state. This reorganization energy is already partially payed when the reaction occurs in an enzyme, as the folding process has provided a preoriented environment for the reaction to occur, maximizing the electrostatic complementarity between the active site and the TS. To what extend this preorganization can be analyzed by simply studying the stability of the enzyme active site residues? This thesis tries to answer this question by careful analysis of the Ras active site in its different activation states (Figure adapted from Rouhi 2000).

Following this hypothesis, the search for features related to stability in the protein/water environment may help us assigning functional roles to regions/residues in proteins or to improve protein docking protocols by locating electrostatics hot spots on protein surfaces.
Estimations of the residual stabilization in a protein structure have been shown to provide a valuable tool to assess functionality of a given aminoacid residue in its environment. Stabilization in terms of differential free energy of solvation between the residue in water and in the protein environment has been extensively used to evaluate binding contributions in the ground state of receptor-ligand complexes (Muegge *et al.,* 1998), in the transition state of enzymatic reactions (Warshel *et al.,* 2000; Elcock 2001) or even in locating patches for protein-protein interactions (Bonet *et al.,* 2006). In the study of enzyme reactivity, in particular, it has been proposed that the electrostatic preorientation of active site residues should be related with their electrostatic distabilization. In this thesis, an exhaustive analysis of the stability of residues in the ON-state and OFF-state Ras active site is performed. We choose this system because of the wealth of mutational and structural information that is available, due to its biological relevance.

Semi-macroscopic solvation methods have provided useful insights into the electrostatic characteristics of residues in protein structures (Sham *et al.,* 2000; Shurki *et al.,* 2004; Bonet *et al.,* 2006; Scheper *et al.,* 2008). Taking profit of this fact the linear response approximation version of the semi-macroscopic protein dipoles Langevin dipoles method (PDLD/S-LRA) is used here to investigate the stability trends of Ras active site in its different

activation states. A generalization of the use of such method for protein interaction induced activation in phosphate hydrolysis proteins is also given. Results confirm the ability of such semi-discrete methods to capture the essential electrostatic features of amino acid residues in complex environments, even responding to relatively small global movements leading to subtle conformational changes.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Ras proteins in the signaling context

Ras is a founding member of a large superfamily (~150 members in human) of small GTPases. Members of the Ras superfamily proteins act as molecular switches (anchored in the internal leaflet of the plasma membrane) and cycle between GTP (guanosine triphosphate) bound active and GDP (guanosine diphosphate) bound inactive forms and control cell growth, proliferation, apoptosis and differentiation (Bos 1989; Peeper *et al.,* 1997; Shields *et al.,* 2000; Cox and Der 2003; Rodriguez-Viciana *et al.,* 2004; Hancock and Parton 2005; Quatela and Philips 2006). This switching between two alternate conformations (GTP and GDP induced) is associated with several biologically important signaling pathways. During the flip over between active and inactive forms, some parts of the Ras protein (see page 7) exhibit significant conformational changes, which are obvious around amino acids 10-17, 30-38 and 60-76 (Mitin *et al.,* 2005; Campbell *et al.,* 1998). These regions of Ras protein are referred as the P-loop, switch-I and switch-II region respectively. In fact, the P-loop region is a common motif among the phosphate hydrolysis proteins and contains some conserved residues that determine the GTP and ATP binding sites (Via *et al.,* 2000).

To convert inactive Ras to active state, GDP has to be substituted by GTP. This reaction requires guanine nucleotide exchange factors (GEFs) to accelerate the release of tightly bound GDP, and then GDP is subsequently replaced by GTP due to higher intracellular concentration of the latter (see page 16). While in its activate Ras interacts and activates downstream targets, called effectors, which in turn propagate the message and trigger cellular responses. After activation, GTP hydrolysis returns Ras protein back to inactive state, consequently ceasing downstream signaling. The catalytic capability of Ras protein to hydrolyze GTP is intrinsically slow with reaction rate constants of 10^{-1} to $3{\cdot}10^{-3}$ min^{-1} (Scheffzek and Ahmadian 2005). Ras protein catalytic capability can be accelerated up to five fold through interaction with GTPase activating proteins (GAPs) (see page 11), and this way regulating the GTP hydrolysis process (Wittinghofer *et al.,* 1997). Thus, GEFs and GAPs are essential modulators of the biological activity of Ras protein (see Figure 2).

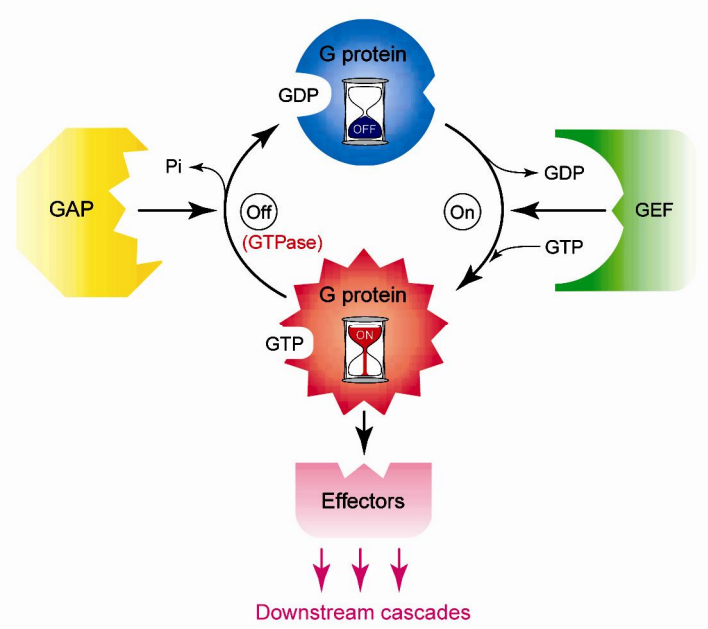


Figure 2: Mechanism of Ras cycling between ON and OFF states
Common principles of mechanism of Ras shuttle between active GTP bound (**Red**) and inactive GDP bound (**Blue**) states. The exchange of guanine nucleotide (from GDP to GTP) is facilitated by guanine nucleotide exchange factor (GEF) (**Green**) turning 'ON' the Ras switch. The activated state is turned 'OFF' by the action of GTPase-activating proteins (GAPs) that trigger the hydrolysis of GTP (**Yellow**). Once activated, Ras interacts with its effectors proteins (**Pink**) and propagates downstream signaling cascades. Adopted from Kosloff and Selinger 2001.

The interest of this thesis is to understand the role of the different residues that form the active site region of Ras and how they are influenced by changes in the activation state of the protein. Thus, a detailed explanation of the primary through tertiary structure of Ras follows.

1.2 Ras family primary structure

The Ras superfamily proteins are broadly classified into six subfamilies, namely Ras, Rho, Arf, Rab, Ran and Rad proteins and these subfamilies are further divided (see Table 1) including over 100 related proteins (Giehl 2005). The N-terminus catalytic domain of Ras superfamily GTPases share a set of five highly conserved G box motifs (Nucleotide binding elements) and abeled as G1, G2, G3, G4, and G5 boxes (Colicelli 2004; Wennerberg *et al.,* 2005) with the following consensus sequences: **G1**, GXXXXGKS/T; **G2**, T; **G3**, DXXGQ/H/T; **G4**, T/NKXD; and **G5**, C/SAK/L/T where 'X' is any

natural aminoacid (Bourne *et al.,* 1991, Wennerberg *et al.,* 2005).

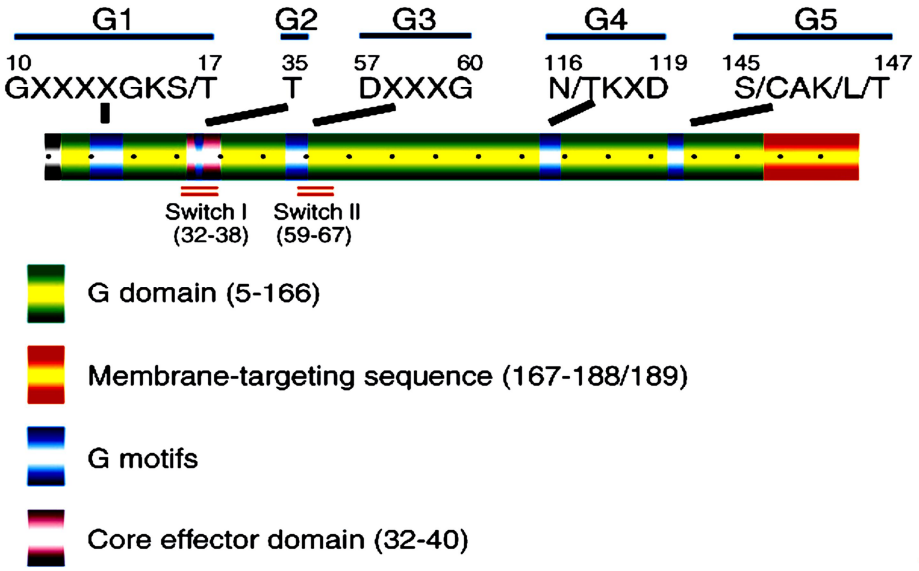


Figure 3: Ras protein functional domain organization
Ras protein conserved G box GDP/GTP-binding elements (Adopted from Wennerberg *et al.*, 2005)

Subfamily	Family members	Cellular function
Ras	H-Ras, K-Ras4A, K-Ras4B, N-Ras M-Ras, R-Ras, Rheb, RalA, RalB Rap1A, Rap1B, Rap2A, Rap2B, TC21, Rin, Rit	Regulation of cell cycle progression Regulation of apoptosis Regulation of exocytosis
Rho	RhoA, RhoB, RhoC, RhoD, RhoE, RhoG, RhoH Rac1, Rac2, Rac3, Cdc42, Rnd1, Rnd2, Rif, CHP, WRCH1, TC10	Regulation of cell cycle progression Regulation of actin cytoskeleton Regulation of apoptosis
Arf	Arf1, Arf2, Arf3, Arf4, Arf5, Arf6	Regulation of vesicle transport Regulation of actin cytoskeleton
Rab	Rab1–Rab63	Regulation of vesicle transport
Ran	Ran	Regulation of nuclear import
Rad	Rad, Gem, Kir, Rem1, Rem2, Ges	Mainly unknown Regulation of cell morphology

Table 1: The Ras superfamily classification
A detailed classification schema of Ras protein superfamily. Ras superfamily is mainly divided into six subfamilies with distinct cellular function. Adopted from Giel 2005.

These G-box motifs are conserved in all Ras super and subfamily proteins (see Figure 3), as well as Gα in other GTPases (Wennerberg *et al.,* 2005).

G1 motif: The Arf subfamily shows the following pattern of sequence conservation ‘GLDxAGKT’ while Rho family displays ‘K C/a VaVGDa V/C GKTxL’ residues. Rab suberfamily (having the largest members than rest of the subfamilies) displays ‘KaaaaGDxxVGK S/T’ pattern of conservation. Ras subfamily displays ‘aaaaGxxGVK S/T’ pattern in the G1 region. This is more similar behaviour that is reflected by the Ras subfamily. Ras subfamily displays ‘LVLVGDGGTGKT’ pattern.

G2 motif: While in the G2 motif only amino acid ‘T’ was found to be conserved. In more detail the Arf subfamily the multiple sequence alignment retrive ‘PtaGF’ conserved residues. While, Rho subfamily displays ‘YaPTV’ sequence pattern. Rab subfamily displays ‘TIG’ pattern with the excemption of Rab20, Rab40 (A, B, and C) and Rab5L. More, Ras subfamily does not show more conserved pattern in this region except conserved ‘T’, where as, Ran subfamily shows ‘YVATLG’ pattern which resembles the Rho and Rab subfamilies.

G3 motif: In the G3-motif the consensus sequence found was ‘DxxGQ’ which lie mainly in the switch-II region of Ras protein and is mainly acidic in nature. Ras subfamily predominately shows the signature motif ‘LxaxDTxG’, the conserved Asp (D) and consensus Thr (T) and Gly (G). Rho subfamily also mimic such pattern in a signature motif like ‘Ar DT A/S/F G/E’. Where Ar means any aromatic residue *e.g.* Phe/Tyr/Trp.
Arf subfamily displays conserved ‘aGG’ residues where ‘a’ is an aliphatic amino acid with an exception of Arl12. The consensus sequences were found to be ‘Ar D’. The overall signature for this subfamily could be represented by ‘Ar DaGGQ’. Ran subfamily signature motif is comprised of ‘WDTAGQE’.

G4 motif: In the G4 motif the signature motif found to be ‘NKxD’. The Ras subfamily residues show ‘aaaGNKxD’ . The Rho subfamily residues show ‘aaGx Amine xD’, where ‘Amine’ means either Glutamine or Lysine. While Ran subfamily shows signature sequences ‘VLCGGNKVD’ which can be rewritten as ‘aaCaGNKVD’. Rab subfamily residues show ‘axGNKxD’ pattern in the G4 Guanine binding region. More, the Arf subfamily shows ‘Lxx A/G NK Q/x’ pattern.

G5 motif: Ras subfamily shows ‘ExSa’ pattern of signature motif in the G5 motif which is a Guanine nucleotide binding region. Rho subfamily shows ‘YxECSa’ pattern of the signature motif. Ran subfamily displays ‘YDISA’ of

the signature motif. Rab subfamily shows ‘ExSaK’ pattern of conservation. Arf subfamily shows no specific pattern of conservation in the G5 motif.

	G1	**G2**	**G3**	**G4**	**G5**
Ras	AaaaGxxGVK S/T	T	LxaxDTxG	aaGNKxD	ExSa
Rho	K C/a VaVGDa V/C GKTxL	YaPTV	Ar DT A/S/F G/E	aaGx	YxECSa
Arf	GLDxAGKT	PtaGF	Ar DaGGQ	Lxx A/G NK Q/x	xxxx
Rab	KaaaaGDxxVGK S/T	TIG	DxxG	N K/R xD	ExSaK
Ran	LVLVGDGGTGKT	YVARLG	WDTAGQE	axGNKxD	YDISA

Table 2: Ras protein concensus domain motifs

The C-terminus region of Ras protein that goes beyond the **G5** motif is termed as hypervariable region (HVR) involved in membrane localization (see **appendix** pp. 92 for details).

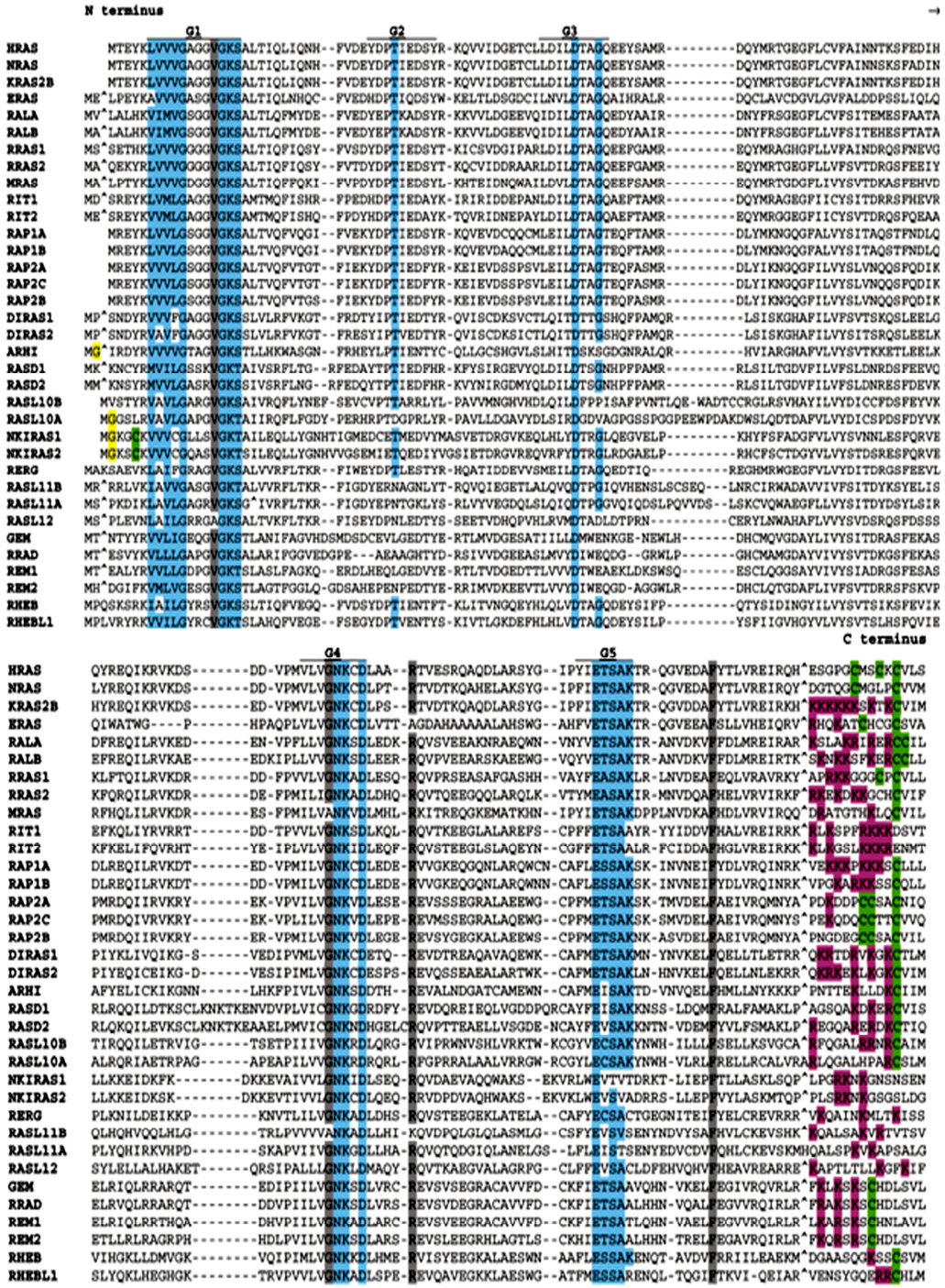


Figure 4: Multiple Sequence Alignment of human RAS subfamily members. G box motifs consensus sequence residues are colored in **Blue**. Cysteine residues (substrates for Posttranslational modification 'Prenylation') in either terminal (C and N) are colored **Green**. The N-terminal Glycine (substrate for Myristoylation) are colored in **Yellow**. The C-terminal basic residues (poly Lys) are colored in **Pink**. The highly conserved residues (in 90% of members in this MSA) are colored **Gray**. Adopted from Colicelli 2004.

1.3 Ras Protein secondary and tertiary structure

1.3.1 GTP and GDP-bound standalone protein structure

Crystal structures of Ras proteins show 5 α–helices and 6 β-strands (*see* Figure 5b) (Pai *et al.,* 1990; Milburn *et al.,* 1990). Several functionally relevant regions can be identified in the Ras protein secondary structure:

- The COOH-terminal region (residues 167 to 186/189) contains the HVR (see Section 1.2) that is involved in the protein's membrane anchorage (for details, see Section 6.1).

- The conserved P-loop (GxxxGKS/T, residues 10-17) (Scheffzek and Ahmadian 2005), a glycine rich motif involved in the stabilization of the substrate phosphate moiety. It is worth noting that the Gly12Val or Gly12Arg mutations, although yield much lower protein activity, do not distort ignificantly the protein structure.

- The two regions, consistently termed switch I (approx. residues 31-38) and switch II (residues 56-76), that interact with the phosphate moiety of the substrate in the GTP-bound form and that are frequently involved in the interactions between the GNBPs and their effectors or/and regulatory proteins.

In addition, the crystal structure of H-Ras protein (Pai *et al.,* 1990) shows nine water molecules within hydrogen bond distance to the nucleotide GppNp (a GTP analog; for simplicity, we will refer to Ras•GppNp and other Ras structures bound to triphosphate analogs as "Ras•GTP" as we proceed along this thesis). Two of such water molecules (Wat175 and Wat189 in Figure 5b) are found close enough to the γ-phosphate to suggest playing a role in the GTP hydrolysis. Indeed, after the appearance of Pai *et al*., structure, the 1990's embraced a heated debate between the idea that such water should be activated by some protein residue and converted into OH^- or the idea that the substrate itself served as the basis for the activation. Such controversy was in principle solved by careful analysis of the GTP pKa, which showed that it had better ability than protein residues to deprotonate a water molecule previous to the nucleophilic attack (see Aqvist *et al.,* 1999 for a summary).

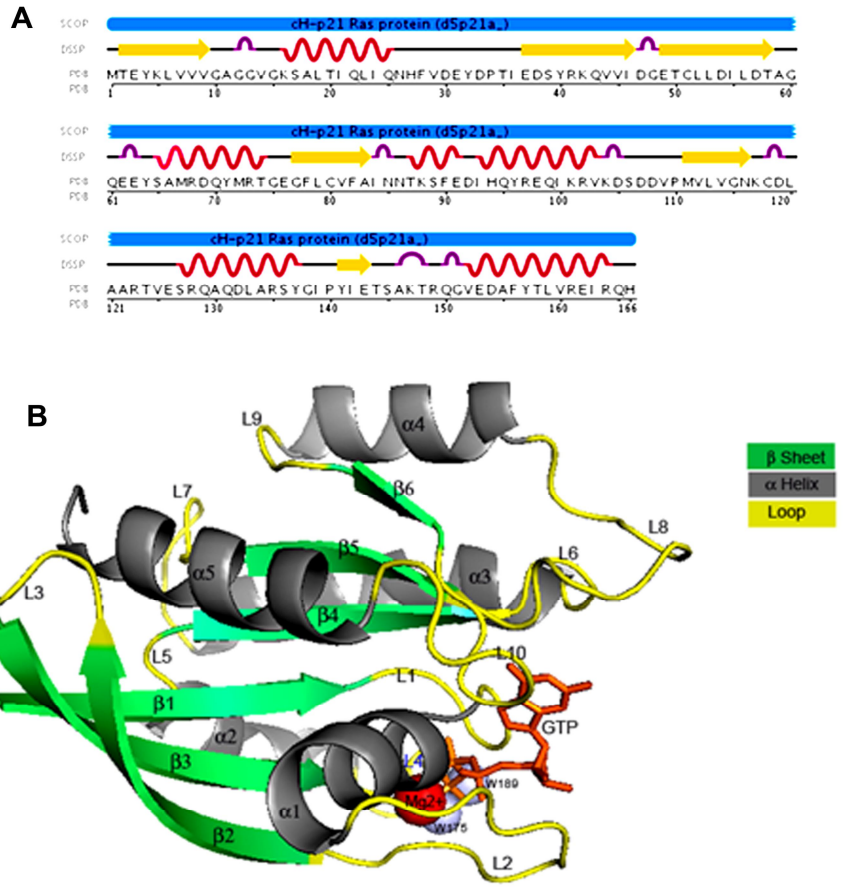


Figure 5 3D structure of GTP bound H-Ras (PDB code 5P21)

a) Ras secondary structure (DSSP information (Kabsch and Sander 1983) obtained directly from the PDB web site Berman *et al.,* 2000; Berman *et al.,* 2003) **b**) Ras protein secondary structure.. β-sheets are shown in **Green**, α-Helices are shown in **Grey**, and Loop structures are shown in **Yellow** color. Water molecules are represented as **Blue** spheres whereas Mg^{2+} is represented as **Red** sphere. The triphosphate analog is shown in **stick**. The numbering of each secondary structural element of H-Ras are as under; β1 (1-9), L1 (10-14), α1 (15-26), L2 (27-37), β2 (38-46), L3 (47-49), β3 (50-58), L4 (59-66), α2 (67-75), L5 (76), β4 (77-84), L6 (85-86), α3 (87-104), L7 (105-109), β5 (110-117), L8 (118-125), α4 (126-137), L9 (138-139), β6 (140-144), L10 (145-150) and α5 (151-171) (Milburn *et al.*, 1990). Structure was generated using PyMol.

In the crystal structure resolved by Milburn *et al.,* (1990), the substrate's α-phosphate is interacting with Ala18 and the β-phosphate is interacting with the amide nitrogen atoms of residues 13 or 14 to 17 (loop-1 and following α1 elements) and the side chain of Lys16 (all residues being part of the **P-loop** motif). Such configuration in which, apart from the Lys16 positive charge, all residues are neutral, seems like a curious way for an enzyme to bind a triphosphate moiety. It is extrange, for example, that the Gly12Val leads to an important decrease in the activity of the protein. However, a close inspection of the structure reveals that all main chian dipoles of the P-loop

motif are pointing towards stabilizing the highly negative substrate, and in this way a small perturbation of the structure (like for example substituting a Gly residue Val) may disrupt the hydrogen bond network (Muegge *et al.,* 1996). In addition, the coordinated Mg^{2+} ion forms an adduct with Ser17, Thr35 and with the β, γ -phosphate oxygen. Furthermore, Milburn's crystal structure also shows that Lys16 is surrounded by the negative charge of carbonyl oxygens of residues 10 and 11, as well as one of the oxygen atoms of the substrate's β, γ -phosphate.

Comparison between Ras protein active and inactive state structures reveals that the positions of the β-phosphate and the Mg^{2+} ion in the GDP complex appear to be shifted by a magnitude of 0.4Å from the corresponding GTP-analog complex. On the contrary, the structural conformation of the ribose moiety and the guanosine base are similar both in the active and inactive Ras protein crystal structures. Similarly, by superposition of the P-loop in the GTP-analog complex of Ras_{wt} protein with the same motif in the GDP complex of oncogenic (Ras_{G12V}) protein, Milburn *et al.,* (1990) found that position 12 contributes to the stability of the γ-phosphate. They showed that the side chain of Val12 in Ras_{G12V} may partially block the entrance of the guanine nucleotide pocket thus preventing the entrance of the attacking nucleophile or Pi after hydrolysis (or both). With respect to Gly13, which binds to the nucleotide's β-phosphate, it was shown that Ras_{G13V} or Ras_{G12D} mutations resulted in distortions of the P-loop structure.

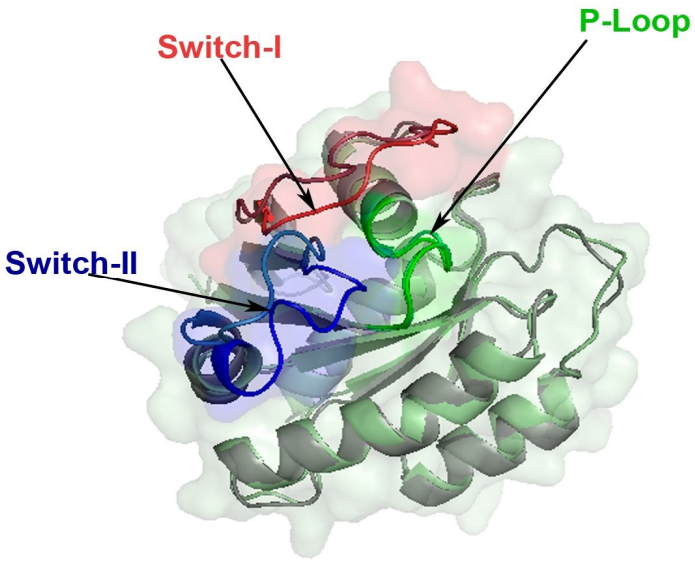


Figure 6: Ras protein ligand induced conformational changes
Detail of the structural superposition of GTP- and GDP-bound Ras structures, showing the most important conformational changes occurring in the active site region

As shown in Figure 6 the largest differences between the active and inactive state structures are localized in the two switch regions. In the H-Ras crystal structure, **switch-I** region conformation "switches" when there is a substitution of GDP by GTP substrate with Cα RMSD values between 2.6 to 2.7Å (Milburn *et al.,* 1990). The structural differences that are associated with the flipping over of this switch region are; (**i**) no hydrogen bond exist between the side chains of Tyr32 and Tyr40 in the active state structure, differing from what is observed in the inactive state; (**ii**) Thr35 is coordinated in the active state to the Mg^{2+} ion and the γ –phosphate, which were absent in the inactive state; and (**iii**) the orientations of two switch-I region residues (36 and 38) in the active state conformation are significantly different than the inactive state conformation. The Ras_{T35A} mutant (a partial loss-of-function mutation) was found to be insensitive to the interaction with the activating protein GAP (see next section) because the residue 35 side chain cannot interact, in the mutant protein, with Mg^{2+} or the γ-phosphate and hence cannot induce the GAP required conformational changes in the switch-I region. Substituting Thr35 by Ser (conserving the hydroxyl side chain), shows partial sensitivity towards GAP. Thus, one may conclude that methyl group of Thr35 is responsible for the dynamic behavior of the switch-I region (Spoerner *et al.,* 2001).

H-Ras **switch-II** region conformation also "switches" when there is a GDP to GTP substitution (see Figure 6). The switch-II region is necessary for GDP/GTP exchange, either by enhancing the dissociation of bound nucleotide and subsequent GTP binding (because of higher intracellular concentration) or by directly participating in the exchange process (Gln61). The region's nine residues (60 to 68) show the largest conformational changes between the two structures with RMSD values ranging between 3.6Å to 7.4Å. H-Ras residues 60 and 61 are highly conserved among all Ras and Ras related proteins while Glu62 and Glu63 are conserved in small nucleotide binding proteins and Ras proteins, respectively. Residues 60 and 61 interaction with the γ-phosphate induces a conformational change in the entire loop-4 and reorients the helix α2.

Since the conformation of switch-I and switch-II regions differ significantly in the active and inactive state, it was already assumed in the original publication that the two switch regions (switch-I and switch-II) are conformationally "linked" to the γ-phosphate and facilitate the H-Ras protein in effector recognition.

As for the guanine base, it is interacting with the side chains of residues 28, 116, 117, and 119, and the backbone of residue 146, the ribose moiety is interacting with the side chain of residue 117 and the backbone of residue 29. There is no significant conformational change for these residues between the Ras active and inactive states (see Figure 6).

Finally, as extensively shown in the appendix, the C-termini of Ras proteins are posttranslationally modified by isoprenylation and farensylation, by which means the protein is anchored to the cytoplasmic membrane. The crystal structures for inactive and active Ras suggest that loops 3, 5, 7, and 9 and helices α2 and α5 are near the membrane and that some of them may be potential sites for interaction with membrane-bound upstream or downstream (or both) regulators.

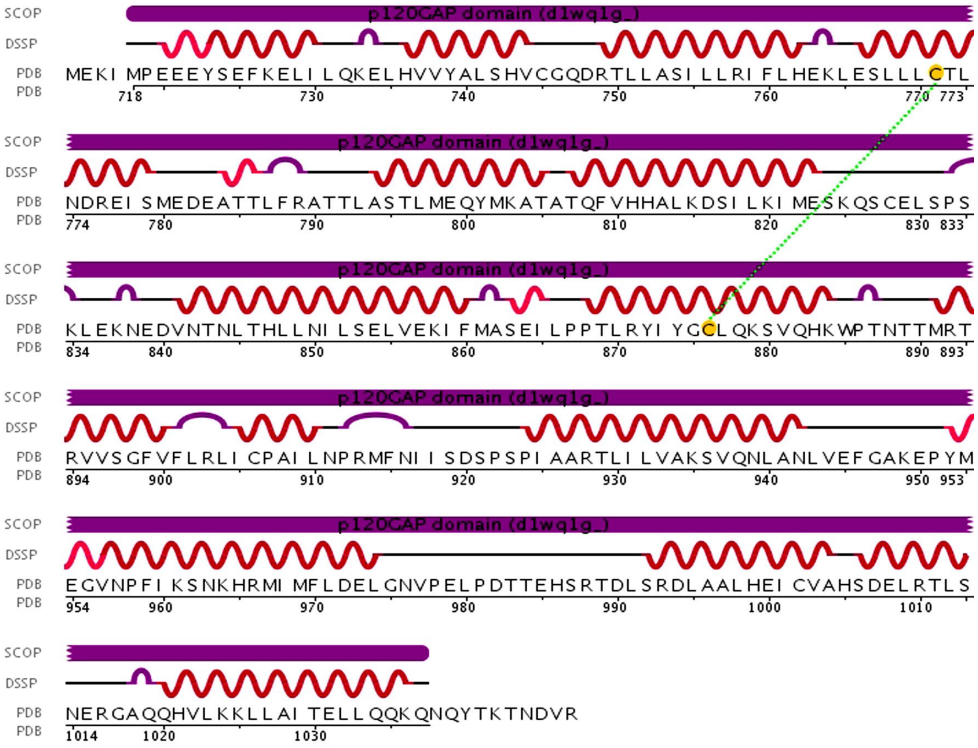
1.3.2 Ras interaction with GTPase Activating Proteins (GAP)

The intrinsic GTPase activity of Ras is too low to promote formation of inactive Ras-GDP and a regulatory protein generally called GTPase activating protein (GAP) is needed for optimal hydrolysis of GTP to GDP. The activation free energy for the hydrolysis of GTP in water is $\approx$ 117 kJ mol^{-1} (27.96 kcal mol^{-1}), whereas the intrinsic GTPase reaction of Ras reduces only to $\approx$ 92 kJ/mol kJ mol^{-1} (21.97 kcal mol^{-1}) (Kötting *et al.,* 2008). This reduction in activation free energy by Ras is mainly due to electrostatic interactions (Kötting and Gerwert 2004). In addition, the Ras-induced charge shift in GTP lowers the activation enthalpy from 105 kJ mol^{-1} (25 kcal mol^{-1}) to $\approx$ 83 kJ mol^{-1} (19.8 kcal mol^{-1}). Thus, catalysis due to the Ras can be seen as mostly enthalpic. The additional term related to the interaction with GAP has been suggested to be mostly entropic (Kötting *et al.,* 2008), although the effect of the Arg789 finger has been demonstrated to be purely enthalpic and to account for a big part of the increase in reactivity (see Scheffzek *et al.,* 1997; Glennon *et al.,* 2000).

The intrinsic GTPase reaction of Ras is slow, with 4.7x10^{-4} s^{-1} at 37°C, but RasGAP catalyzes the reaction $\approx$10^{5}-fold to 19 s^{-1} (measured at 25°C at a pH of 7.4) (Schweins *et al.,* 1995; Kötting *et al.,* 2008). In fact, mutations in hotspots either in RasGAP or in Ras proteins lead to malfunctioning of the catalytic process and ultimately results in cancerous cells. It has been shown that Ras protein interactions with their regulatory proteins (GEF and GAP) are associated with conformational changes affecting the active site (Jeffrey *et al.,* 1995; Scheffzek *et al.,* 1997). Making use of the similarity

between the structure of AlF_3 with the transition state inversion of the departing terminal phosphate during phosphate hydrolysis, Scheffzek *et al.,* (1997) solved the crystal structure for the RasGAP/Ras complex, (PDB code: 1WQ1). The structure, comprising residues 718 to 1037 of GAP-334, residues 1 to 166 of H-Ras, GDP, AlF_3, Mg^{2+}, and 35 water molecules, shows AlF_3 oriented such that the three fluoride atoms contact Mg^{2+}, Thr35 (NH), Lys16 (NH_3^+), Gly60 (NH), Gln61 (NH_2), and Arg789 (the so-called "arginine finger" provided by most GAP proteins). The interface of Ras•RasGAP complex consists of weak *van der Waals* interactions but stronger polar interactions including several bridging water molecules (Scheffzek *et al.,* 1997 and references cited therein) (see Figure 7). Activated Ras binds in the shallow groove of the central catalytic domain GAP_c of GAP-334, with the tip of its triangular shape penetrating most deeply into the groove. The overall structures of GAP-334 and Ras were reportedly similar to their unbound structures (Cα atoms RMSD difference were 0.9 Å for GAP-334 and 0.54 Å for Ras with comparison to corresponding individual structures) except switch-II region of Ras and the loop $L6_c$ of GAP-334. P-loop, switch-I, switch-II regions and, possibly, helix α3 (residues 87 to 98) of Ras protein participate in the interaction with RasGAP. Conversely, RasGAP helices $\alpha6_c$ and $\alpha7_c$ and loops $L1_c$ and $L6_c$ are involved in the interaction with Ras.

A



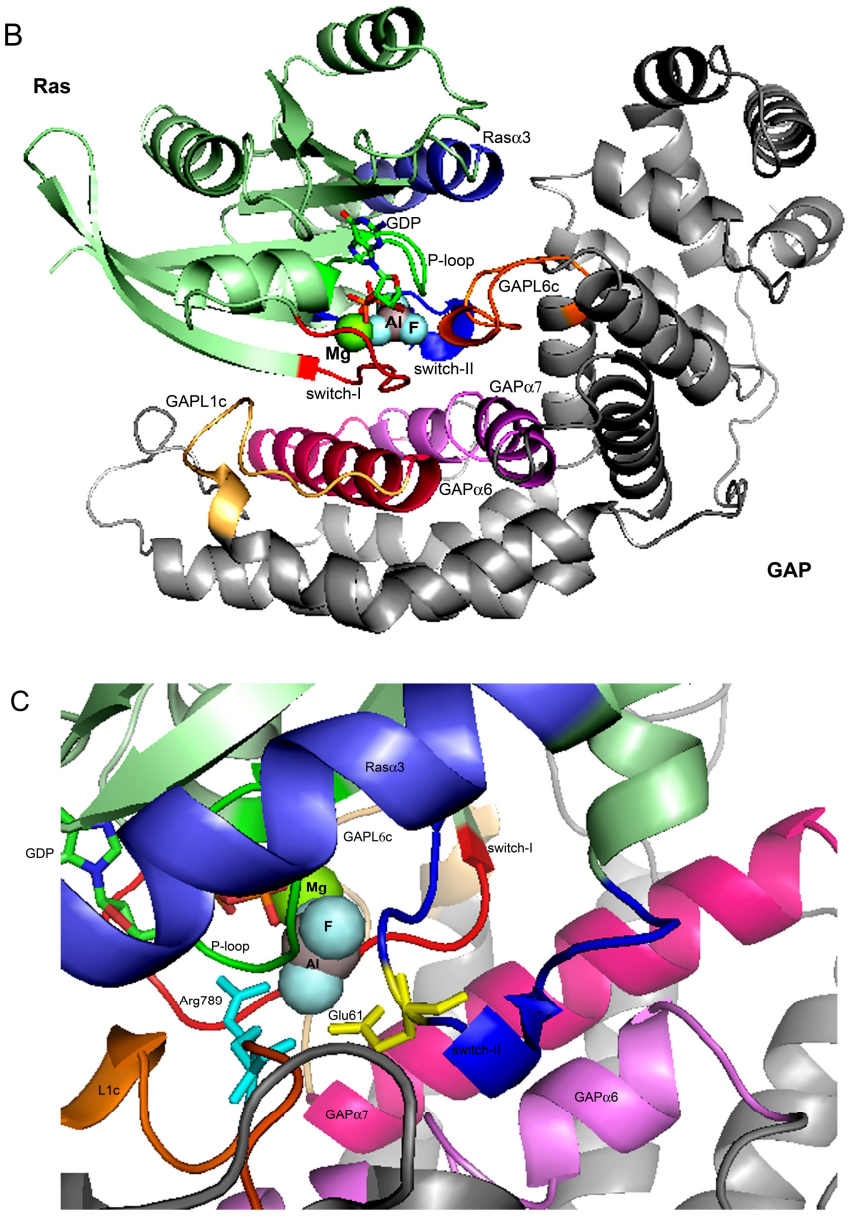


Figure 7: Ras/RasGAP complex

a) RasGAP secondary (DSSP information (Kabsch and Sander 1983) obtained directly from the PDB web site Berman *et al.*, 2000; Berman *et al.*, 2003) **b**) Ras/RasGAP complex ribbon representation. The Ras protein is shown in pale green and the GAP protein is shown in light grey color. Ras protein P-loop region is shown in deep green, switch-I in red, switch-II in blue, Ras Helixα3 in smudge blue colors. Ligand GDP is shown in stick. Aluminum (brown) and Magnesium (Green) metal ions are shown as spheres. Flouride ions are shown as cyan sphere. GAP functionally important regions L1c also called as Finger loop (light orange), L6c also called as variable loop (orange), Helixα6 (Hot pink) and Helixα7 (Violet) are shown. **c)** Details of the interaction of Ras with RasGAP are shown. Coloring schema is same as in **b**. The functionally important residues Arg789 of GAP (cyan color) and Glu61 of Ras (yellow) are shown in stick model of representation. These structures are made using RasGAP 3D structure (PDB: 1WQ1, Scheffzek *et al.*, 1997). Structure was generated using PyMol.

The RasGAP Loop-1 ($L1_c$) is also called “finger loop” because it points into the active site of ras protein and $L6_c$ is called the “variable loop” because its length is variable in the structure-based alignments (Scheffzek *et al.,* 1997 and references cited therein). Ras protein switch-I region is negatively charged and interacts with effector proteins and RasGAP (GAP-334) (Nassar *et al.,* 1995). In particular, Ras switch-I residues interact with the GAP helices $\alpha 6_c$ and $\alpha 7_c$ and with the variable loop. On the other hand, The crystal structure from Scheffzek *et al.,* (1997) emphasizes the importance of GAP_{K949} and GAP_{E950} (both $L6_c$ region residues) in the Ras•RasGAP complex formation. Thus, for example, GAP_{K949} (from the $L6_c$ region) points into this negatively charged surface of Ras switch-I region.

In addition to the polar interactions, hydrophobic interactions are formed between GAP_{L902}, GAP_{L910} and Ras_{P34}, Ras_{I36}, Ras_{Y32}, and Ras_{Y64}. These interactions are presumably the target for inhibitory lipids such as phosphatidic acid and arachidonic acid. The arachidonic acid inhibits the Ras•RasGAP interaction in a partly competitive mode with respect to Ras•GTP (Sermon *et al.,* 1996).

$L1_c$ residue Arg789, also called as ‘‘arginine finger’’, is the key residue for GAP-mediated catalysis (Ahmadian *et al.,* 1997). The x-ray structure of the Ras•GDP•AlF3•RasGAP complex mimics the transition state structure for the GAP mediated catalysis and shows arginine finger pointing to the GTP binding site of Ras. GAP_{R789} stabilizes the position of the catalytic Ras_{Q61} and neutralizes the negative charge on the β, γ phosphates that are developed during the phosphate hydrolysis process (Scheffzek *et al.,* 1997; Glennon *et al.,* 2000). The catalytic arginine finger is a highly conserved motif among several GAP proteins for small G proteins (Kosloff and Selinger 2003), although it has been shown to be missing in some systems. For example, p85 showed how a subunit in PI-3K is homologous to GAP proteins (including the Arg finger), but does not act as a GAP (Fidyk and Cerione 2002; Scheffzek and Ahmadian 2005). While on the other hand, it was shown that mutation of Arg74 to R74K and R74M in (E. *coli* strain L7/L12) ribosomal protein does not cause a major change in GTPase activity for EF-G and conclude that arginine residue is not essential for the activation, excluding an "arginine finger"-type mechanism (Savelsbergh *et al.,* 2002; Scheffzek and Ahmadian 2005). Finally, and perhaps more strikingly, Seewald and colleagues in RanGAP ternary complex (Ran, Guanine nucleotide and RBD of RanBP1) found no arginine in the active site vicinity (Seewald *et al.,* 2002; Scheffzek and Ahmadian 2005). The positive

charge of the arginine guanidinium group is delocalized (Heesen *et al.,* 2007). Since, coulombic interactions drive the Arg789 towards the negatively charged GTP. Hence, Arg789 alone is not capable to induce specific charge distribution in GTP, although it plays a major role (Glennon *et al.,* 2000; Heesen *et al.,* 2007).

GAP specific function is to stabilize the orientation of Ras_{Q61} by a complex network of hydrogen bonds (see Figure 7). Ras_{Q61} is essential for GTP hydrolysis, its mutation to any other amino acid (except Glu) blocks Ras-mediated GTP hydrolysis (Polakis and McCormick 1993; Bollag and McCormick 1991). Ras_{Q61} position a water molecule for in-line nucleophilic attack on the γ-phosphate. The arginine finger forms a hydrogen bond with Ras_{Q61} (Kötting *et al.,* 2008). In addition, the RasGAP crystal structure suggests an indirect role of Arg903 in catalysis.. RasGAP crystal structure also reveals that Gln61 contributes to the stability of the GTPase reaction transition state.

After hydrolysis, the protein-bound P_i intermediate is still in the ON conformation. Then the movement of switch-I is reversed to OFF and Arg789 moves out of the binding pocket. Thus an efficient GAP should not only catalyze the bond breakage but also the release of the phosphate (Kötting *et al.,* 2008).

An interesting alternative view of the entropic contributions of RasGAP to catalysis is given by Heesen *et al.* (2007). The arginine finger movement changes active site dielectric by pushing out waters from the binding pocket. This results in an increase in the Ras-induced charge shift on the GTP moiety. In principle any residue that can move into Ras protein binding pocket and push the waters out may be able to catalyze the Ras GTPase reaction (Kötting *et al.,* 2008), which is in agreement with Seewald et al 2002(talks of RanGAP Asp225 and Lys130 of Ran and Q61 is equivalent of Q69). Thus, the loss of protein-bound structured water is most likely the source for a positive activation entropy $\Delta S^{\ddagger}$. The cost of entropic change by relocating one protein bound water; into the bulk water is up to 8 kJ mol^{-1} (1.91 kcal mol^{-1}). In their MD simulations study, Heesen *et al.,* (2007) showed that the coordination of the arginine finger with the GTP displaces five water molecules, and thus would provide an overall entropy change of up to 40 kJ mol^{-1} equivalent to 9.5 kcal mol^{-1} (very close to the experimental value of 39 kJ mol^{-1} equivalent to 9.1 kcal mol^{-1}). Kötting *et al.,* (2008) suggests that this decrease of entropy after arginine interaction may be

counter balanced by the structured waters formerly coordinating the arginine. The free energy gained by displacing the water molecules may be used to induce the electrostatic interactions to stabilize the transition state for bond breakage.

1.3.3 Ras interaction with Guanine nucleotide Exchange Factor (GEF)

We have seen above how Ras proteins are tightly controlled by dynamic switching between the GDP- and GTP-bound activity dependent conformational states. Guanine nucleotide exchange factors (GEFs) stimulate the intrinsic GTP/GDP exchange activity of Ras and promote formation of active Ras-GTP. There are three main classes of RasGEFs: son of sevenless protein (SOS), Ras-GRF, and RasGRP (Wennerberg and Der 2004). In brief, all RasGEFs have a common ~250 amino acid CDC25 homology catalytic domain (also called the RasGEF domain) and an adjacent ~50 amino acid N-terminal Ras Exchange Motif (abbreviated as REM and also called the RasGEF N-domain) see Figure 8 (Margarit et al 2003; Mitin *et al.,* 2005). The interface between the REM domain and the helical hairpin of the CDC25 domain is known to be important for SOS activity (Hall *et al.,* 2001).

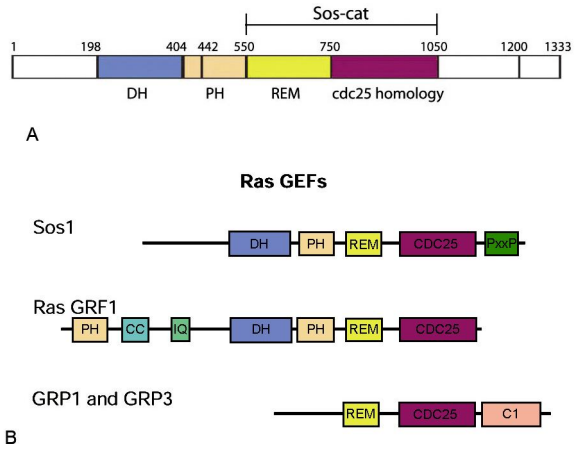
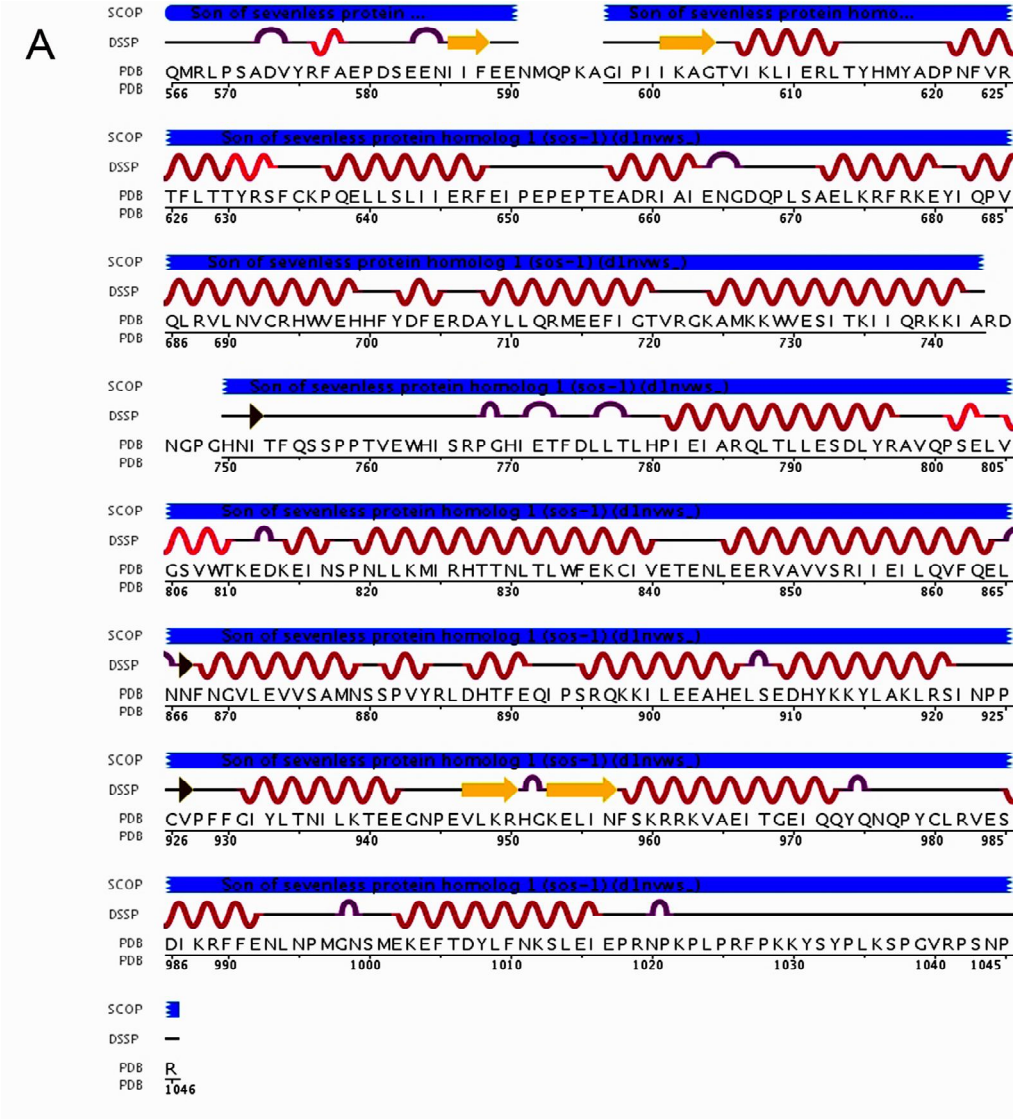


Figure 8: Domain organization of SOS
a) Complete domain organization of Human SOS1 protein **b)** A comparative analysis of different Ras specific GEFs.

Ras-GTP binds with high affinity to distal (REM) and allosteric site (CDC25 homology domain), which is a SOS1 regulatory region (Margarit *et al.,* 2003) (see). Binding of Ras-GTP has a stabilizing impact on active state conformation of SOS, whereas binding of Ras-GDP to this SOS1 regulatory

site results in a lower SOS activity (Sondermann *et al.,* 2004). The crystal structure of SOS$^{DH\text{-}PH\text{-}cat}$ resolved by Sondermann *et al.,* (2004) reveals that the allosteric Ras binding site in SOS is blocked by the DH domain, which is located far from the active site of SOS. It appears that the DH domain has no direct inhibitory effect on the Ras-specific nucleotide exchange reaction but rather functions as a gate for Ras binding to the allosteric site.

The crystal structure of Ras-GEF (PDB code: 1NVV) shows Ras residues Pro34, Ile36, Glu37, Met67, Tyr64, and SOS residues Leu687, Met726, Trp729, Arg688, Asn691, His695 to be relevant for the interaction (see).



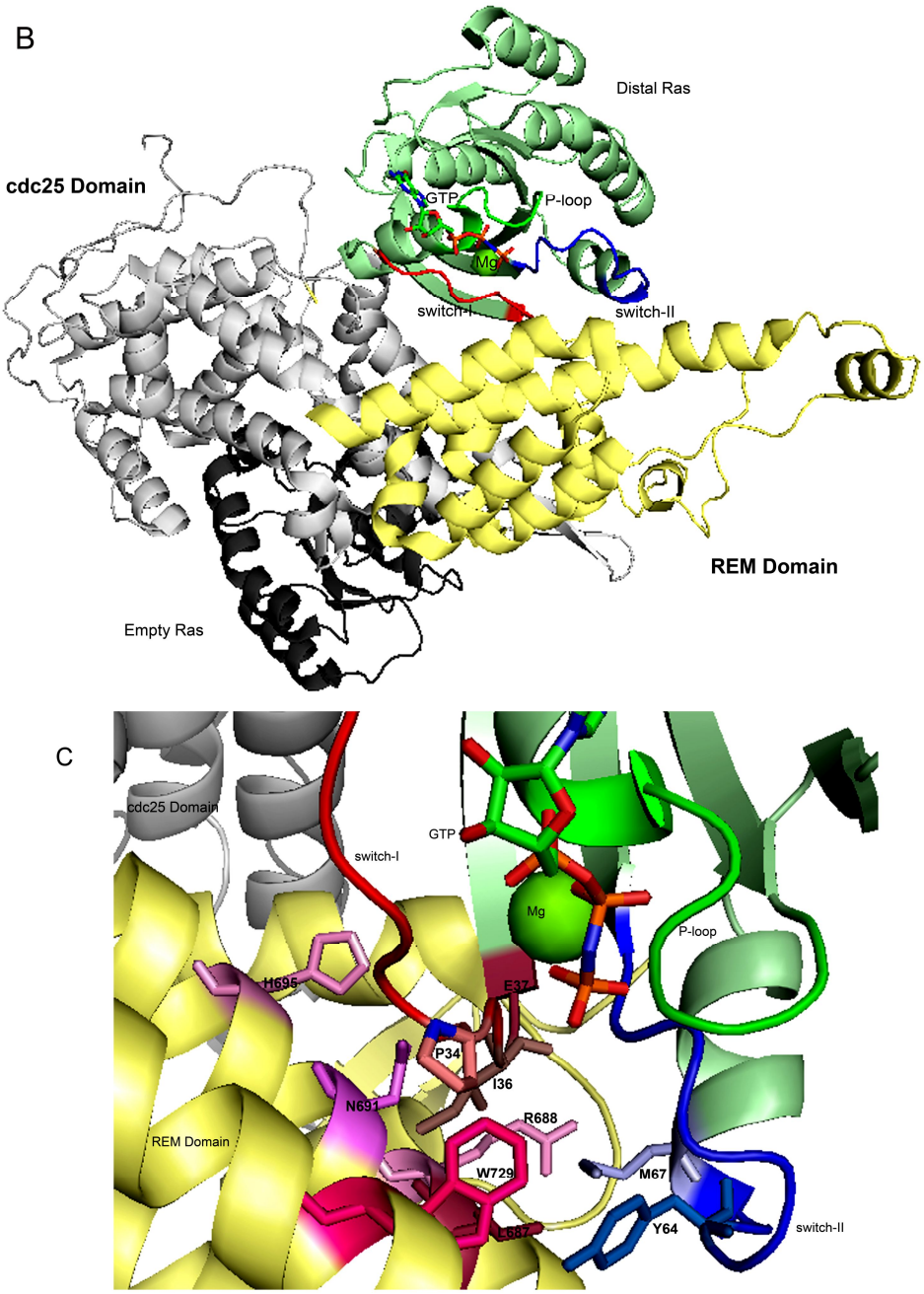


Figure 9: Figure showing details of Ras/GEF

a) RasGEF (SOS cat domain) secondary (DSSP information (Kabsch and Sander 1983) obtained directly from the PDB web site Berman *et al.*, 2000; Berman et al., 2003) **b**) Ras/GEF complex ribbon representation. The Ras protein (pale green) and two domains of GEF protein (cdc25 domain in light grey and the REM domain in light yellow) are shown. Ras protein P-loop region is shown in deep green, switch-I in red, switch-II in blue colors. Ligand GTP is shown in stick. Magnesium metal ions (Green) is shown as sphere. Empty Ras protein is shown as dark grey color. **c**) Details of the interaction of Ras with GEF are shown. Coloring schema is same as in **b**. The functionally important residues of Ras protein P34, I36, E37, Y64 and M67 are shown. Similarly, important residues of GEF (L687,R688,N691,H695,W729) are shown in the stick model of representation. These structures are made using Ras/GEF/Ras 3D structure (PDB: 1NVW, Margarit *et al.*, 2003). Structure was generated using PyMol.

In particular, Figure 9b shows how residues Pro34 and Ile36 of the switch-I region of Ras pack against hydrophobic side chains in SOS (Leu687, Met726, and Trp729) and Met67 of switch-II region of Ras packs against a hydrophobic crevice formed by Leu687 (which is also part of the binding site for switch-I) and Arg688. Tyr64 of switch-II is also located in this region. A central aspect of the Ras-GEF interface is the formation of a tight bidentate hydrogen bonding interaction between Glu37 and Arg688 (see Figure 9c). Glu37 also forms hydrogen bonds with the side chains of Asn691 and His695 (Margarit *et al*., 03).

1.3.4 Ras interaction with Effectors

Both switch regions in Ras are involved in interactions with downstream effector proteins. All Ras effectors share a region of ~100 amino acids known as the Ras-binding domain (RBD). RBDs in all Ras effectors reported so far share a ubiquitin fold (Herrmann *et al.,* 2004). Despite their structural similarity, the binding affinity between the different RBDs and Ras varies among the different members of the subfamily. This observed difference provides an evidence of differential amino acid effects on the Ras-effector interface, stressing the understanding of the energetics of the interaction. Moreover, the RBDs are reported to be important but not sufficient in order to explain the activation of the downstream cascade.

1.4 Ras structure-function relationships

1.4.1 GTP hydrolysis mechanism in solution and in enzyme

Ras and other small GTP-binding proteins cycle dynamically between an inactive GDP-bound state and an active GTP-bound state. GEFs stimulate the intrinsic GTP/GDP exchange activity enabling GTP loading of the protein and promoting the formation of an active state Ras-GTP complex (see page 1). Once activated, these proteins interact with downstream effectors and propagate the signaling response. We have also seen how the intrinsically slow GTP hydrolysis in Ras is potently augmented by GAPs that increase the reaction rate by several orders of magnitude (see Section 1.1.2). The exact nature of GTP hydrolysis has been the target of a heated scientific debate (Schweins *et al.,* 1994; Warshel and Florian 1998; Glenon *et al.,* 2000; Du *et al.,* 2004; Topol *et al.,* 2004; Grigorenko *et al.,* 2005). It is still controversial whether the transition state of the GTPase reaction is dissociative or associative. The dissociative reaction path for the transition

reaction involves a metaphosphate-like PO_3^{2-} state, in which upon nucleophile arrival the β, γ -anhydride bond breakage occurs. Contrarily, in the associative mechanism, a SN2-like reaction is followed, making a bond between the nucleophile and the γ-phosphate the γ–β phosphates cleavage. Whatever is the proposed mechanism, a nucleophile has to be created by deprotonating a water molecule. What is the general base for such deprotonation became the center for the scientific argument in the 1990's. Biochemical studies of wild and mutant Ras proteins have helped state that the phosphate hydrolysis in Ras works via a substrate assisted catalysis (SAC) mechanism (see Figure 10) where a proton transfers from nucleophilic water to the γ-phosphate of GTP, favoring an associative reaction path (Schweins *et al.,* 1995; Pasqualato and Cherfils 2005). In addition, Wittinghofer (2006), analyzing on different crystal structures of Ras and Ras-like proteins, found that Ras-like proteins structures in the GTP bound state, show a consensus water molecule ~3.5Å from the γ-phosphate in a position ready for in-line attack on it.

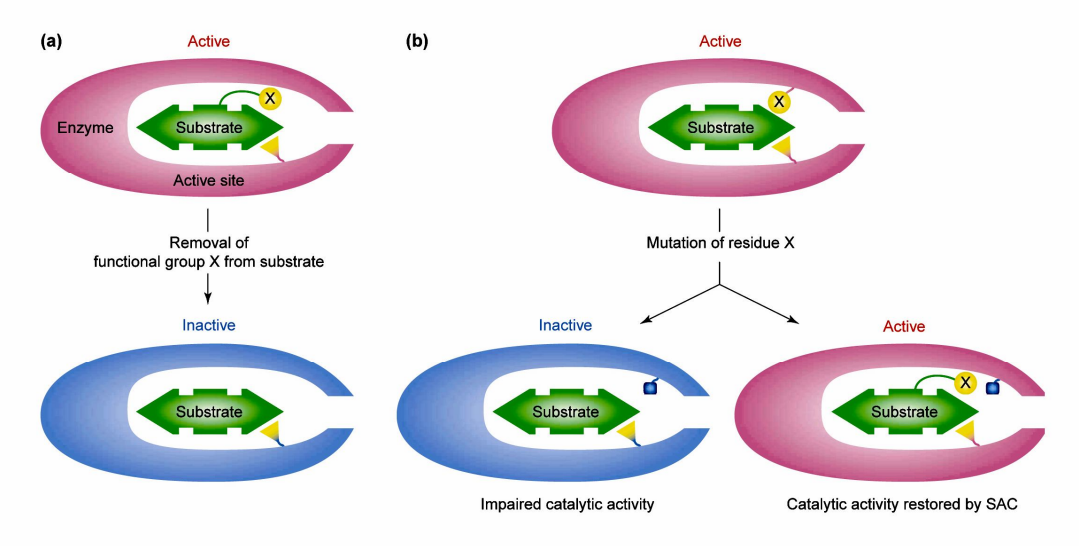


Figure 10: Figure showing two types of substrate assisted catalysis (SAC) Natural and Engineered
a) Natural SAC: In some enzymatic reactions the catalytic functional group (in this figure marked as X) is part of the substrate itself. Any change in the substrate that removes or changes this functional group results in the impairment of the catalytic activity. **b**) Engineered SAC. Taking profit of the above effect, a functional group that may have been removed from the active site by means of some mutation, can be restored into the catalytic site by the substrate itself. Adopted from Kosloff and Selinger 2001.

But as discussed above, this is not the end of the story and many other labs based on their computational (pKa shifts, binding affinity or activation free energy calculations) and experimental (FTIR, real time atomic resolution

techniques) studies reached conflicting conclusions (*i.e.* dissociative transition state) (Cepus *et al.,* 1998; Du *et al,.* 2000; Allin *et al.,* 2001; Du *et al,.* 2004).

In the two families of G proteins (Heterotrimeric G-proteins and Ras-like monomeric proteins) two residues, glutamine (Gln_{cat}) and arginine (Arg_{cat}), are of vital interest because both are directly involved in catalysis. Both residues are conserved throughout the G protein family and relevant GAPs and their side chain orientation with respect to substrate is highly similar (Sprang 1997; Kosloff and Selinger 2001). In heterotrimeric G proteins, both residues originate in *cis,* while in Ras proteins with crystal structures it became apparent that Ras proteins lack Arg_{cat}. This dogma was resolved when RasGAP protein was crystallized and it was found that in Ras proteins the Arg_{cat} is supplied in *trans* by the GAP proteins (the Arg_{cat} is called 'arginine finger', as seen in previous sections) (Scheffzek *et al.,* 1997; Rittinger *et al.,* 1997; Albert *et al.,* 1999; Wittinghofer 2006).

1.4.2 Stability and function

As explained in the preface of this work, various theoretical studies of the energetics of enzymatic reactions and the corresponding reactions in aqueous solutions indicate that electrostatic stabilization is the most important factor in enzymatic catalysis (Warshel 1978). This provided the initial hint towards establishing the relationship between stability and function. This is, electrostatic preorganization can be related to the instability of the residues that are important for catalysis. Elcock (2001), showed how functionally important residues in proteins are charged residues that are located in electrostatically unfavorable environments. In another work, the stability profile is added to the information on position conservation and shape in order to discriminate between catalytic and non-catalytic residues (Ota *et al.,* 2003). Recently Dessailly *et al.,* (2007) also used energetic criteria to identify functional sites in protein.

These studies served to establish a central hypothesis for the computational component of this PhD thesis. Namely, if we were able to design a simple but robust protocol to analyze residue stability in a protein environment we could provide a medium- to high-throughput method to analyze series of proteins in search of functional regions. To reach this ambitious goal, the linear response approximation version of the semi-macroscopic protein dipoles Langevin dipoles method (PDLD/S-LRA) was adopted to evaluate the solvation free energies for individual residues. The rest of the sections in

this manuscript include a brief overview of the PDLD/S-LRA method and the description some statistical tests used to assess the significance of the solvation free energy values. Following this methods section, a results section is presented, in which we mainly discuss the values obtained for relevant residues in the active site of the Ras protein in different conformational states. In addition, we include a brief subsection expanding the method to study a more complex problem, the preorganization of protein interfaces in protein protein interactions of phosphate hydrolysis enzymes.
For completeness on the discussions on the Ras structures, although out of the scope of this thesis, the conclusions section is followed by an appendix containing an additional review on the posttranslational modifications leading to different Ras protein localizations.

2 Objectives

The general aim of this thesis is double:

1. To characterize the active site of Ras by getting the interaction energy fingerprint in different states (GTP/GDP and bound/unbound to partners)

2. To find a general protocol to produce medium- to high throughput structure-function correlations in proteins

These main objectives unfold into several more precise milestones:

1. On the ras active site fingerprint:

    a. To analyze the level of distabilization of different key residues in the Ras active site between the GDP- and GTP-bound systems

    b. To assess the effect of interacting partners in the Ras active site stabilization energies

2. On the generalization of the PDLD/S-LRA based method

    a. To set up a protocol suitable for fast check of all relevant residues in a single protein

    b. To compare stability values in sets of residues for different proteins using a proper statistical test

    c. To extend the method to study hot spots in protein-protein interfaces

3 METHODOLOGY

3.1 Structure selection and classification

3.1.1 Sequence Selection and Multiple Sequence Alignments

We took HRas (PDB Code: 5P21) as our query sequence. We searched the NCBI-NLM 'Protein Blast' Non-redundant database (Altschul *et al.,* 1990; Altschul *et al.,* 1997) against our query HRAS protein sequence and selected only Human Ras-subfamily member proteins (Giel *et al.,*2005). In order to explore the pattern of sequence conservation within the Ras-subfamily, we took the obtained protein sequences in fasta format and used 'MUSCLE' (Edgar 2004) for the multiple sequence alignment (MSA). MUSCLE aligns sequences using a pair-wise algorithm that through a progressive multiple alignment provides a distance matrix. This procedure is iterated until the tree stabilizes or until a specified maximum number of iterations is reached.

3.1.2 Protein Classification and Active Site Delimitation

We selected 3D structures of all Ras related Ras-subfamily proteins from the Protein Data Bank (Berman *et al.,* 2000; Berman *et al.,* 2003). We classified the selected structures in a 2x2 table (see Figure 11), labeling them as follows:

- Group A includes unbound proteins with a triphosphate ligand
- Group B includes unbound proteins with a diphosphate ligand
- Group C contains the GTP-bound structures in complex with a partner (regulators, effectors…)
- Similarly, Group D contains GDP-bound proteins in complex with partners.

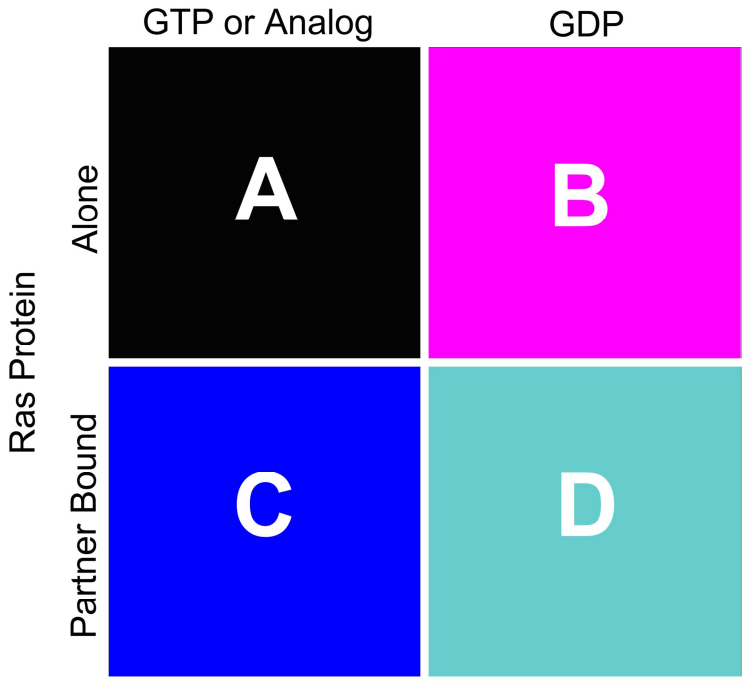


Figure 11: Classification of Ras structures used in this study.

Group-A has following PDBs (5p21, 1qra, 121p, 221p, 421p, 521p, 621p, 721p, 821p, 6q21, 1lf0, 1jai, 1iaq, 1gnr, 1ctq, 1gnq, 1plj, 1p2s, 1agp, 1p2u, 1jah, 1p2t, 1p2v, 1plk, 1gnp, 2cl0, 2cl6, 2cl7, 2clc, 2evw, 1xcm, 1x1s, 1zw6, 2pmx, 2rga, 2rgb, 2rgc) Group-B PBDs (1q21, 2q21, 4q21, 1ioz, 2cld, 2ce2, 1x1r, 1xj0, 1zvq, 1crr, 1lf5, 1pll, 1crq, 1crp, 1aa9, 2quz) , Group-C with following PDBs (2C5L, 1lfd, 1he8, 1nvv, 1nvw, 1nvu, 1nvx, 1k8r, 1wq1), and Group-D with PDB (1xd2).

One of the main goals of the thesis is to relate the stability of individual residues in the active site of Ras to their function in different states of activation of the protein. Thus, the first step after having collected the PDB structures corresponding to Ras and Ras with its effectors is to determine which residues are forming part of the active site region. In lack of a precise biological criteria we defined the residues to be included in this study by selecting all residues in a sphere of 12.5Å sphere (to be consistent with typical electrostatic cutoffs in MOLARIS version alpha9.06.01) (Chu *et al.,* 2003) around each substrate heavy atom in the PDB structure 5P21.

3.2 Evaluating the stability of single residues by solvation calculations

In this work we focus on the calculation of the electrostatic stabilization of a given residue by the protein environment. Electrostatic stability is a vague concept that we formalize by assigning to each residue a free energy value that describes the energy gained or lost by the residue when moving it from

the solution environment to its protein environment. In a way, we are interested in moving “bulky” charges and dipoles from one environment to another. To get such free energy value one could use in principle any method that allows the calculation of solvation energies. In particular, we evaluate the solvation energy by means of the linear response approximation version of the semi-macroscopic protein dipoles Langevin dipoles (PDLD/S-LRA) method (Sham *et al.,* 2000). The PDLD/S-LRA method evaluates the electrostatic group contribution by averaging the effective potential of transferring a group from solvation to protein (Lee *et al.,* 1993; Sham *et al.,* 2000; Schutz and Warshel 2001). The method is described elsewhere with great detail, but in order to understand the implications its use have in our work we will briefly introduce the main points here.

3.2.1 The PDLD solvation model

It is an increasingly accepted fact that electrostatic energy is one of the most important structure-function correlators for biomolecules. It is far from trivial, though, obtaining reliable estimations of such energy. One may in principle use one of the following approaches:

1. A fully microscopic model which represents explicitly all the solvent and/or proteins atoms. This option is extremely attractive because one may rely on solid developments in statistical mechanics to produce free energy values using rigorous formulations. However, the amount of computer time needed to explore efficiently the potential energy surface is huge and even the arrival of new hardware (Cell processor, GPU graphical cards) and specialized algorithmics it is not possible to reach converged results when strong environment reorientation occurs. This is the typical situation when studying charge separations in proteins or in general processes between different charging schemas. In addition, the definition of the potential energy surface itself is typically problematic.

2. One can represent the solvent molecules by dipoles that would account for the main physics of the solute-solvent interaction.

3. The solvent region can be divided into relatively large volume elements and the average polarization, as well as the average field, of each volume element can be treated within the continuum approximation, which typically involves the use of a macroscopic

dielectric constant.

In this work we use a combination of the three approaches, as the PDLD/S-LRA method includes the use of a discrete representation of protein and solvent to generate configurations that are later treated using a dipole-based solvation model. The continuum approach is reserved for the description of the bulk solvent properties and how this affects the inner system. In particular, the protein dipoles Langevin dipoles (PDLD) approach is a simplified microscopic approach that retains the clear physics of the microscopic world, where one does not have to assume an arbitrary dielectric constant (the dielectric is just the vacuum dielectric constant). However, the solvent is drastically simplified and the time average polarization of each solvent molecule is represented by a Langevin type dipole.

$$\left(\vec{\mu}_i^{Langevin}\right)^{n+1} = \vec{e}_i^{\,n}\mu_0\left(\coth\chi_i^n - \frac{1}{\chi_i^n}\right) \qquad \text{Eq. 3-1}$$

$$\chi_i^n = \frac{C\mu_0}{k\ T}\left|\xi_i^n\right| \qquad \text{Eq. 3-2}$$

Where $\vec{e}_i^{\,n}$ is a unit vector in the direction of the local field $\vec{\xi}_i^{\,n}$, C is a parameter, and μ_0 is taken as 1.8D. The equation is solved iteratively at each given Langevin dipole site. A useful alternative is provided by assuming a linear polarization law that replaces the Langevin equation

$$\left(\vec{\mu}_i^{Langevin}\right)^{n+1} = \alpha\xi_i^n \qquad \text{Eq. 3-3}$$

Although this equation can, in principle, lead to over polarization of the solvent molecules near highly charged sites, it has been found to behave in a stable way and to converge faster than Eq. 3-1. The PDLD model does not try to reproduce the exact position of the solvent atoms but places the solvent dipoles on a simple spherical grid. The microscopic results of the PDLD model might involve convergence difficulties which should require extensive averaging over the dipoles configurations. We will introduce in the following sections how to do such averaging.

3.2.2 Molecular mechanics and molecular dynamics

Computer modeling of macromolecules is based on using a mathematical description of the dependence of their energy on their structure, the so called potential energy surface (PES). Its definition rigorously comes from the use of quantum mechanics. However, to answer a number of relevant

questions in biochemistry, a description in terms of classical interaction between fragments of the macromolecules (atoms bonded to other atoms, residues interacting with other residues…) suffices. Thus one can describe large molecules as a collection of small molecular fragments where the overall potential surface is expressed as a sum of contributions from bonded atoms and interactions between nonbonded atoms. Such a representation is usually done by sets of analytical functions that present an approximation to the true potential surface and are called potential functions or force fields. The functional forms and parameters of molecular force fields are taken from studies of small molecules with the implicit assumption that these functions are transferable from small to large molecules.
Molecule potential functions are usually given in the form

$$U(\vec{r}) = \underbrace{U_{bond}(\vec{r}) + U_{bend}(\vec{r}) + U_{tors}(\vec{r})}_{} + \underbrace{U_{vdW}(\vec{r}) + U_{elect}(\vec{r})}_{} \quad \text{Eq. 3-4}$$

Where $\vec{r}$ represents the set of Cartesian coordinates of the whole system. The description of the functional forms for the different terms in Eq. 3-4 can be found in many sources and is not given here. For a historical perspective see Levitt 2001.
Here we simply mention that the non-bonded terms can be described by an atom-atom interaction potential of the form

$$U_{elect}(\vec{r}) = \frac{1}{4\pi\varepsilon_0\varepsilon_r}\frac{q_i q_j}{r_{ij}}$$

Where the dielectric constant ε_r represents the factor for the screening of electrostatic interactions due to the presence of a dielectric with respect to what should happen in vacuum and r_{ij} is the distance.

Simulation of macromolecules must reflect solvent effects which are not just small perturbations but major contributors to the overall energetics and force. In fact, modeling of macromolecules in a vacuum is quite irrelevant as much as the behavior of such molecules in solution and proteins is concerned. Molecular Dynamics (MD) simulations evaluate the motion of the atoms in a given system and provide the positions or *trajectory* of these atoms as a function of time. The trajectories are calculated by solving the classical equation of motion for the molecule under consideration. One obtains the relevant forces on each atom from the first derivatives of the given potential functions. The actual evaluation of classical trajectories is done numerically,

expressing the changes in coordinates and velocities at a time increment, δt, by

$$r_i(t+\Delta t) = r_i(t) + \dot{r}_i(t)\Delta t$$
$$\dot{r}_i(t+\Delta t) = \dot{r}_i(t) + \ddot{r}_i(t)\Delta t = \dot{r}_i(t) - \frac{1}{m}\frac{\partial U(t)}{\partial r_i}dt \qquad \text{Eq. 3-5}$$

where the dot designates a time derivative and we use Newton's law:

$$m\ddot{r}_i = F_i = -\frac{\partial U}{\partial r_i}$$

Starting with a given set of initial conditions for the position and velocity vectors, $r_i(t=0)$ and $\dot{r}_i(t=0)$, one can evaluate the trajectory by numerical integration of Eq. 3-5. There exist several variants for the equation to be integrated that are more efficient and accurate than Eq. 3-5 (see, *e.g.* Allen 1987) but the essentials of MD are included in the above minimal set of equations. If we consider a global system (ensemble) with a constant total number of particles, a constant volume and integrate the above equations we are running simulations in the so called "microcanonical (or NVE) ensemble", with constant energy determined by the initial conditions.

However, in order to simulate a given macroscopic property at a specified temperature we must introduce some type of "thermostat" in the system that will keep it at the given temperature. In this way we will be working on the so called "canonical (or NVT) ensemble". One may profit from the fact that the kinetic energy of a system of N particles is related to the temperature by means of $E_k = \frac{3}{2}Nk_BT$ in order to force T to be a given value through the simulation. Again, there exist several variants for such thermostat, but in this work we use a simple scaling of velocities in order to adjust T at each step of the integration. The mean squared displacement provides a means to establish how the simulation is evolving:

$$1\,\Delta r^2(t) = \frac{1}{N}\sum_{i=1}^{N}[r_i(t) - r_i(0)]^2$$

For a fluid without structure, the RMSD (MSD of the square root) increases with time. MD simulations allow one to sample the potential energy surface of a given macromolecular system and from this we can perform statistics and, in principle, obtain macroscopic quantities comparable with experimental data. In fact, such sampling is nearly impossible to be made

complete in real systems with a large number of degrees of freedom and some strategies need to be done in order to accelerate the convergence of our results. In the next section we discuss how to sample the PES for for a process using free energy perturbations method as an introduction to the linear response approximation method employed in the PDLD/S-LRA method.

3.2.3 Free energy perturbations

Absolute value of A (the Helmholtz free energy, the "natural" quantity in the canonical, NVT, ensemble) is difficult to get, but obtaining its relative value is and easier task:

$$\Delta A = A_Y - A_X = -k_B T \ln \frac{Q_Y}{Q_X}$$

where Q_X is the so called partition function of state X and can be obtained from the sampled data from, let's say, a long enough MD run on the region of the PES that describes X. Zwanzig in 1954, developed a better way to evaluate this quantity (Valleau and Torrie 1977):

$$\Delta A = -k_B T \ln \left\rangle \exp\left[\left(U_Y - U_X \right) / k_B T \right] \right\langle_X \qquad \text{Eq. 3-6}$$

based on running simulations on the state X and evaluating, on each of the sampled points, the energy of the system by the expressions describing the PES for both X and Y states. Most times X and Y do not overlap in phase space and thus the evaluation of the above average is difficult (see Figure 12a). However, as A is a thermodynamical state function, which implies we can use any path to go from one state to the other with identical result.

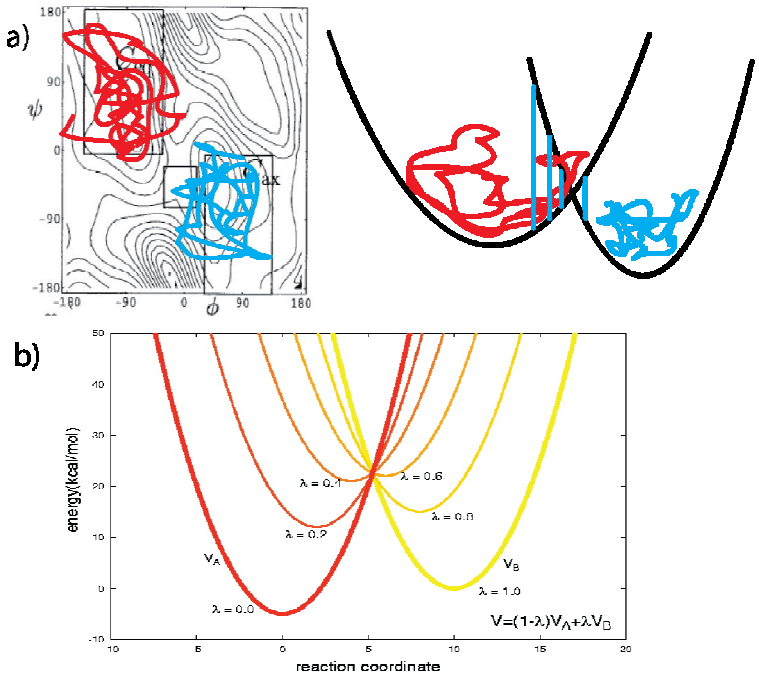


Figure 12 Free energy perturbations.
Zwanzig's formulation allows one to obtain a free energy change between two different states by evaluating the energy of a given sample point on one state by the PES functions associated with both states (vertical blue lines in the rihght hand side picture. However, the overlap between the trajectories sampled in two different states is not always optimal (left hand side picture). B) The FEP method allows on to create "alchemical" intermediate states and to use them to increase the overlaps needed for the Zwanzig formula to succeed in providing a converged ΔA.

Thus, we can envisage an imaginary path that slowly leads from X to Y: To efficiently implement a Free Energy Perturbation method, we need to move the systems from X to Y in small steps (even if they have no real physico-chemical meaning!) through a mapping potential:

$$V_m = \lambda_m V_Y + (1 - \lambda_m) V_X$$

where λ_m goes from 0 to 1. This expression allows us to create small perturbations to the initial state PES in our way to reach Y. Thus, we are able to use Zwanzig's formula for "alchemical" surfaces that have better overlap. The final free energy difference is then evaluated by the sum of all individual free energy differences between consecutive mapping potential states. It must be emphasized here that the convergence of FEP approaches is quite slow and that obtaining meaningful results requires proper treatment of long-range electrostatic effects.

3.2.4 The Linear Response Approximation (LRA)

FEP calculations of solvation free energies or related properties for large

cofactors or ligands are expected to converge extremely slowly.
A promising strategy may be provided by exploiting the fact that electrostatic effects in solutions (and probably proteins) seem to follow the linear response approximation (LRA). From Figure 13 it easily follows that:

$$\Delta A_{X\to Y} = \frac{1}{2}\left[\left\langle U_b - U_a \right\rangle_a + \left\langle U_b - U_a \right\rangle_b\right]$$

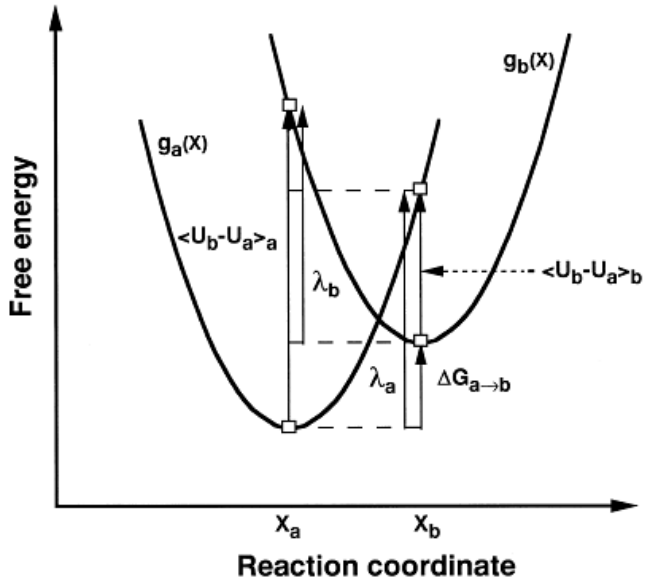


Figure 13 Linear response approximation.
Schema for the graphical interpretation of the linear response approximation formula. For two states represented by parabolic PES, the free energy difference may be easily related to the average differences between the state energies at the minima of the two (taken from Sham *et al* 2000).

This equation has been found to provide a powerful way of estimating the results of FEP calculations of large solutes and drugs. Although rigorously the Helmholtz free energy difference ΔA is the one directly obtained from NVT simulations, and the Gibbs free energy difference ΔG should be obtained from NPT simulations, we can assume that the error introduced by this approximation is negligible with respect to the rest of error sources we face. So, from this point we will just use the Gibbs free energy in our notation, for clarity.

3.2.5 PDLD/S-LRA

In order to obtain the solvation energy values that will be associated with the stability of each individual residue in the protein we need first to establish a proper thermodynamic cycle of the process we want to study. In our case, Figure 15 is used. In the cycle, the free energy difference of solvation between the water and protein system can be evaluated by the expression:

$$\left(\Delta\Delta G_{sol}^{w\to p}\right)_i = \left(\Delta G_{self}^{p} - \Delta G_{self}^{w}\right)_i + \sum_{j\neq i}\Delta G_{ij}^{p}$$
$$= \left(\Delta G_{q\mu}^{p} + \Delta G_{q\alpha}^{p} + \Delta G_{qw}^{p} - \Delta G_{self}^{w}\right)_i + \Delta G_{qq}^{i} \qquad \text{Eq. 3-7}$$

where ΔG_{self}^{x} is the self-energy associated with charging the ith group in its specific environment. In the case of a charge in a protein we decompose _this value into the interaction between the charge and its surrounding permanent dipoles_and induced dipoles, _as well with the water molecules in and around the protein. ΔG_{ij}^{p} is the free energy of interaction between the ith and jth ionized residues. We can envisage the above equation as the sum of the loss of "solvation" energy associated with removing the charge from water_plus the "solvation" of the charge by its surrounding protein environment (the protein dipoles and water molecules) and finally the interaction between the charge and the ionized groups

The picture of the solvation energy just described involves the use of a single protein structure for the evaluation of the solvationm energy. Of course, we are interested in having a proper average of solvation energies obtained from a set of conformational states of the protein and ligand. This is accomplished by the use of the LRA method, in what would be a PDLD-LRA method.

In addition, the above formula involves adding big values with different sign, which involves having a very large source of error in the calculations. On the other hand, macroscopic models can lead to more precise (though not necessarily more accurate) results since they implicitly *assume* the above compensation by using a large dielectric constant (see Schutz *et al.,* 2001 for discussion). Considering the high precision of the macroscopic approaches, it is useful to find a way to scale the large energy contributions of the PDLD model, and to obtain similar precision while retaining the clear energy-based concepts of the microscopic approach. Such a "scaled microscopic" model, which is referred to here as the semi-microscopic or PDLD model and is designated by the abbreviation PDLD/S (Sham *et al.,* 2000).

In the PDLD/S method (see Figure 15) the solvation energy is obtained by

$$\Delta U_{pdld/s,i}^{w\to p} \cong [\Delta\Delta G_p^w(q_i = 0 \to q_i = \bar{q}_i) - \Delta G_{q,i}^w](\frac{1}{\varepsilon_p} - \frac{1}{\varepsilon_w}) + \frac{\Delta V_{q\mu}^p(q_i = \bar{q}_i)}{\varepsilon_p} \qquad \text{Eq. 3-8}$$

Where $\Delta\Delta G_p^w(q_i = 0 \to q_i = \bar{q}_i)$ represents the change in the microscopic (all-atom) solvation energy of the entire protein plus its bound ionizable group

upon changing the charge of this group from zero to its actual charge $\bar{q}$. Note that we use the term for energy (U) instead of free energy (G) here, to emphasize we still need to do their proper average by making use of the LRA approach described above.

$$\Delta G^{w\to p}_{pdld/s,i} = \frac{1}{2}[\langle \Delta U^{w\to p}_{pdld/s,i}\rangle_{q_i=0} + \langle \Delta U^{w\to p}_{pdld/s,i}\rangle_{q_i=\bar{q}_i}] \qquad \text{Eq. 3-9}$$

A strong discussion can be found in the literature about the actual value for the dielectric constant ε_p to use in the PDLD/S expression. The answer to such question lies in the understanding that ε_p represents all the factors not included in an explicit way into our MD simulations. Thus, if one would consider all possible microscopic behavior (from water penetration to induced dipoles treatment through a consistent quantum mechanical approach) during the sampling procedure the value of ε_p would be 1, the gas phase value. As this is not the case, depending on the simulation performed a reasonable value for ε_p is 2 to 6. In this work we have used the value ε_p=4, that has provided good results in previous works (Sham *et al.,* 2000; Schutz *et al.,* 2001; Bonet *et al.,* 2006). We have used 5 configurations of PDLD/S-LRA with step at 300°K each were run, This protocol was followed for all 71 residues and 63-structures (total of 4391 residues). For each residue whose solvation was evaluated the following protocol was used. In all simulations a sphere of water molecules with a radius of 16Å was located centered in the residue whose solvation was evaluated (both including or not the protein environment) The surface constrained all-atom solvent (SCAAS) protocol (already implemented in Molaris) (Warshel and King 1985) was used to restrict the water molecules within the boundary of the sphere to behave as in an infinite solvent. The local reaction field (LRF) method (Lee and Warshel 1992) was used to treat long range electrostatic effects. Figure 14 shows an schema of the way PDLD calculations are done in the protein environment. A similar schema (but simple, as we lack protein atoms) is carried out for the evaluation of the solvation free energy in water.

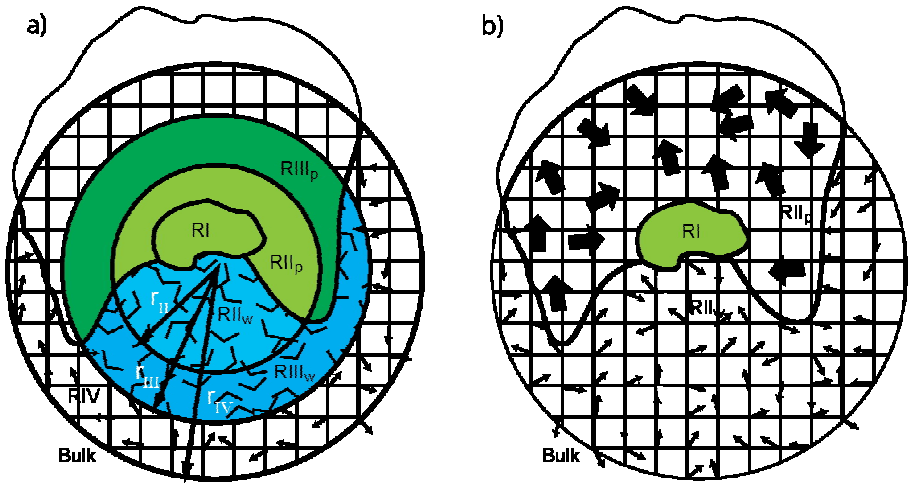


Figure 14: Schema of regions in the PDLD/S-LRA calculations
a) MD is used to explore the conformational space of a sphere of atoms (including explicit water) around the region of interest (in our case a given protein residue whose stability we try to analyze). **b**) for selected snapshots during the MD run (typically 5-10 separated by several picoseconds of simulation) we take the current protein conformation and evaluate the PDLD free energy of solvation for the system in RI. Note that the SCAAS method for spherical boundary conditions already includes a layer of Langevin dipoles surrounding the inner sphere of explicit water molecules (taken from Villà-Freixa, J. Biomolecular Simulations CRC Press in preparation).

After a heating protocol to relax the structure, a total of five production simulations of 10 ps each were performed for the residue in solution and the residue in the protein environment (5000 steps and 2.0-fs time), in each case in the charged and the uncharged states.

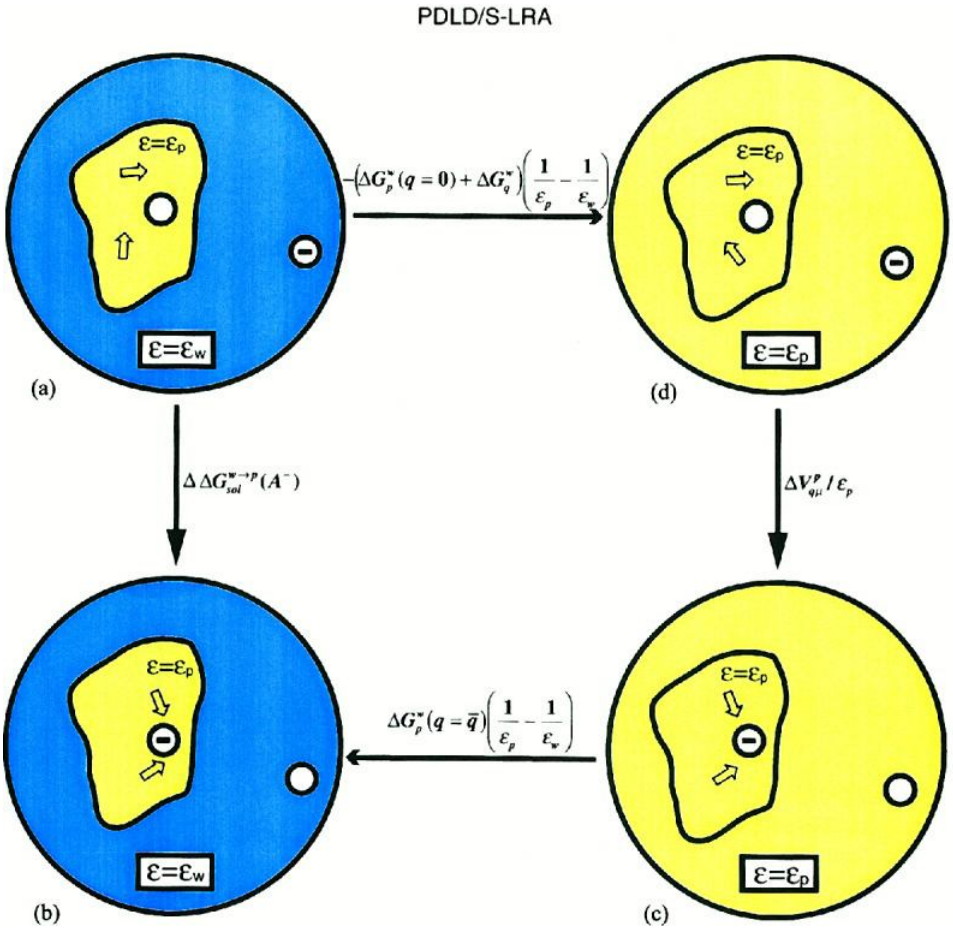


Figure 15: The PDLD/S-LRA thermodynamic cycle for the evaluation of solvation free energy
The cycle involves the change of the dielectric constant around the protein, moving the charge from solvent to protein environment, changing back to the solvent dielectric and uncharging the ionized group inside the protein. The energy contribution of each step is indicated (see the text in figure). The LRA process is designated by the reorientation of the protein permanent dipoles. Adopted from Schutz and Warshel (2001)

3.3 Statistical Tests

We have evaluated the stability values for a big number of residues and we have performed several averages between functional states of the Ras protein. Thus, it is obvious that we need a proper statistical test to assess if the differences we find in different calculations correspond to a real effect or are simply due to a random process. We have used two different scores, a parametric vs. a non-parametric one. While parametric tests assume an underlying probability distribution and, thus, are more accurate only if the data really follows such distribution, non-parametric tests are more robust and can be applied to a wider range of situations without any underlying assumption for the probability distribution.

3.3.1 Z-score Test

Given two sets of results, representing data averaged over two conformational states of the protein, we used the parametric 'Z statistics' in order to test the null hypothesis (H_0) $H_O$that any difference between the two state, based on the nature of ligand is due to random variation or chance. While alternate hypothesis (H_A) state that the difference in the Z-score is due to the nature of the ligand- or partner protein induced conformational change.

The statisctics for the Z-scoe test reads:

$$Z = \frac{x - \mu}{\sigma} \qquad \text{Eq. 3-10}$$

Where we assume a normal distribution of the data. Using 95% probability that if the mean of the PDLD/S-LRA value for our samples are so low that the chance of obtaining that value from a sample selected at random from the population is less than 5 percent, we'll reject the null hypothesis in favor of the alternate hypothesis and conclude that there is evidence to support the null hypothesis that there is no effect of ligand or/partner protein on the thermodynamic properties of the protein.

3.3.2 Wilcoxon Rank test

The Wilcoxon test is a nonparametric test that compares the median distribution of two paired groups. It is also called the Wilcoxon matched pairs signed ranks test. The Wilcoxon test tests the null hypothesis that scores are distributed symmetrically around a specified center *C*. It is here applied to gain or difference scores. Thus, in our case we deal with data obtained for a given residue in two different conformational states of the protein and we use the Wilcoxon test to test the null hypothesis that the gain scores are distributed symmetrically around zero--a result that would be expected if the conformational change had no effect. Thus, the test calculates the difference between each set of pairs, and analyzes that list of differences. The Wilcoxon test analyzes only the differences between the paired measurements for each subject. the statistics to be analyzed by the Wiilcoxon test reads;

$$Z = \frac{W_1 - W_e}{\sigma_w} \qquad \textbf{Eq. 3-11}$$

We classified those structures based on an ex-ante criteria *i.e.* nature of ligand, into two groups (either Group-A if the ligand is in Triphosphate or

Group-B if the ligand in Diphosphate bound state) and the nature of association (A-extended and B-extended if the HRAS protein is attached to partner protein with the same criteria for the ligand as discussed earlier). We preceded further with the solvation free energy calculations that are discussed elsewhere in the methodology section, analyzing each position in A vs. B and A-ext. vs. B-ext groups

4 RESULTS AND DISCUSSION

4.1 Introduction

The generic goal of this thesis is to unravel the energetic fingerprint of the Ras active site. Thus, the first task is to report a detailed analysis of the electrostatic characterization of the active site residues of the protein, by studying different clusters of conformational states manually curated from the PDB. The clusters are built taking into account the functional state of the structure, thus distinguishing between those bound to GTP or GDP and those interacting with different partners (see Figure 11). We base our analysis on the stability of the active site residues in the protein environment with respect to solution for the different structures. Then we perform a statistical analysis to classify them according to their electrostatic characteristics. The results appear to give a consistent picture of the stability pattern of the active site regions in all systems, while the use of parametric and non-parametric statistical analysis yields a list of residues that significantly explain the differences between the clusters of structures and how these residues are strongly correlated with mutations responsible for different types of Ras-related diseases. To do so we performed a detailed analysis of the literature and compare our results with those found by others. Our results demonstrate the ability of semimacroscopic methods to explain the electrostatic features for this family of enzymes and provide a clue for further simplifications of the protocol to be used in the medium to high-throughput search of functional residues in proteins with a rigorous but efficient physico-chemical framework.

A final section of this chapter includes the results obtained using an analogous method in which we focused on the extension of the semimacroscopic calculations and the stability hypothesis applicability to unravel hot spots in protein interfaces for complexes including a phosphate hydrolysis enzyme as a partner.

4.2 Analyzing conservation in the Ras family of proteins

In order to analyze the conservation trends of the functional characteristics in the family of Ras proteins, we generated a multiple sequence alignment of the Ras sub-family members in the Human genome. We started by performing a Blast search with the default parameters in the NCBI server for

our query sentence extracted from the structure of H-Ras by Pai *et al* (PDB code 5P21).

After getting the sequences we performed a multiple alignment calculation with program MUSCLE (Edgar 2004), based on the physico-chemical properties of the active site amino-acids. Thus, in addition to the specific amino acid conservation we also looked at the side chain chemistry.

The multiple alignment shows, among other results, the already known highly conserved residues of the P-loop region (around L1) as well as the less conserved switch I and switch II regions. We will discuss in more detail the conservation results from this figure in the next sections.

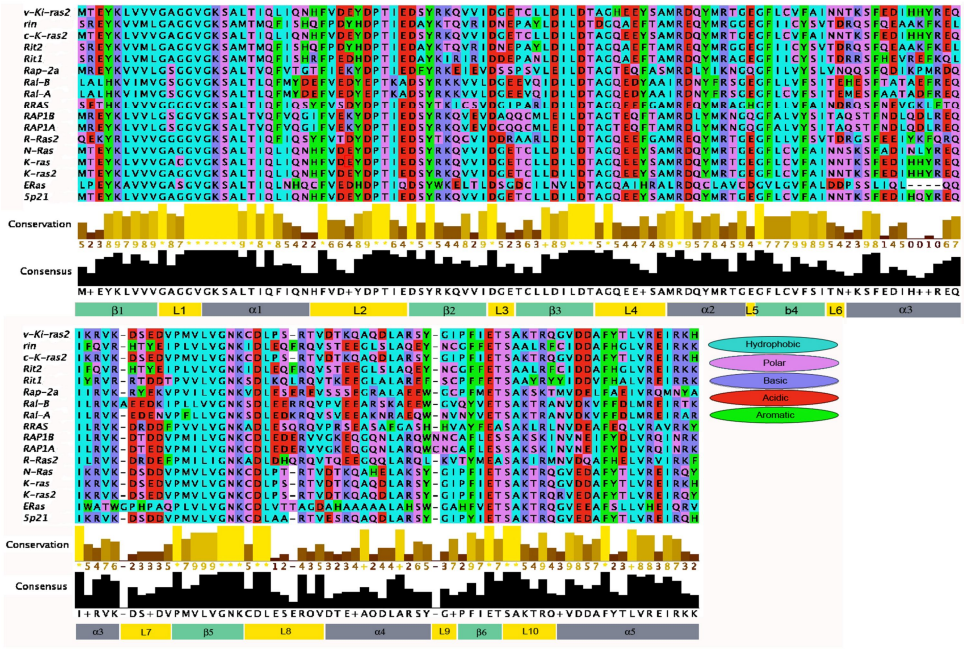


Figure 16: Multiple Sequence Alignment of Ras subfamily genes
Multiple Sequence Alignment of Ras subfamily genes found in Human, obtained with Muscle (Edgar 2004) and viewed with alignment editor (Clamp *et al.*, 2004)

4.3 Correspondence between functional and structural classification

We are interested in the identification of energetic patterns that allow us to distinguish between Ras protein structures at different conformational states. To do so, we come back to Figure 11 in the methodology chapter and decide to distinguish between 4 main groups of structures. On the one side, we aim at classifying the structures by their substrate binding state, being GDP or GTP bound. A preliminary attempt to obtain structures that bind the transition state of the hydrolysis reaction in order to achieve more precise results for the reactivity trends yielded too few structures to be considered in this work and those are included as GTP-bound structures. On the other side, Ras

structure is present in the PDB as a standalone structure or as a partner in protein complexes, both with effectors and regulators.

As we deal with several structures of just one protein sequence, the first analysis one could think that will yield a useful result is the RMSD value of the superposition of all those structures. In structural bioinformatics one typically is interested in the superposition of the subset of C_α that are shared among the different structures. After initial inspection we found that the C_α RMSD matrix did not bring any signifficative difference between the different clusters of protein structures (data not shown). In fact, even if one may believe that the superposition of all the main chain heavy atoms (CA, N, C, O) would provide a better description of the orientation of the main chain permanent dipoles of the protein (see Figure 18), neither this was able to distinguish between the different clusters of structures.

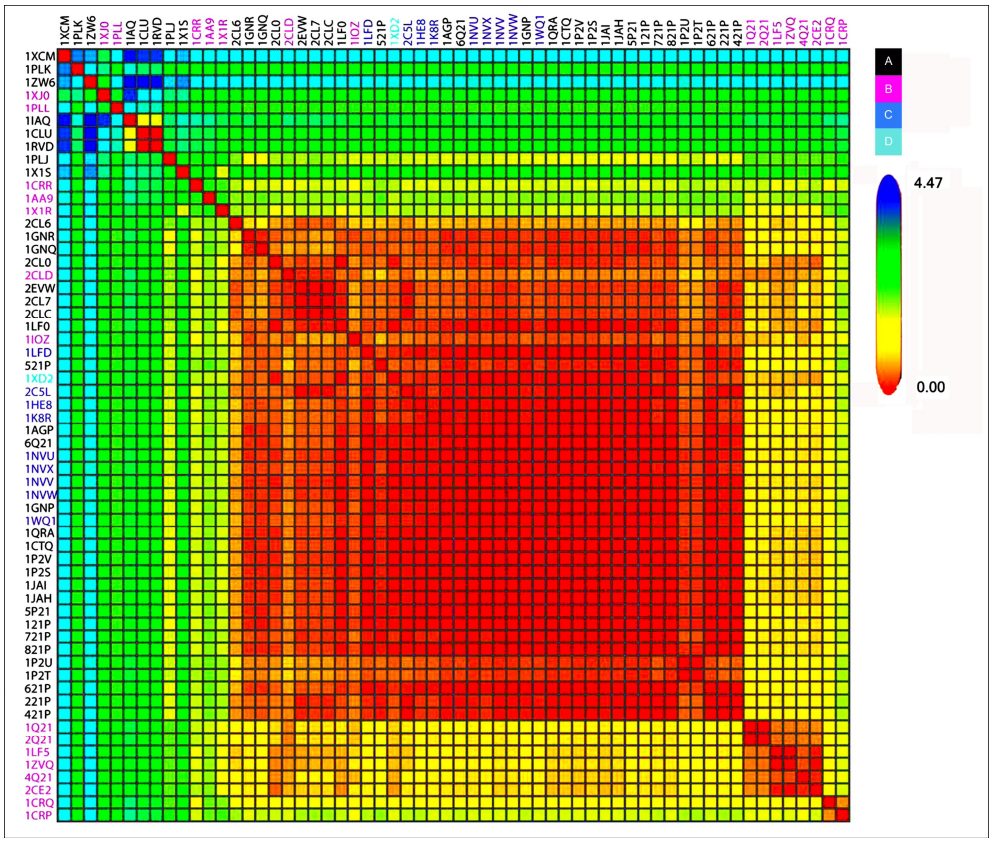


Figure 17: C_{NO} RMSD matrix
RMSD matrix of the superposition of all shared main chain heavy atoms for the Ras protein structures analyzed in this work. Clearly the functional (conformational) clusters do not follow any clear structural pattern. Results obtained with the structural alignment plugin in Adun (Johnston *et al* 2005), developed by Sai Phanindra (unpublished results)

Thus, it is obvious that whether some difference should occur this would be mainly related with the active site residues side chains which are, obviously, much more flexible, discarding this way the direct use of the X-ray structures

for structural comparison. The approach taken here is, instead of searching for the conformational space explored by the side chains, directly use such conformational sampling to perform free energy calculations of residue solvation by means of the PDLD/S-LRA method.

4.4 Stability analysis of the active site residues

As we realized in section 'Correspondence between functional and structural classification' (page 40) there is not any significant global structural difference between the different clusters, even when considering just the main chain dipoles. In the following sections we move to an electrostatic description by means of the analysis of the stability of all active site residues. We performed the solvation free energy calculations using PDLD/S-LRA method (see pp. 26 for detail) and tried to classify residues, regions and conformations with a realistic statistical measure that shows the significance of the differences in free energy. Figure 18 contains the raw data obtained for the active site residues selected by using the criteria described in the Methods section above.

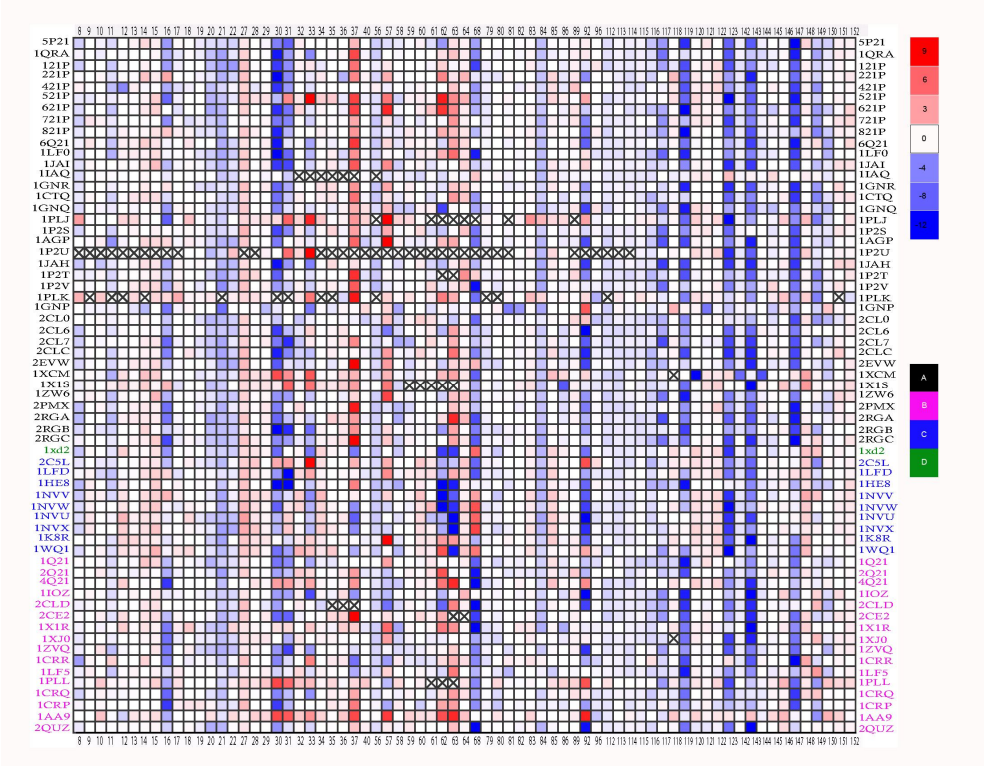


Figure 18: Solvation energy matrix.
Solvation free energy $\Delta\Delta G_{solv}^{w\to p}$ values for all active site residues (columns) in the selected Ras protein structures (rows). Blue is used for residues that contribute to the stabilization of the structure. Destabilizing solvation energies are shown in red (the energies were calculated in kcal mol^{-1}). The different classes of 3D structures are represented with different colors (see right hand side panel)

By eyeballing we found that some positions show a clear thermodynamic stability pattern (*e.g.* 119 or 143) while others yield the opposite trend, although some of them present in fact a mixed behavior. As expected for a iological macromolecule, most of the residues are stabilizing the structure, and only some residues seem to provide hints for functional behavior from the free energy values.

If now we analyze the mean of solvation free energy calculations within each column, we expect obtaining a way to compare the Ras GTP bound state (Group-A) to GDP bound state (Group-B)[1]. Considering that a difference in the average free energy of solvation of 1 kcal mol-1 enough to distinguish between residues more stable in one of the two groups, we clustered the residues in:

- a “pro-ON cluster”, containing the residues that show more stability in Group-A than in Group-B PDB structures, this is, that favor the Ras protein to be in the ON state: residues **10**, 11, 20, 22, 30, 31, 32, 36, 62, 63, 92, 117, 123, 146, 147 and 149.
- a “pro-OFF cluster”, containing the residues that show more stability in Group-B than in Group-A PDB structures: residues 37, 57 and 68 .

We may repeat the classification but considering the “extended” groups (A-ext=A+C) and (B-ext=B+D), and we find

- an “ext-pro-ON cluster” , formed by residues 10, 11, 30, 31, 36, 62, 63, 84, 92, 117, 123, , 146, 147 and 149
- and “ext-pro-OFF cluster”, formed by residues 16, 37, 57, 68 and 119

It is worth noting that the standard deviation is less than 1.5 kcal mol^{-1} only in the case of residue 10, while the dispersion of results for the rest of the values listed above is much wider.

Concerned about the conservation pattern of the residues and its relationship with their stability, we noted that residues highly positionally conserved may belong to the set of stable residues (16, 20, 22, 40, 56, 58, 68, 115, 116, 117, 119, 143, 145). The same occurs with sequence positions that conserve the character of the aminoacid side chain (8, 21, 33, 79, 82, 84, 112, 113, 114, 149).

More interesting are the residues that show instability and are conserved positionally (10, 14, 15, 18, 28, 34, 35, 57, 60, 120, 146) or physicochemically (29, 32, 64, 89 and 152).

[1] recall that all calculations are done in absence of the substrate, which ensures a consistent direct comparison of the stability trends for all different structures in a systematic way.

From the above qualitative description we were able to see what are the positions that show thermodynamic stability and sequence conservation pattern (both at amino acid level and based on the nature of the side chain) and further can we see any pattern-change in the mean of the solvation energies across each group. It is encouraging to note how a catalytically important residue like Lys16 is stable in all functional groups except in C, the one corresponding to the activated complex RasGAP/Ras, confirming our previous results in (Bonet *et al.,* 2006) (see section 4.7).

4.5 Statistical tests

The qualitative analysis should be followed by a quantitative characterization of the data, trying to find signifficatively informative patterns in the calculations.

4.5.1 Z-test

Generally a Z-test is employed when one wants to compare two means to suggest whether the samples come from the same population. A 'Z-test' reflects the number of standard deviations above or below the mean a particular score is. One-sample Z-test can be used to assess whether a sample drawn at random from a population tends to have the same characteristics as the population from which it is taken. This implies one needs to assume a given underlying probability distribution, which limits the applicability of the method. Here, thus, assuming that the data follows a normal distribution, we test the difference between the means of Group-A and Group-B. The underlying null hypothesis, H_0, was that there is no difference in the distribution of the solvation free energy between the two groups based on the nature of ligand, and that any variation is due to random variation or chance: $H_0 : \mu_A = \mu_B$.Our alternate hypothesis (H_A) was that the difference in the Z-score is due to nature of the ligand attached ($H_A : \mu_A \neq \mu_B$) Similarly our assumptions were same for the A-ext and B-ext groups

$H_0 : \mu_{A-ext} = \mu_{B-ext}$

$H_A : \mu_{A-ext} \neq \mu_{B-ext}$

The Z-score yielded the solvation energies for the following residues to be statistically significant:

- When comparing groups A vs B:
    - P-Loop: 10 11 16
    - switch-I: 30 36 37

    - switch-II: 56 57 68
    - G4: 112
    - G5: 146 147 149
- When comparing groups A-ext vs B-ext:
    - p-Loop: 10 11 16
    - switch-I: 27 30 36
    - switch-II: 56 57 68
    - miscellaneous: 84
    - G4: 122
    - G5: 146 149

4.5.2 Non parametric Wilcoxon Rank test

As described in the previous section the use of parametric tests involves assuming an underlying probability distribution. We have also applied a non-parametric test, the Wilcoxon rank-sum test, to assess whether two samples of observations are from the same distribution. The null hypothesis assumes that the two samples are drawn from a same population, and therefore both have equal probability distributions.

Wilcoxon test take the paired observations, calculate the differences, and rank them from smallest to largest by absolute value. Finally, the p-value associated with this statistic is found from the table. The Wilcoxon test is an R-estimate (robust estimation based on rank test). Thus our null hypothesis is that the median difference between pairs of observations of solvation free energy ratings both in Group-A vs. Group-B as well as Group A-extended vs. B-extended are same. On the contrary, the alternate hypothesis is that the median difference between pairs of observations of solvation free energy ratings in both the groups are not same.

The residues found to be signifficanly differents are:

- When comparing A vs. B:
    - P-Loop: 10
    - switch-I: 30 36 37
    - miscellaneous: 96
    - G4: 123
    - G5: 146 147 149
- When comparing A-ext vs. B-ext:
    - P-Loop: 10
    - switch-I: 30 36 37

- miscellaneous: 96
- G4: 123
- G5: 146 147 149

In the next sections we will try to rationalize the meaning of such values.

Wilcoxon Test Results A-B (53)										
position	**Wilcox- Z-value**	**p-value**	**significance**	**region**	**aminoacid**	**Physico-chem**	**conservation**	A vs A-ext	A vs B	A vs B-ext
10	-2.379	0.017	significant	p-loop	Gly	Hydrophobic	Conserved	:	+	+
30	-2.689	0.007	significant	sw-I	Asp	Acidic	Semi-cnserved	:	+ + +	+ + +
36	-2.726	0.006	significant	sw-I	Ile	Hydrophobic	Conserved	:	+	+
37	-2.272	0.023	significant	sw-I	Glu	Acidic	Conserved	-	- - -	- - -
96	-2.896	0.004	significant	misc	Tyr	Aromatic	Semi-cnserved	:	:	:
123	-1.810	0.070	significant	Guan_bind	Arg	Basic	Conserved	:	+ +	+
146	-3.258	0.001	significant	Guan_bind	Ala	Hydrophobic	Conserved	:	+	+
147	-2.095	0.036	significant	Guan_bind	Lys	Basic	Conserved	+	+ +	+ +
149	-2.689	0.007	significant	Guan_bind	Arg	Basic	Semi-cnserved	:	+ +	+ +

Table 3: Summary of Wilcoxon Test of A-B group (53-members)[2]

[2] The number of positive or negative signs represents a qualitative measure of the weight from the test. **The data only refers to active site residues. Other resdues that are signifficantly over estabilized in some of the conformations, although are not directly involved in substrate binding are listed in the detailed analysis section**

Wilcoxon Test Results A-B extended									
position	**Wilcox- Z-value**	**p-value**	**significance**	**region**	**aminoacid**	**Physico-chem**	**conservation**	A-ext vs B	A-ext vs B-ext
10	-2.485	0.013	significant	p-loop	Gly	Hydrophobic	Conserved	+	+
30	-2.296	0.022	significant	sw-I	Asp	Acidic	Semi-cnserved	+ + +	+ + +
36	-2.637	0.008	significant	sw-I	Ile	Hydrophobic	Conserved	+	+
37	-2.430	0.015	significant	sw-I	Glu	Acidic	Conserved	- -	- -
96	-2.509	0.012	significant	misc	Tyr	Aromatic	Semi-cnserved	:	:
123	-1.965	0.049	significant	Guan_bind	Arg	Basic	Conserved	+ +	+ +
146	-2.722	0.006	significant	Guan_bind	Ala	Hydrophobic	Conserved	+	+
147	-2.273	0.023	significant	Guan_bind	Lys	Basic	Conserved	+	+
149	-2.864	0.004	significant	Guan_bind	Arg	Basic	Semi-cnserved	+ +	+ +

Table 4: Summary of Wilcoxon Test A-B-extended (63-members)..[2]

4.6 DETAILED ANALYSIS OF RAS ACTIVE SITE RESIDUES

This extended section includes the detailed description of all the active site residues, selected to be (see Methods section) those that are at an electrostatic cutoff distance from the substrate or those that are known to be functionally relevant plus those whose stabilization energies show estabilize differences with other residues. Thus, in some cases the results for the solvation free energy was not included in the analysis of the previous section, but is given here (see Table 5).

In what foloows, as a rule of thumb, a estabilized residue indicates a favorable interaction with the substrate. Thus, a residue that is unstable in the GTP bound system may indicate that participates strongly in the binding of such substrate. One would like to have a number of structures for transition state analogs in order to assess if previous results on the role of active site residues in reactivity (Glennon *et al,* 2000) are generalizable by getting more statistics, although unfortunately that was not possible in this work.

4.6.1 Beta sheet 1 (β1)

Lys5

Our results show significant difference in the thermodynamic solvation energy values between active (GTP) and inactive (GDP) state at position 5 ($\Delta\Delta G_{solv}^{w\to p}$ **is -12.6, -10.0, -14.0 and -10.5 for groups A, B, C and D)**. Lys5 is thermodynamically more stable in the active state than in the inactive state. More, this position shows an overall stability pattern in the Ras•Regulator bound state (Guanocine Exchange Factor – GEF) protein in either active or inactive state. Thus we can show how this residue is important for the overall structure stability of Ras and relevant for the interaction with bothe effectors and regulators.

4.6.2 Flexible turn 1 (L 1)

Gly10

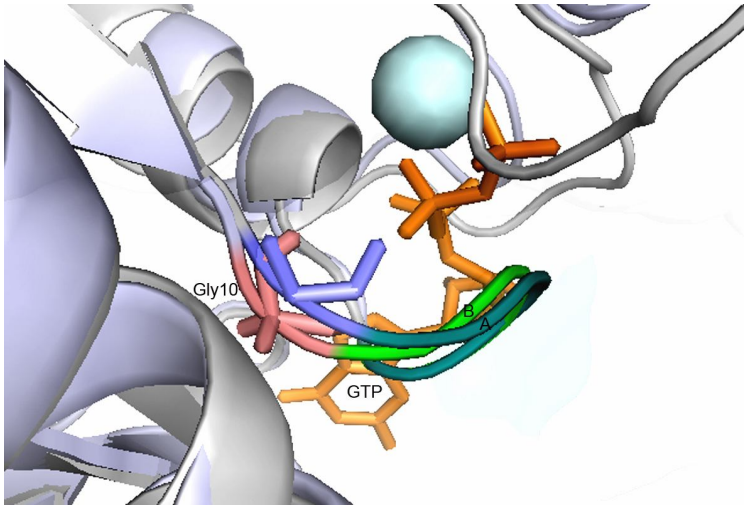


Figure 19: Gly10

Pai *et al*., 1990, showed Ramachandran diagram for the Ras protein refined structure and reported that Gly10 is in a conformation which is not allowed for other amino acids. Based on this they postulated that this could be a possible reason for the sequence conservation at this position in GXXXXGKS/T motif in nucleotide binding proteins and explained why the structure of the Phosphate binding loop (P-loop) is conserved between these proteins.

Interestingly, the Wilcoxon test identifies this residue as being different between the A and B structures, more stable in A. This may suggest this is a Pro_OFF residue, that is destabilized in the GDP bound structured, favoring this way the diphosphate when this is bound.

Gly12

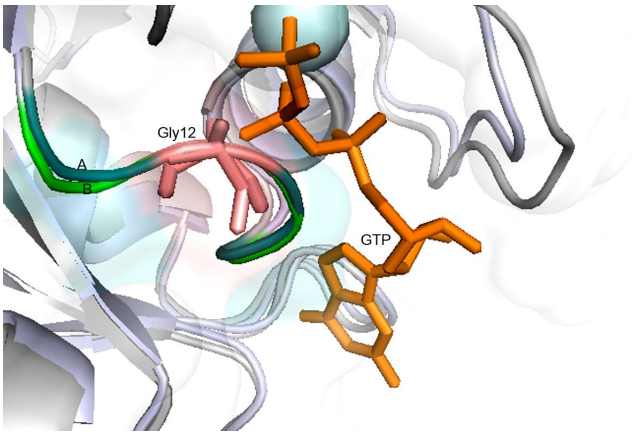


Figure 20: Gly12

Despite the relevant role of Gly12 in many mutational studies no significant difference in the thermodynamic energy was reported earlier (Bonet *et al*., 2006) and no ifference changes in . Similarly we were also unable to see any significant thermodynamic pattern at this position in four different states.

Gly13

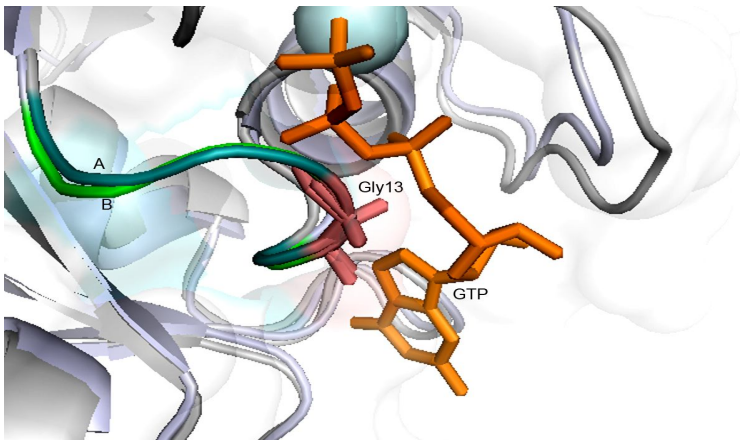


Figure 21: Gly13

The backbone of Gly13 donates an H-bond to the -β - γ- bridging oxygen of GTP and this H-bond is catalytic. Hence such hydrogen bond is expected to be stronger in the transition state (Maegley *et al*. 1996) and correspondingly we expected finding a pro-OFF residue (more unstable in the GTP-bound conformation). As occurred for Gly12, we are unable to find any ifference ifference between conformations, though.

4.6.3 Helix 1 (α 1)

Gly15

Although a similar irregular Ramchandran plot position is found for this residue as occurred for Gly 10, no ignificant difference in the electrostatic stability was found in our calculations.

Lys16

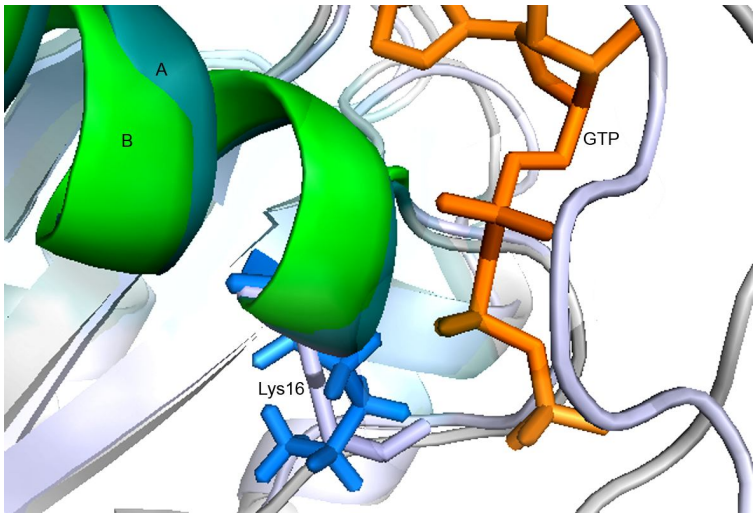


Figure 22: Lys16

Lys16 helps the GTP molecule to take the correct conformation inside the Ras active site (Paduch *et al*., 2001). The ε-amino group of Lys 16 binds/coordinate to the β- and -γ-phosphate oxygens but is closer to the gamma and therefore it is assumed that this could assist GTP hydrolysis (Pai *et al*., 1990). Hence Lys 16 is important in the catalytic mechanism interacting with the phosphate (Scheffzek *et al*., 1997; Maegley *et al*., 1993).

Simulation study show thermodynamic ignificant of Lys16 in Ras•GAP (Bonet *et al*., 2006) and we confirm that ignificant was more than 2-kcal mol^{-1} stronger than the one found in standalone Ras . In addition, we also found thermodynamic ignificant of at least 2-kcal mol^{-1} when compare the same residue to the average thermodynamic values/stability of GTP bound structures with Ras•GEF Y64A:SOS:Wild (Margarit *et al*., 2003 PDB: 1NVV), Ras•RalGDS complex (Huang *et al*., 1998 PDB: 1LFD), Ras•PI3K-γ complex (Pacold *et al*., 2000 PDB:1HE8) and lower eukaryotic Byr2 kinase in Ras•Byr2RBD complex, a homolog of mammalian Raf-kinase (Scheffzek *et al*., 2001 PDB:1K8R) bound states. This seems to be a clear indication of the (as expected) strong participation of the residue in the phosphate hydrolysis reaction.

Ser17

This position was determined to be the part of Helix-α1 and an important element of the active site (Pai *et al*., 1990) and a member of GXXXXGKS/T motif (Bourne *et al*., 1991; Colicelli 2004; Wennerberg *et al*., 2005). Ser 17 also binds to Asp 57 via its side chain (Pai *et al*., 1990). Ser17 has been reported to be an active mutagenic residue (Polakis & McCormick 1993)

S17A mutation restricts magnesium binding (Farnsworth & Feig, 1991; Kraulis *et al*., 1994). Bonet *et al*., in their simulation studies found that the effect of Ras•GAP complex formation cause the PM1 region to be ~10 kcal mol^{-1} higher in energy in the complex state than that in free Ras, and hence concluded that Ser 17 can help in the reactive process (Bonet *et al*., 2006). However, our present calculations brought the conclusion that this residue does not present ignificant differences between the different conformational states. More work is needed to understand the origin of this discrepancy.

Gln22

Ras_{Gln22} in Ras•PI3K-γ complex (Group D) shows an overall thermodynamic instability of at least 2-kcal mol^{-1} when compare the same residue to the average thermodynamic stability of wild Ras GTP bound form. Although the values are not statistically ignificant, this may imply an important role of this residue for complex formation.

Asn26

Position-26 of Ras•PLC ε complex (Bunney *et al*., 2006 PDB:2C5L), Ras•RalGDS complex (Huang *et al*., 1998 PDB:1LFD), lower eukaryotic Byr2 Kinase in Ras•Byr2RBD complex, a homolog of mammalian Raf-kinase (Scheffzek *et al*., 2001 PDB:1K8R) and with GTPase activating protein (GAP) show instability of at least 2-kcal mol^{-1} when compare the same residue to the average thermodynamic stability of GTP bound structure. Thus, based on our MD simulation studies we may say that this position is playing a role in restructuring the complex structure. However, we were unable to see any visible difference in the thermodynamic profile of this position with other effectors and GEF proteins.

4.6.4 Flexible turn 2 (L 2)

His27

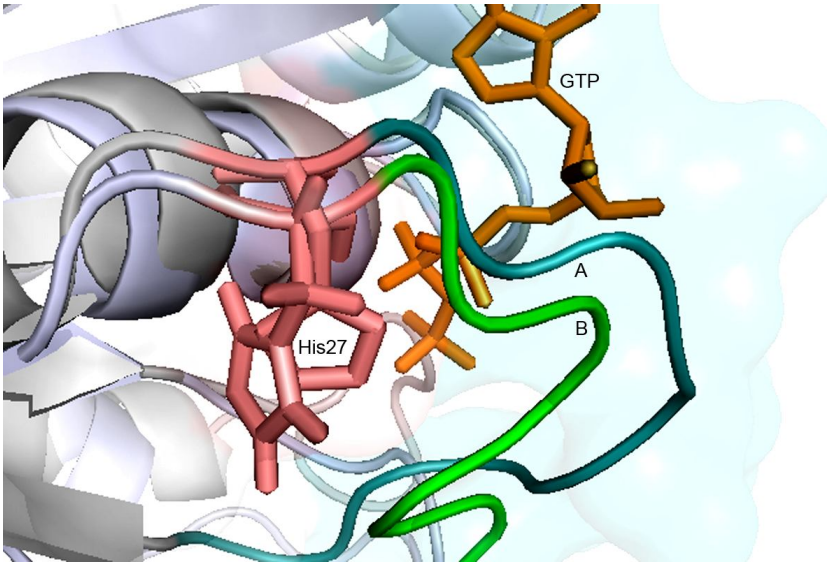


Figure 23: His27

The Ras loop 2 (L2) has an overall impact on the conformational capability of the Ras protein (Pai *et al*., 1990). In our current thermodynamic solvation energy studies we found this position overall destabilized. We found position-27 shows extra destabilization in Ras•GEF complex, which is around 2-kcal mol^{-1}. Although we could not find relevant literature discussing the thermodynamic feature of this position, we predict it having a relevant role in both reactivity and GTP/GDP exchange.

Phe28

Phe28 form an aromatic interaction with the nucleotide base which results tight binding of the base (Burley and Petsko, 1988), Asp33, Pro34 and Thr35 (Pai *et al*., 1990). More, Phe28 is held in place by an aromatic-aromatic interaction with Lys147, which is stretched out over the plane of the phenyl ring (Pai *et al*., 1990). Mutation at this position reduces GEF driven stimulation and GTP dissociation (Polakis & McCormick 1993). Bonet *et al*., (2006) reported no significant change in the stability of this position/residue upon binding to GAP, what we observed and confirm in this study. Moreover, we found that Ras_{Phe28} does not show any visible change in the thermodynamic solvation energy except Ras•PLC ε complex, where overall thermodynamic instability of at least 2-kcal mol^{-1} when we compare the same residue to the average thermodynamic stability of GTP bound structures vs effector bound GTP state. It appears that this residue may play a similar role as His27 in both reactivity and GTP/GDP exchange.

Asp30

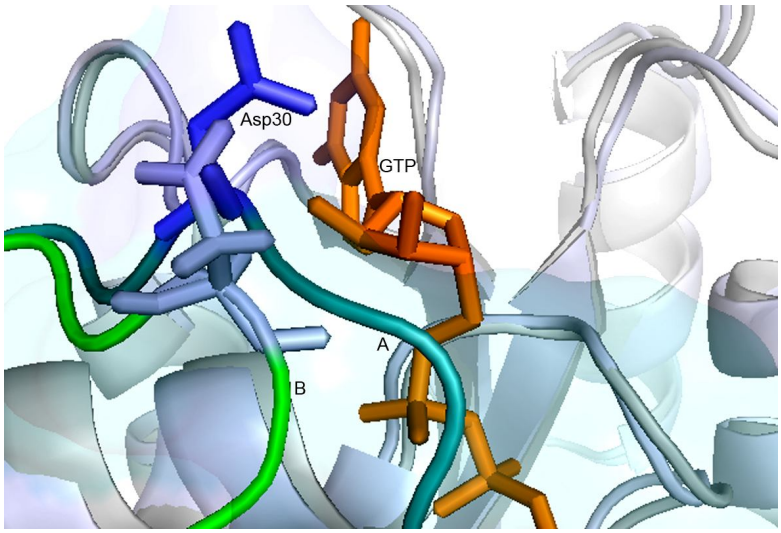


Figure 24: Asp30

Crystal structure of Ras wild type active state protein shows that Asp30 makes weak hydrogen bond with the 2'- and 3'-hydroxyl groups of the ribose sugar of the nucleotide (Pai *et al*., 1990). Asp30 and Glu31 residues of switch-I and loop L2 are very important as these two residues determine target specificities of Rap and Ras proteins toward the RBDs of Raf (Nassar *et al.,* 1996) and RalGDS (Herrmann *et al.,* 1996).

Scheffzek *et al*., (1997) in the crystal structure of the Ras•GAP complex found that Asp30 is not directly involved in the Ras•GAP interaction but along with other nearby residues, namely Glu31, Asp33, Glu37, and Asp38 creates a negatively charged surface and this negative charged surface is latter used for interaction with effectors. In the current study we see similar stability at this position between Ras•GAP complex and Ras·GTP bound conformation/state. It was also found that Asp30 makes a direct hydrogen bond with Lys602 of GEF, thus is involved with the Ras•GEF complex interaction (Margarit *et al*., 2003) and more thermodynamic stabilization. In current study this position also displays difference in the stability between GDP and GTP state and shows more stability in GTP than GDP bound structure. We observed a dramatic pattern of thermodynamic stability in the Ras-effector interaction where all the interaction is via the Ras Binding Domain (RBD). Marked estability was observed in the thermodynamic solvation energy of Asp 30 in Ras•GTP vs. Ras•PLC-ε complex while estability was also observed Ras•Byr2RBD complex. Strikingly marked stability was observed in Ras•PI3K-γ complex.

In the Ras•PLCε complex crystal structure it was shown that D30E Ras mutant has a lower affinity for RA2 domain (Bunney *et al*., 2006). It seems confirmed, then, the role of this residue in the reaction process, but also in the GTP/GDP exchange.

Glu31

Glu31 shows an overall thermodynamic estability in the GEF bound conformations. This member of the loop region shows a mixed trend of stability in effector bound states. This may imply a Pro-OFF character of the residue, although further binding calculations are needed to ensure this point.

In the Ras•PLCε complex crystal structure it was shown that E31K ras mutant has a lower affinity for RA2 domain (Bunney *et al*., 2006). We observed marked stability at this position when Ras•GTP is in complex with RalGDS and PI3K-γ, while on the other hand Ras Glu31 shows marked estability when Ras•GTP vs. Ras_{GTP}•PLCε complex are compared. Similarly Ras Glu31 shows thermodynamic estability when Ras•GTP vs. Ras•Byr2 Kinase and Ras•GAP complexes are compared.

Tyr32

Tyr32 is very flexible and may exhibit different conformations as could be shown in crystal structures of p21 complexed with different GTP analogues (Franken *et al*., 1993 1AGP). Milburn *et al*. (1990) reported that in the complex p21 –GppCp, Tyr 32 binds to the -γ-phosphate of the nucleotide analogue and RasQ61G mutant also adopt this conformation, makes a direct hydrogen bond contact with the -γ-phosphate of the nucleotide (Ford *et al*., 2006) while in RasA59G mutant the γ-phosphate contact is not direct rather through a water molecule (Hall *et al*., 2002). Pai *et al*., (1990) proposed the roles for such a side chain that either could be to desolvate part of the active site, or it could increase the positive potential created by the main chain nitrogens of loop L1 and the side chain of Lys16. In the wild-type p21 Ras structure, Tyr32 points to the solvent, whereas both in the RasA59G and RasQ61G structures, the bulky phenol group of Tyr32 rests above the phosphates in a crevice between switch-I and the P-loop (Hall *et al*., 2002; Ford *et al*., 2006). Both in RasA59G and RasQ61G structures, the hydroxyl group of Tyr32 is making a water-mediated hydrogen bond with the γ-phosphate, and its bulky phenol group is protecting the phosphates from the surrounding solvent (Hall *et al*., 2002).

Using fluorescence labeling technique, the effect of various Ras mutations at position-32 on GTP hydrolysis were studied and it was found that Y32W ras (quasi-wild-type) mutation has only a small effect on the kinetic properties of Ras (Polakis & McCormick 1993; Rensland *et al*., 1995). In order to investigate the Ras•Gap interaction Scheffzek *et al*., (1997) found that Tyr32 makes van der Waals interactions with (GAP) Leu902. However, negligible change in solvation energy of Ras Tyr32 upon binding to GAP was shown in previous studies (Bonet *et al*., 2006). In the current calculations a non-signifficant estability on trend for partner-bound conformations is detected.

Asp33

$Asp33_{Ras}$ is a member of switch-I region negative patch (Scheffzek *et al*., 1997), being part of the estability sphere of Mg^{2+}. $Asp33_{Ras}$ makes a direct contact with the solvent exposed part of loop L6c (GAP) specifically with $Lys949_{GAP}$. More, from the invariant residues of GAP on helix -α–7c, only $Asn942_{GAP}$ makes a direct contact with Ras protein making a hydrogen bond with Asp33 of Ras, and $Asn942_{GAP}$, being is stabilized by the carbonamide group of $Gln938_{GAP}$ (Scheffzek *et al*., 1997). Interestingly, results from Bonet *et al*., (2006) show that even being stabilized by the GAP (from residues Ser936-Glu950), the overall effect of complex formation is unfavorable and assume a probable implication in catalysis. Mutational studies of $Asp33_{Ras}$ show partially reduced GTPase activity although a similar affinity with GAP (Polakis & McCormick 1993). More, relationship between Asp33 of Ras and catalytic Gln61 has been reported earlier (Scheffzek *et al*., 1997; Shurki & Warshel 2004). Interestingly, this study also show an overall thermodynamic instability when compare the thermodynamic stability of this residue in Ras•GTP bound structure with Ras•GAP complex. Thus, based on our MD simulation studies we suggest that this residue is essential for the mechanism of GTPase activation by GAP, although the results are not conclusive (not statistically estability) and further calculations with a bigger dataset are needed .

In this study we also studied the role of Asp33 of p21ras with the effectors. Crystal structure of Ras•Byr2RBD complex reveals that three residues play an important role for the interaction with Byr2 kinase, namely Asp33, Asp38, and Tyr40 (Scheffzek *et al*., 2001). The side chain of $Lys101_{Byr}$ at the C-terminal end of helix –α-1 is in an orientation where a salt bridge may be formed with $Asp33_{Ras}$ or $Asp38_{Ras}$ (Scheffzek *et al*., 2001). Mutational

analysis of the Ras•Byr2RBD complex interface reveals that RasD33A and D38A mutants exhibit 10-fold increase of the K_D value for each mutant. While, no significant change in the solvation free energy of this residue in the Ras•GTP vs. Ras•Byr2RBD complex was seen hence supporting probability for the latter residue.

When compared the solvation free energy of Ras•GTP vs. Ras•PLC ε complex, we observed marked thermodynamic estability in complex state. Thus, based on our MD simulation studies we may say that this position is playing a role in restructuring the Ras•PLC ε complex structure. More, we also observed 2-kcal mol^{-1} thermodynamic estability in the solvation free energy calculations of Ras•RalGDS complex while Ras•PI3K-γ complex shows thermodynamic stability. In fact we see more estability trend of switch-I region in Ras•PLC ε complex while on contrary more stability trend in the Ras•PI3K-γ.

An important aspect of this study was the thermodynamic solvation free energy of $Asp33_{Ras}$ in Ras•GEF complex, we found that Ras Position-33 shows thermodynamic stability in all GEF bound structures more obviously in the GDP bound state and then in wild type Ras bound GEF. Finally, we see $Asp33_{Ras}$ is more stable in the GDP bound state than GTP bound state.

Pro34

Pro34 is reported to stiffen the loop L2 in Ras protein (Pai *et al*., 1990) and that the X-ray crystallographic structure of c-H-Ras shows that Pro34 is always in a trans conformation, which could possibly increase the transition state activation (John *et al*., 1988; Pai *et al*., 1990).

Ras•GAP complex structure shows that $Pro34_{Ras}$ makes hydrophobic interactions with $Leu902_{GAP}$ (Scheffzek *et al*., 1997) while Bonet *et al*., (2006) reported an overall gain of >1 kcal mol^{-1} during the Ras•GAP binding process confirmed by the current study. More, Ras•GEF crystal structure shows that $Pro34_{Ras}$ and $Ile36_{Ras}$ of the switch-I region pack against hydrophobic side chains in SOS (Leu687, Met726, and Trp729). While in this study we see thermodynamic solvation free energy estability in Ras_{Y64A} : SOS: Ras_{Wild} structure.

Thr35

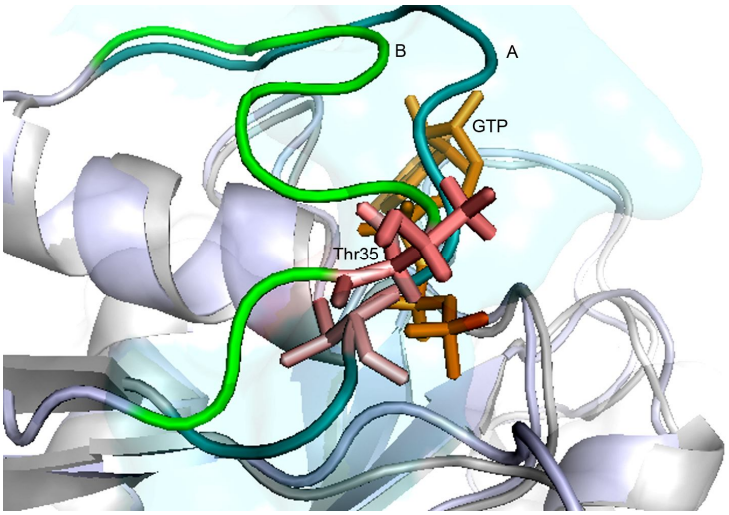


Figure 25: Thr35

Thr35 is well conserved in Ras family proteins (Wittinghofer and Nassar 1996). In the triphosphate active state Hydroxyl side chain (OH) of $Thr35_{Ras}$ is coordinated to magnesium ion (distance 2.2 Å) and its methyl packs against the effector loop, and where $Tyr32_{Ras}$ is close to the phosphates (closest approach to β-Phosphate 5.2 Å) (Spoerner *et al*.2001). In GDP bound state $Thr35_{Ras}$ swings away from Mg^{2+} (Milburn *et al*., 1990. Schlichting *et al*., 1990) as a result the hydrophobic side chain of Ile36 which in the triphosphate form is exposed to the solvent move towards the surface of the protein in the GDP form. This arrangement might balance the loss of the interaction energy when Thr35 leaves the Mg^{2+} and -γ-phosphate coordination, by removing the isoleucine side chain from the solvent (Pai *et al*., 1990). Carbonyl group of $Thr35_{Ras}$ assists Gln6l to activate Wat175 and subsequent GTP-Hydrolysis (Pai *et al*., 1990; Scheffzek *et al*., 1997; Glennon *et al*., 2000) interacting with the phosphoryl group in the Ras•GAP transition state (Scheffzek *et al*., 1997).

Thus, dynamics of the switch-I region appear to be responsible for the conservation of $Thr35_{Ras}$ in GTP-binding proteins (Spoerner *et al*.2001). Minor changes in the switch region, such as removing the side chain methyl group of $Thr35_{Ras}$, drastically affect dynamic behavior and as well as interaction with effectors (Spoerner *et al*., 2001). Crystal structure of Ras T35S mutant (partial loss-of-function mutant) shows distortion of switch-I and switch-II regions, and these two regions are invisible in the crystallography. More, in $T35S_{Ras}$ Mg^{2+} slips somewhat (0.3 Å) further away from β- and γ-phosphate oxygens as compared with wild type Ras (Spoerner *et al*., 2001).

It was found only 3-fold decrease in Mg^{2+} binding affinity for the T35S and 4-fold decrease in T35A mutant as compared with wild type (Spoerner *et al*.2001). The T35S mutation is however compromised in its ability to interact with all effectors, even though it is not involved in any direct interaction (Scheffzek *et al*., 2001). Similarly, in the Ras•Byr2RBD complex Ras Thr35 does not make any direct contact with Byr2 Kinase, nevertheless T35S mutation increases the K_D 12-fold, most likely due to a change in the dynamic properties of Ras (Spoerner *et al*.2001). However, in this study we could not see any behavioral change in the solvation free energy values of Ras in GEF and effector bound states, which may advocate for its sequence conservation.

$Thr35_{Ras}$ makes a water-mediated polar contact with $Lys949_{GAP}$ (Polakis and McCormick 1993; Scheffzek *et al*., 1997) and in this study we also notice almost 2-kcal mol^{-1} destabilization in the thermodynamic free energy values of $Thr35_{Ras}$ in Ras vs. Ras•GAP complex, suggesting that it is the interaction with GAP, which makes this residue participate more in the reaction, in agreement with Glennon *et al*., (2000). Bonet *et al*., (2006) report no significant changes in the free energy values of $Thr35_{Ras}$.

Ile36

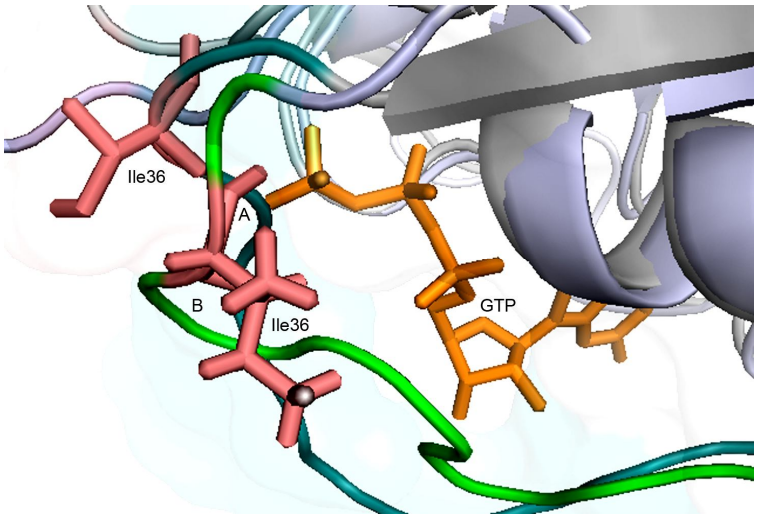


Figure 26: Ile36

In the wild-type GTP-bound structure of Ras (active state labeled as state-A), Ile36 is solvent exposed, whereas in the GDP-bound structure (inactive state labeled as state-B) Ile36 moves towards the surface (Pai *et al*., 1990) thus Ile36 is essential for substrate specificity. In wild type Ras structure, main chain carbonyl of Ile36 interacts with the hydroxyl group of Tyr64, whereas the side chain of Arg68 is stabilizing the loop L4 (60-62) through

direct or water mediated hydrogen bonds (Hall *et al*., 2002). Whereas, the side chains of Glu37, Tyr64, Arg68, Tyr71, Leu56, Ala59, and Thr58 form a pocket that is stabilized by *van der Waals* interactions and water-mediated hydrogen bonds (Hall *et al*., 2002). In the RasA59G GppNp structure (PDB: 1LF5), this pocket unfolds and side chain of Arg68 adopts a new conformation in which the guanidinium group is making strong hydrogen bonds with Glu37, Ile36 and Gly59 (Hall *et al*., 2002). These interactions with guanidinium group pull the Ile36/Glu37 polypeptide (by 1Å) towards the switch 2 region (Hall *et al*., 2002). Consequently, Tyr64 and Tyr71 point to the solvent and the phenol group of Tyr71 at *van der Waals* contact distances from the carboxylic group of Glu37 (Hall *et al*., 2002).

In the Ras•GAP complex, $Ile36_{Ras}$ forms a hydrophobic contact with $Leu902_{GAP}$ (Scheffzek *et al*., 1997). Bonet *et al*.,(2006) did not observe any (de)stabilization effect of this residue upon Ras•GAP binding and in this study we also could not see any (de)stabilization effect.

In case of Ras•GEF complex it was found that $Pro34_{Ras}$ and $Ile36_{Ras}$ of the switch-I region pack against hydrophobic side chains of $Leu687_{SOS}$, $Met726_{SOS}$, and $Trp729_{SOS}$ (Margarit *et al*., 2003), whereas in this study we could not see any change in the solvation free energy values in the Ras•GEF complex structures. More, we also could not see any change in the solvation free energy values in Ras•Effector(s) complex.

Glu37

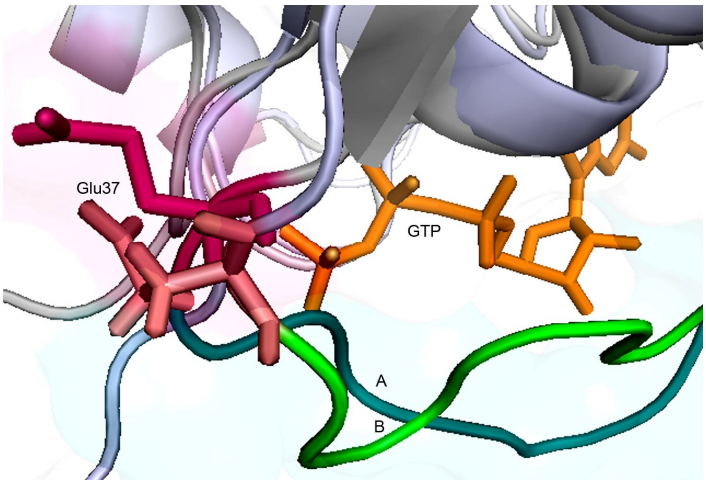


Figure 27: Glu37

$Glu37_{Ras}$ is solvent exposed in wild-type Ras, whereas in Ras_{A59G} Glu37 makes hydrogen bonds with Arg68 and *van der Waals* interactions with Tyr71, which protects it (via bulky phenol group) partially from the

surrounding solvent (Hall *et al*., 2002). While in RasQ61G mutant, Glu37 does not make any contact with Arg68 (Ford *et al*., 2006). In our calculations we also see a marked thermodynamic stability in Ras_{A59G} (GDP bound state: state-B) than the GTP bound Ras wild type as well as Ras_{A59G} mutant.

Ras switch-I contain five acidic residues which are Asp30, Glu31, Asp33, Glu37, and Asp38. All these negatively charged residues create a negatively charged surface patch (Scheffzek *et al*., 1997). This negatively charged surface patch is used for interaction with effectors (Nassar *et al.,*1995) and with GAP-334 (Scheffzek *et al*., 1997). Lys949 of conspicuously exposed loop L6c of GAP points into the highly negatively charged surface patch. More, Scheffzek *et al*., (1997) also reported that the side chains of Asp30, Glu31, and Glu37 (members of the negatively charged surface) are not directly involved in the interaction with GAP. Differential solvation of $Glu37_{Ras}$ in the three systems (Ras alone, Ras' and Ras•GAP complex) indicates that $Glu37_{Ras}$ is stabilized by the conformational change but not so strongly by the interaction with GAP itself (Bonet *et al*., 2006). In our thermodynamic solvation energy calculations, we found a marked thermodynamic stability at this position in the in the Ras•GAP complex.

On one hand $Glu37_{Ras}$ form a hydrogen bonding with Arg688 of SOS and at the other hand $Glu37_{Ras}$ also forms hydrogen bonds with the side chains of Asn691 and His695 of SOS_{cat} (Margarit *et al*., 2003). This bidentate hydrogen bonding is a central aspect of the Ras•SOS interface (Margarit *et al*., 2003). Sequence alignment data of four different species GEF shows that Arg688, which forms hydrogen bonds with $Glu37_{Ras}$, is conserved. Thus sequence conservation at this position raises the possibility that the distal interaction with Ras•GEF may also be conserved (Margarit *et al*., 2003). In this study, we see a strong thermodynamic stability at this position in all Ras•GEF complexes when compared to Ras•GTP. The Ras_{A59G}:SOS:Ras_{A59G} and Ras_{GDP}:$SOS_{DH\text{-}PH\text{-}cat}$:Ras complexes show marked thermodynamic stability.

The role of Glu37 with the Ras effectors could be explained as that salt bridge is formed between Lys223 of PI3Kγ and $Glu37_{Ras}$ in the crystal structure of Ras•PI3Kγ complex (Pacold *et al*., 2000), while our calculations shows thermodynamic estability at this position in Ras•PI3Kγ complex. More, in the Ras•Byr2RBD complex the $Arg74_{Byr}$ in strand β1 forms a salt bridge with the carboxylate group of $Glu37_{Ras}$ (Scheffzek *et al*., 2001). The E37G mutation was one of the first partial loss of function mutant that fails to

interact with Raf but is unimpaired for Byr2 binding (White *et al.,* 1995). Our calculations, show thermodynamic stability at this position in Ras•Byr2RBD complex than Ras in active state.

4.6.5 Beta sheet 2 (β 2)

Asp38

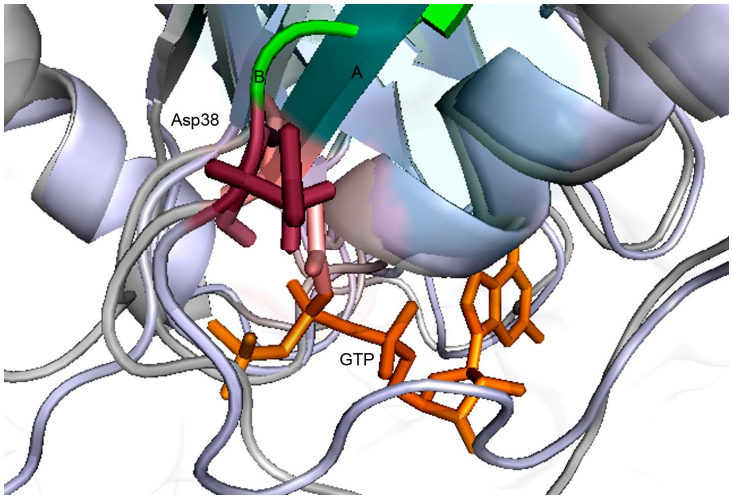


Figure 28: Asp38

$Asp38_{Ras}$ is thermodynamically more stable in the triphosphate bound active state than inactive (diphosphate bound) state. Asp38 plays a significant role in Ras•GAP interaction, allowing us to classify it as a pro-ON residue, Ras D38E mutation has been shown to have a pronounced effect on Ras•GAP interactions (Polakis & McCormick 1993). Bonet *et al*., (2006) found that this residue gain an extra stability upon binding to GAP, probably due to a water-mediated polar contact between $Asp38_{Ras}$ and $Lys949_{GAP}$ (Scheffzek *et al*., 1997). We also found the same pattern of thermodynamic values as reported by Bonet *et al*., (2006). This, then, involves that the residue does not participate directly in reactivity, but indirectly b y helping the GAP partner to position in the complex. I makes, then, an analogous role to residues that for the folding core stable proteins.

In general, Asp38 shows thermodynamic instability when Ras protein is bound with effector proteins, here few effectors of Ras are discussed. The side chain of basic residue $Lys101_{Byr}$ at the C-terminal end of helix α1 is in an orientation where a salt bridge may be formed with acidic residues of $Asp33_{Ras}$ or $Asp38_{Ras}$ (Scheffzek *et al*., 2001). Scheffzek *et al*., (2001) also report that mutating D38E shows rather large effects on Ras•Byr2RBD binding even the charge being unaltered. In our study we found thermodynamic instability in the solvation free energy values when compare

the $Ras_{active\ state}$ with the Ras•Byr2RBD complex. More, here we also report thermodynamic instability in the solvation free energy values at this position when compare $Ras_{active\ state}$ with Ras•PIK-γ and $Ras_{active\ state}$ with Ras•PLC ε complex.

Ser39

In this study Ser39 does not show any difference in the solvation free energy values between Ras active (GTP) and inactive (GDP) bound state. Ras•GAP complex crystal structure shows that $Lys935_{GAP}$ forms an intramolecular salt bridge with $Glu950_{GAP}$ in loop L6c. $Glu950_{GAP}$ as well via carboxylate group, contacts main chain amide group of $Ser39_{Ras}$ (Scheffzek *et al*., 1997). In this study we could not see any change in the thermodynamic energy values in Ras_{Alone} vs. Ras•GAP complex.

The β2 strand residues of Byr2RBD contact Ras switch-I region amino acids in the stretch from residue 38 to 41 (Scheffzek *et al*., 2001). In Ras•Byr2RBD complex, the guanidinium group of $Arg83_{Byr}$ is in close proximity of the effector loop main chain region, allowing a polar contact with the carbonyl oxygen of $Ser39_{Ras}$. More, Scheffzek *et al*., (2001) report that the side chain of Ser39 provide no contribution for Byr2 binding, since the alanine mutation increases the K_d value only 3-fold (which reported and cited else where in case of D33A and D38A to be 10-fold) and the contribution of the main chain contacts originating from Ser39 are not accessible by mutational analysis. Our thermodynamic calculations of Ras•Byr2RBD complex shows that Ser39 gains stability in Ras•Byr2RBD complex than in Ras alone. More, we also observed that Ser39 in Ras•RalGDS complex gains more stability than Ser39 in Ras alone.

Tyr40

Scheffzek *et al*., (2001) showed that Tyr40 play an important role for the interaction with Byr2 kinase, and Y40C mutation show rather large effects on Ras•Byr2RBD binding. Surprisingly in Ras Y40F mutants, greater than 10-fold increase of the K_D value is observed despite the reality that hydroxyl group is not likely to form polar contacts within the interface (Scheffzek *et al*., 2001). In this study we were unable to see any thermodynamic pattern.

Arg41

Arg41 is highly conserved within the Ras subfamily and shows strong thermodynamic stability in Ras protein either alone or partner bound state.

Solvation free energy values in the active state Ras was found to have extra stability as compared to inactive state. We compare the solvation free energy values of this residue in Ras alone vs. Ras effector, GEF, and GAP bound form. $Arg41_{Ras}$ shows gain of stability in thermodynamic values in complex form when compare $Ras_{active\ state}$ vs. Ras•GEF bound form. In the Ras•Byr2RBD complex and Ras•GAP complex, we found this position comparatively destabilized yet stable, while in Ras•PIK-γ complex Arg41 shows gain of stability. More, Arg41 has a modest contribution to the interaction with Byr2 kinase and Arg41 R41A mutant show a 7-fold reduction of affinity (K_D) (Scheffzek *et al*., 2001).

The residue is clearly related with the overall stability of the protein and in addition helps in the formation of complexes.

Lys42

This residue, in a similar way to Arg41, shows thermodynamic stability in Ras protein either alone or partner bound form. Solvation free energy values in the active state Ras was found to gain extra stability as compared to partner bound form. We compare the solvation free energy values of this residue in Ras alone vs. Ras effector, GEF, and GAP bound complexes. $Lys42_{Ras}$ shows gain of stability in thermodynamic values in complex form when compare $Ras_{active\text{-}state}$ vs. Ras•GEF, Ras•GAP, Ras•Byr2RBD, Ras•PIK-γ, Ras•RalGDS, and Ras•PLC ε complexes.

4.6.6 Beta sheet 3 (β 3)

Asp54

Crystal structure of Ras•Byr2RBD complex reveals that $Arg160_{Byr}$ of C-terminal helix is in proximity to $Asp54_{Ras}$ and that $Arg160_{Byr}$ moderately contributes to binding to Ras (Scheffzek *et al*., 2001). In this study we report thermodynamic stability at position-54 ($Asp54_{Ras}$) in Ras•PLC ε, Ras•Byr2RBD, Ras•GAP complexes, Ras•GEF ternary complex (A59G: SOS: A59G) meaning that this position may play some role in binding in above cited complexes. More, we also found $Asp54_{Ras}$ to be thermodynamically unstable in the complex with Ras•PIK-γ and Ras•RalGDS complexes, which may involve an active role in the binding of GDP.

Asp57

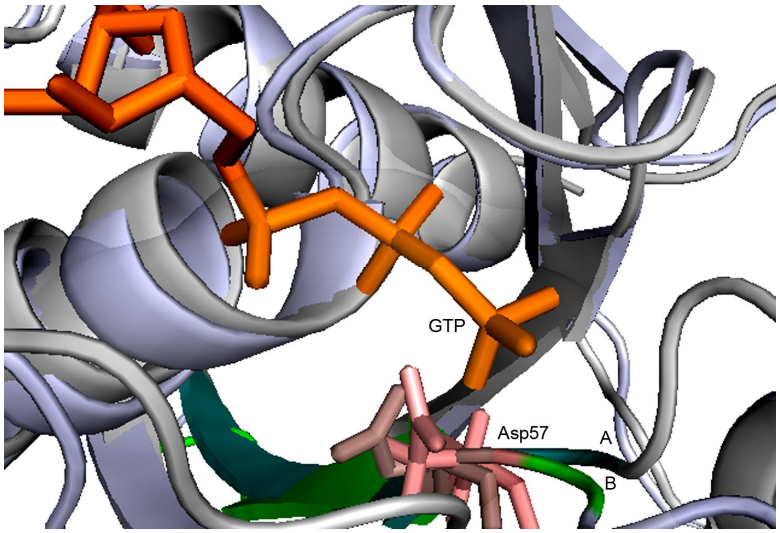


Figure 29: Asp57

Asp 57_{Ras} is the part of β strand 3 of Ras protein, a completely conserved residue of the $_{57}DXXG_{60}$ motif in guanine-nucleotide binding proteins, is completely rigid (Pai *et al*., 1990). The X-ray structures of the C-terminal truncated GDP bound form of Ras reveals that the interaction between the carboxylate group of $Asp57_{Ras}$ and the Mg^{2+} appears to be direct (Schlichting *et al*., 1990), while in the intact Ras (aa: 1-189) in GDP form and several GTP-analogue complexes (Milburn *et al*., 1990; Pai *et al*., 1990), via an intervening water molecule. Kraulis *et al*., (1994) in their NMR structural studies also favours the indirect interaction, showing the distance between the closest oxygen in the carboxylate group of $Asp57_{Ras}$ and the Mg^{2+} in the solution structures is 4.3 ± 0.9 Å. More, crystal structure also reveals that $Asp57_{Ras}$ binds to the side chain of $Ser17_{Ras}$ (Pai *et al*., 1990). Mutation of this residue diminishes the intrinsic rate of GTP hydrolysis (Polakis and McCormick 1993). In this study we also found GTP bound form $Asp57_{Ras}$ more stable than GDP bound form augmenting the idea that this residue plays role in GTP hydrolysis.

More, it is reported that $Asp57_{Ras}$ contributes to catalysis both in Ras` and Ras•GAP complex (Glennon *et al*., 2000) and this residue becomes more unstable upon Ras binding to GAP (Shurki and Warshel 2004; Bonet *et al*., 2006). While in this study we did not see any change in the solvation free energy values in the Ras vs. Ras•GAP complex, we can see the trend of a more unstable residue in GTP bound systems, which advocates for its possible role in reactivity. Further work is neede to assess more clearly the effect of GAP binding on this residue.

Thr58

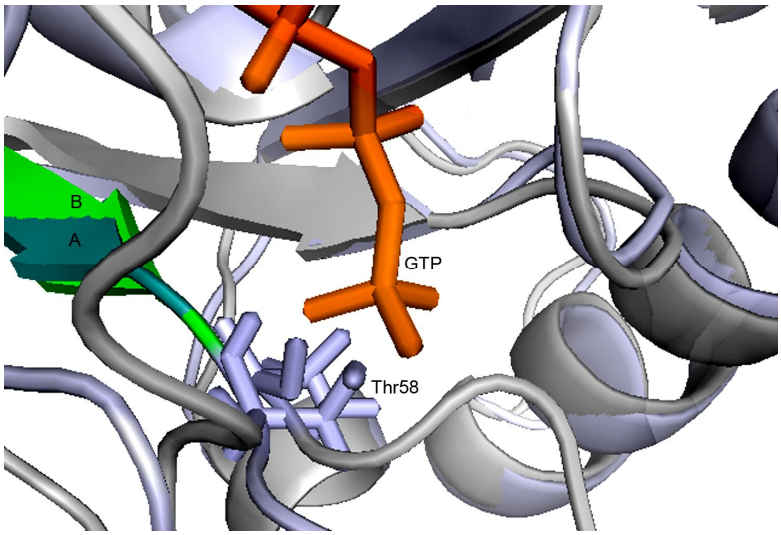


Figure 30: Thr58

In this study we found that $Thr58_{Ras}$ in Ras•PIK-γ and Ras•GAP complexes shows thermodynamic stability when compare the same residue to the average thermodynamic values in Ras•GTP bound form (state-A).

4.6.7 Flexible turn 4 (L 4)

Ala59

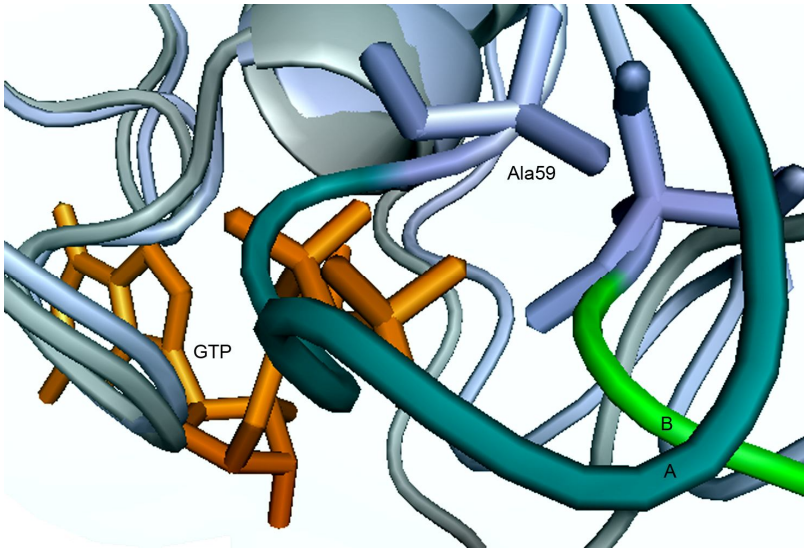


Figure 31: Ala59

Ala59 is located in the conserved $_{57}DXXGQ_{61}$ motif that is important for Mg^{2+} coordination, GTP hydrolysis, and nucleotide binding. This motif is part of an important loop/flexible turn (L4) connecting strand β3 and helix α2 of Ras. The flexibility of this loop is essential for the conformational change that accompanies GTP hydrolysis and γ-phosphate release (Milburn *et al*. 1990;

Pai *et al*., 1990). However, we are unable to see any estabilize trend in our calculations.

Gly60

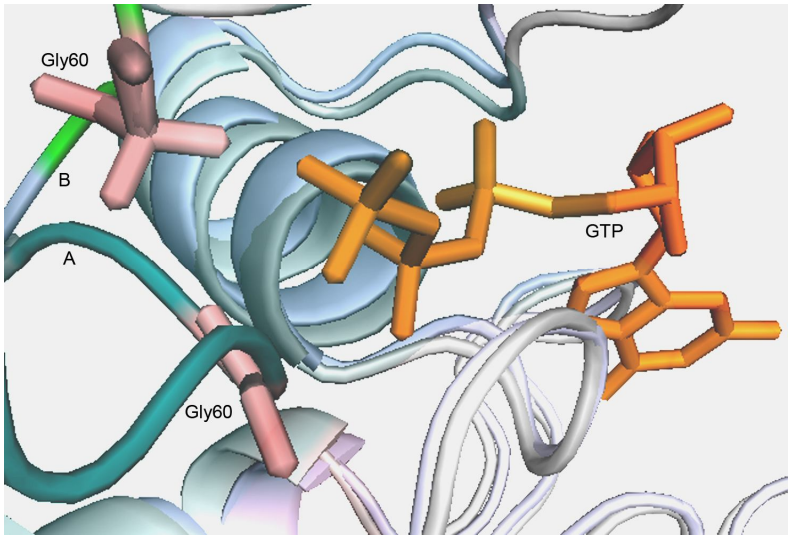


Figure 32: Gly60

The $_{57}$DXXG$_{60}$ sequence motif (where X is any amino acid) in the GTP binding pocket of Ras is conserved in all GTPase proteins (Pai *et al*., 1990). In this study we could not see any significant change in the thermodynamic energy vlues in Ras either in GTP or GDP bound conformation and Ras vs. Ras•GAP, Ras•GEF, and Ras•Effector complexes.

Gln61

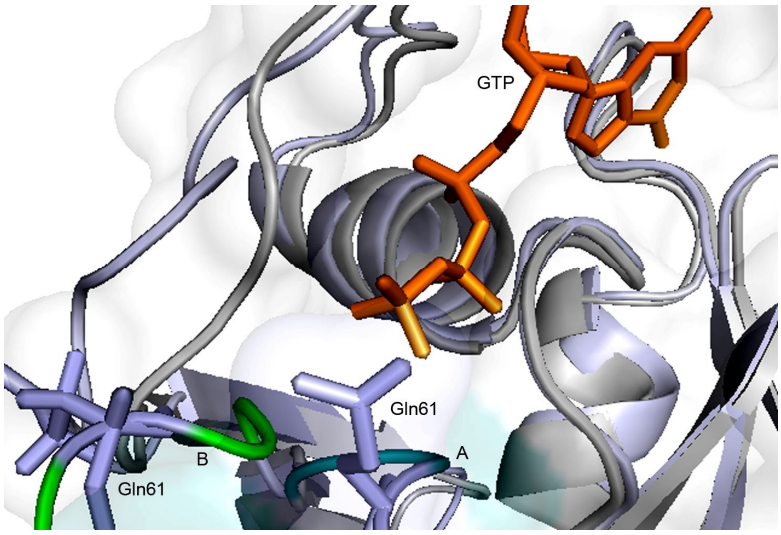


Figure 33: Gln61

Both Gln61 and Glu62 are highly conserved in the small nucleotide binding proteins. In the triphosphate bound conformation of Ras protein, the γ-carboxamide of Gln6l, was initially suggested to activate Wat175 (Pai *et al*.,

1990) although it has been seen later that glutamine is a poor general base and argued that Gln61 is not required to fix this water molecule at the attacking position (see refs in Glennon *et al.,* 2000).

Crystal structure of Ras•GAP complex reveals that Gln61 points toward the phosphate chain of the nucleotide and is stabilized in its orientation by a hydrogen bond with the main chain carbonyl group of the invariant Arg789 (Scheffzek *et al*., 1997). Ras Gln61 is coupled to other residues in the P-loop, switch-I, β3, and switch-II regions and seems to play an important role in the preorganization of active site since mutation of this residue destroys catalysis and leads to major changes in the interaction of Ras in the Ras•GAP complex (Shurki and Warshel 2004). Here we report no estabilize trends in solvation energy, in agreement with previous electrostatic calculations (Glennon et al 2000).

Glu62

Interaction of Ras Glu62 with (GAP) Arg749 was suggested to stabilize the switch-II region (Scheffzek *et al.*, 1997) Our results confirm the participation of this residue in the stabilization of both C and D groups structures.

Glu63

Ras Glu63 is an important mutagenic residue (Polakis & McCormick 1993). The finding that a E63K mutation partially activates the transforming potential of Ras p21 (Fasano *et al*., 1984) seems to implicate this residue, at least indirectly, in GTP hydrolysis (Pai *et al*., 1990; Shurki & Warshel 2004) though the role of Glu63 is debatable (Privé *et al*., 1992).

Ras Glu63 interacts with (GAP) Arg903, stabilizing the Ras switch-II region (Scheffzek *et al*., 1997). Arg903 (turn L1c of GAP) is vital part of highly conserved FLR fingerprint regions and has an indirect role in catalysis, whereby it stabilizes the position of the crucial arginine *i.e* $Arg789_{GAP}$ (Scheffzek *et al*., 1997) and mutating R903K results in only a small reduction of GAP activity. Simulation studies shown that Ras Glu63 becomes more stable upon binding to GAP (Bonet *et al*., 2006), and this extra-stabilization is partially attributed to the conformational changes (Bonet *et al*., 2006). In this study we also confirm the previous studies and observed marked stability at this position in Ras•GAP complex than GTP and GDP bound conformations of Ras protein..

In Ras•PI3K-γ complex crystal structure, $Lys234_{PI3K-\gamma}$ makes hydrogen bonds with Ras switch-II region residues $Glu63_{Ras}$ and $Tyr64_{Ras}$ (Pacold *et al*., 2000). While in this study we observed marked thermodynamic stability at position-63 (Glu63) in Ras•PI3K-γ complex than Ras•GTP conformation and we assume that this extra-stabilization in Ras•PI3K-γ complex is partially attributed to the conformational changes after complex formation. Crystal structure of Ras•PLC-ε complex reveals that a polar contact is made between RA2 region $Lys2154_{PLC-\varepsilon}$ and $Glu63_{Ras}$ (Bunney *et al*., 2006). Some how similar interaction as discussed above with Ras•PI3K-γ complex. Thus, the crystal structure of Ras•PLC-ε (RA2/ras) complex reveals unique properties compared to other RA/RBD interactions with Ras protein, particularly, the involvement of Ras switch-II region. This binding to switch-II region in Ras•Effector complex is a mutual property in RA2 and PI3KγRBD, while, in complexes formed with effectors other than RA or RBD, only switch-I region was found interacting (Bunney *et al*., 2006). More, in this study we also observed thermodynamic stability in the solvation free energy values in Ras•PLC-ε complex than Ras•GTP bound active state and this could be attributed to the stabilization by Lys2154.

Strikingly we found Glu63 more stabilized in all Ras•GEF complexes than Ras•GTP bound state though GTP bound Ras was more stable than GDP bound state. Yet the marked stability was biased towards Ras_{A59G}:SOS_{cat} complex. After Ras_{A59G}:SOS_{cat} complex we found Ras_{Wild}:SOS_{cat} complex then $Ras_{Wild.GDP}$:SOS_{cat} and lastly with Ras_{Y64A}:SOS_{cat} complex.

Tyr64

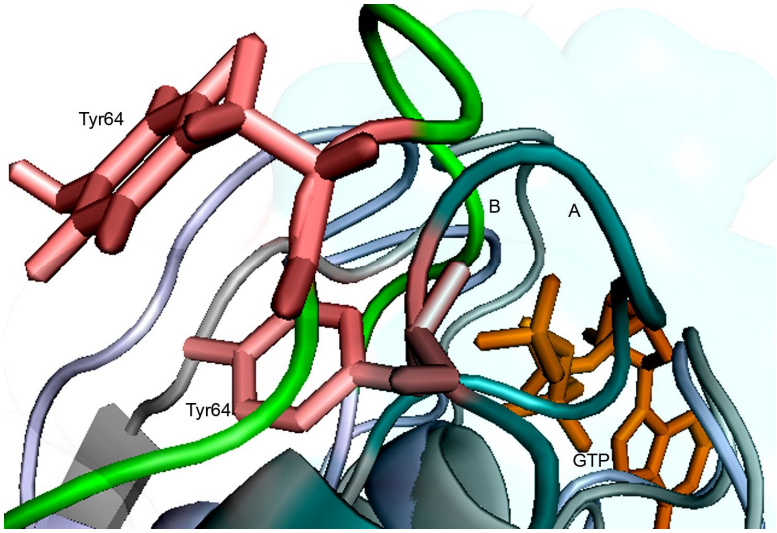


Figure 34: Tyr64

In the partner-unbound GTP-bound structure of Ras the side chains of Glu37, Tyr64, Arg68, Tyr71, Leu56, Ala59, and Thr58 form a pocket that is stabilized by van der Waals interactions and water-mediated hydrogen bonds. The main chain carbonyl of Ile36 is interacting with the hydroxyl side chain group of Tyr64 (Pai *et al*., 1990). In the RasA59G GppNp structure, this pocket unfolds as a consequence, the side chain of Arg68 adopts a new conformation which is incompatible with the conformations of Tyr64 and Tyr71 as observed in wild-type Ras. Consequently in RasA59G mutant, these two tyrosines (Tyr64 and Tyr71) change their conformations and point phenol group to the solvent (Hall *et al*., 2002). More, the structural arrangements that were observed were predicted by the MD simulations observed are all depicted in the structure predicted by the MD simulations (Ma and Karplus 1997). In RasA59G, Tyr64 which is essential for SOS-binding by Ras, is displaced 6Å when compare it in the Ras•SOS structure and adopts a position that inhibits the docking of the two proteins (Hall *et al*., 2002). More mutation Ras_{Y64A} disrupts Ras•SOS interaction (Hall *et al*., 2001).

In Ras•GAP complex, Ras Tyr64 participates in the formation of the hydrophobic interface between Ras and GAP-334 and forms a polar contact with the main chain carbonyl group of $Leu902_{GAP}$ and makes van der Waals or hydrophobic interactions with $Leu910_{GAP}$ on α6c of GAP (Scheffzek *et al*., 1997). More, Ras mutational studies shown that Tyr64 can be mutated to Phe (Y64F) without affecting Ras•GAP interaction but not to Y64E (Marshall & Hettich 1993).

In Ras•GEF complex, $Tyr64_{Ras}$ is located at the interface between the two proteins and forms part of a cluster of hydrophobic interactions between Ras and distal site of SOS (Hall *et al*., 2001, Margarit *et al*., 2003). Addition of stoichiometric amounts of Ras_{Y64A}•GTP significantly accelerates the SOS_{cat} catalyzed nucleotide release (Sondermann *et al*., 2004).

PI3K-γ Phe221 and Lys234 interacts with $Tyr64_{Ras}$ (Pacold *et al*., 2000). Cα atoms RMSD values showed that the conformation of Ras•GTP in the Ras•PI3K-γ and Ras_{Y64A}**:**SOS_{cat}**:**Ras_{Wild} complexes is essentially the same (Margarit *et al*., 2003). In this study we could not see appreciable difference in the thermodynamic pattern in Ras alone vs. partner bound forms. However, we found $Tyr64_{Ras}$ estabilized around 1-kcal mol^{-1} in Ras•GAP and Ras•GEF (Ras_{Wild}**:**SOS) complex.

4.6.8 Helix 2 (α 2)

Met67

In GTP bound conformation, Ras_{Met67} takes active part in the formation of interface between distal Ras and the REM (stands for Ras Exchanger Motif) Domain of SOS_{cat}, REM domain is located in the N-teminal of SOS_{cat}. Key features of the SOS induced conformational changes include the movement of Ras switch-I away from the nucleotide binding site and a restructuring of Ras switch-II that results in the placement of the methyl group of Ala59 of Ras at the site where Mg^{2+} is normally bound. Ras_{Met67} packs against a hydrophobic crevice formed by SOS $Leu687_{REM}$ (which is also part of the binding site for switch-I) and $Arg688_{REM}$. Thus REM is responsible for the displacement of the Ras switch-I region away from the nucleotide (Margarit *et al*., 2003). In this study we were failed to see any significant change in the solvation free energy values in Ras vs. partner bound state. We have not observed ignificant solvation changes for the residue.

Arg68

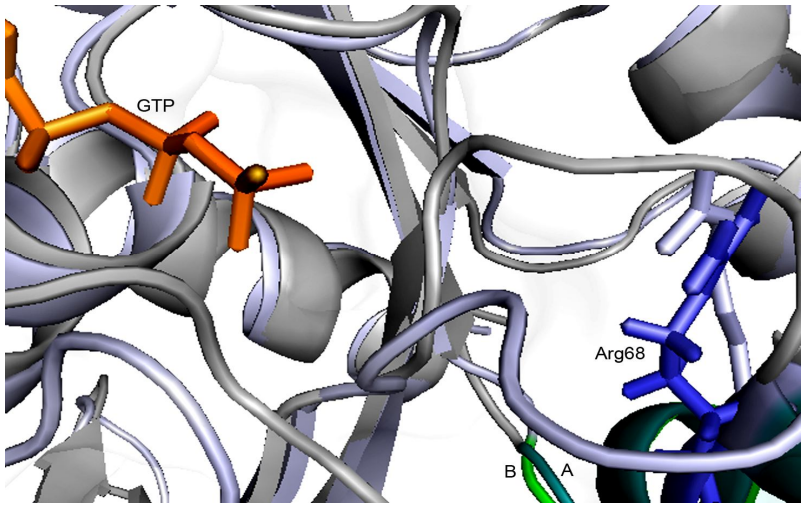


Figure 35: Arg68

In partner-unbound wild type Ras, Arg68 stabilizes the N-terminus of the switch-II region (60–62) through direct or water-mediated hydrogen bonds. In Ras_{A59G}, Glu37 makes hydrogen bonds with Arg68 and van der Waals interactions with Tyr71, which protects it partially from the solvent (Hall *et al*., 2002). Where as in Ras_{Q61G} mutant, the Arg68 side chain shifts towards the core of the protein. The restructuring of Arg68 enables it to make direct hydrogen bonds with the carbonyls of Ras Ala59 and Gly61 and the side chain of Gln99 and forces Ras Tyr71 to be solvent exposed (Ford *et al*., 2006). Thus conformational changes displayed in both Ras_{A59G} and Ras_{Q61G}

mutant protein have some how similar restructuring effect of Arg68. In our calculation we compare the solvation energy values of Ras_{wild}•GTP with Ras_{A59G}•GDP and we found that both the structures showing analogous values. More, Ras_{wild}•GTP vs. Ras_{Q61L}•GTP (PDB: 721P) shows loss of stability of around 3-kcal mol^{-1} in the mutant form. When comparing Ras_{wild}•GTP vs. Ras_{Q61V}•GTP (PDB: 2RGC), Ras_{Q61I}•GTP (PDB: 2RGA) show gain of stability around 2-kcal mol^{-1} in the mutant form while Ras_{A59G}·GTP (PDB: 1LF0) show gain of stability around 4-kcal mol^{-1}. While in Ras_{wild}•GTP vs. Ras_{Q61G}•GXP (X= T and D) mutants, we could not see any significant change, less than 1-kcal mol^{-1}, in the thermodynamic energy. More in Ras_{wild}•GTP vs. Ras_{A59G}•SOS•GTP complex shows marked ignificant in the Ras partner bound conformation. In general we found a strong thermodynamic instability trend in Ras•GEF complex structures. Ras vs. Ras•GAP complex shows marked instability, suggesting a role in GDP binding.

Similarly Ras vs. Ras•Effector complex structure with Ras•PI3K-γ shows ignificant of 2-kcal mol^{-1}. While Ras•PLCε and Ras•RalGDS complex shows gain of stability.

Asp69

This residue is out of the active site region but presentes interesting trends for PPI. The $Ras_{A59T/G12P}$ •GTP (PDB: 521P) mutant was more destabilize than the wild type thus this mutation has destabilization effect on the system. The Ras_{A59G} mutant in the GDP bound conformation was more unstable than GTP bound conformation, but Ras_{A59G} mutant form gain extra stability around 8-kcal mol^{-1} in the partner bound state with SOS (Ras_{A59G}:SOS_{cat}:Ras_{A59G}). More, we found that Asp69 in wild type Ras complexed with SOS is also stable (Ras_{Wild}:SOS_{cat}:Ras_{Wild}), except Ras_{Y64A} mutant where there is no difference in the solvation free energy in Ras•GTP alone vs. Ras•GEF (Ras_{Y64A}:SOS_{cat}:Ras_{Wild}) complex.

Next we compare the thermodynamic energies of Asp69 in wild type Ras with the Ras•Effector bound state, these results were also interesting as in all four cases that we analysed Asp69 was different than Ras alone active conformation. We found Ras•PLCε, Ras•Byr2RBD, and Ras•GAP complexes were more stable than Ras_{Wild}•GTP while Ras•RalGDS and Ras•PI3K-γ complexes were unstable at position Asp69.

More, we also analyzed this position in Ras_{Wild} type vs. Ras_{Q61L}•GTP (PDB: 721P) show loss of stability around 2-kcal mol^{-1} in the mutant form. Similarly, in this study we came to know that Ras_{wild}•GTP vs. Ras_{Q61K}•GTP (PDB: 2RGB), Ras_{Q61I}•GTP (PDB: 2RGA), and Ras_{Q61V}•GTP (PDB: 2RGC) mutants show gain of stability respectively. Since all these mutants are in the active state conformation hence the thermodynamic phenomenon that observed are comparable more accurately in the active state conformation. Moreover, upon comparing the solvation free energies of Ras_{wild}•GTP vs. Ras_{Q61G}•GXP (X= T and D) mutants, we could not see any significant change in the solvation free energy values in the active conformation, though inactive conformation (GDP bound state) was slightly more unstable (around 1-kcal mol^{-1}).

Gln70

No ignificant trend was found for this residue, although some extra stability os found in the inactive (GDP) state, suggesting a pro-OFF character..

Tyr71

As we discussed that $Glu37_{Ras}$ is solvent exposed in wild-type Ras, whereas in Ras_{A59G}, Glu37 makes hydrogen bonds with Arg68 and van der Waals interactions with Tyr71, which protects it (via bulky phenol group) partially from the surrounding solvent (Hall *et al*., 2002) while in Ras_{Q61G} mutant Glu37 does not make any contact with Arg68 (Ford *et al*., 2006) and due to this we found that position gain stability in Ras_{A59G} (GDP bound state: state-B). In this study we found Tyr71 thermodynamic stability in Ras_{A59G} mutant in both active and inactive conformation. Moreover, the GDP bound conformation was more stable than the GTP bound Ras_{A59G} mutant. Interestingly the Ras_{A59G}•SOS complex does not show any difference in the thermodynamic values.

Tyr71 in Ras_{G60A} mutants, show thermodynamic stability in the inactive (diphosphate bound) state than the active (triphosphate bound) state. Same phenomenon was also observed in the Ras_{Q61G} mutants.

Phe75

We could not find any previous reference regarding this position in Ras protein, yet thermodynamic values of this position grab our attention. Interestingly this position shows a mixed trend. In Ras_{A59G} mutant, the GDP bound conformation was more stable than the GTP bound Ras_{A59G} mutant

and Ras wild type protein. More, upon comparing the Ras_{G60A}·GTP structure vs. Ras_{G60A}·GDP and Ras_{Wild}·GTP structures, mutant GTP bound state is showing more stability around 2-kcal mol^{-1}. Similarly, Ras_{Q61L} mutation also show gain of stability when compare with the Ras wild type and with Q61V, Q61I, and Q61K mutants and that Ras Q61L has one of the highest transformation efficiencies of any gain-of-function mutant at position 61 (Buhrman *et al*., 2007). More, we found thermodynamic stability around 2-kcal mol^{-1} at Phe75 in Ras_{Q61G} weakly transforming mutants (both in active and inactive state) than Ras wild type. More, we found Ras_{E31K}•RalGDS and Ras_{Y64A}**:**SOS complexes gain stability around 2-kcal mol^{-1} at Phe75.

4.6.9 Helix 3 (α 3)

Lys88

This position is showing a dramatic change for different mutant structures. In Ras G12V, G12D, G12P mutants in active state this position show gain of stability. Similarly in the $Ras_{Y32C/C118S}$ double mutant structures in active and inactive state this position show gain of stability than the Ras wild type, though the inactive conformation show more stability at this position and Ras D38E mutant (PDB: 221P) also displays thermodynamic stability with respect to Ras wild type. An exciting case is when compare Ras wild type vs. Ras_{A59G} mutants. Ras_{A59G} mutants in both the state (active and inactive) and partner bound state (Ras_{A59G}**:**SOS) show gain of stability than the Ras wild type, though this stability is not uniform and active state conformation show greater stability at Lys88. We then analyzed the effect of Ras_{G60A} mutation on the stability of Lys88 and we found that Ras_{G60A}•GTP mutant show an obvious gain of stability, while the diphosphate conformation does not show any difference than the wild type Ras protein. Similarly we compared the stability of Lys88 in Ras wild type vs. Ras Q61H, Q61L, Q61I, Q61K, Q61G, and Q61V mutants. All the Gln61 mutants show gain of stability at Lys88 but Q61I mutant show remarkable gain of stability.

Interesting case was in Q61G mutants, where we compare the active and inactive state structures than the wild type Ras and inactive conformation gains more stability than active conformation in mutant as well as in wild type structure. We compare the wild type Ras•GTP with the inactive K117F mutant structure, and in this case the mutant inactive conformation gains extra stability than the wild type active conformation.

We also compare the stability of Ras Lys88 between the Ras alone vs. partner bound state (effectors and regulators). All the Ras-partner structures we analyzed, were showing gain of energy, remarkably Ras_{Wild}**:**SOS complex showing gain of 12-kcal mol^{-1} on one hand and on the other hand Ras_{Y64A}**:**SOS complex was showing loss of stability. Moreover, we previously showed that Lys88 becomes more stable after complex formation between Ras•GAP (Bonet *et al*., 2006) and we also observed the same phenomenon.

Glu91

In this study we found Glu91 tends to be more stable in the triphosphate bound conformation than diphosphate conformation. In Ras G12D, G12P mutants in active state and in $Ras_{Y32C/C118S}$ double mutant structure in inactive state this position show loss of stability than the active state conformation of wild type Ras and $Ras_{Y32C/C118S}$ double mutant active state. Similarly, Ras_{D38E} mutant displays thermodynamic instability than wild type Ras. An exciting case is when compare Ras wild type vs. Ras_{A59G} and Ras_{G60A} mutants. Both the mutants in active state show thermodynamic estability than the wild type active state conformation as well as respective mutations in inactive state. Moreover, we compared the thermodynamic stability of Glu91 in Ras wild type vs. Ras_{Q61V} mutant where Gln61 mutant show loss of stability at Glu91, but Ras_{Q61G} mutant in inactive conformation show gain of stability than the active state and Ras wild type active conformation. We compare the wild type Ras•GTP with the inactive K117F mutant structure at position-91, and in this case the mutant inactive conformation show gain of stability than the wild type active conformation.

We then compare the thermodynamic values of Ras wild type active conformation with the partner bound conformations and found that Glu91 becomes more destabilize upon complex formation with Ras•Byr2RBD and Ras•GAP complexes.

Asp92

The solvation free energy values of this position grab our attention. This position is within 12.5Å distance from the heavy atom of GTP analog in the crystal structure. In Ras G12D, G12P, D38E, Q61H, Q61L, Q61I, Q61K, Q61V mutants in active state and in the G12V, and K117R mutant structures in inactive state this position show gain of stability around 2-kcal mol^{-1} than the active state conformation of wild type Ras.

In double mutant structures $Ras_{G12P/A59T}$ double mutant in active state conformation shows loss of stability around 2-kcal mol^{-1} while in $Ras_{Y32C/C118S}$ double mutant in either state (active as well inactive state) Asp92 show stabilization trend, but the $Ras_{Y32C/C118S}$ active state conformation show remarkably gain of stability around 12-kcal mol^{-1}. Another interesting case regarding Asp92 is in Ras_{Q61G} mutant structure, where mutant structure in either state (active as well inactive state) show gain of stability but interestingly inactive conformation of Ras_{Q61G} is more stable than the active state conformation.

We then proceed to compare the stability pattern in Ras vs. Ras-partner bound state, remarkable Asp92 shows marked stability in GEF bound complex that we analyzed. The stability flows less to more in a series Ras_{Y64A}**:**SOS_{cat} > Ras_{Wild}**:**SOS_{cat} >> Ras_{A59G}**:**SOS_{cat} complex structures. We then compare the thermodynamic values of Ras Asp92 active state conformation with the partner bound conformations and found that Asp92 gain more stability upon complex formation with Ras•RalGDS, Ras•Byr2RBD and Ras•PI3K-γ complexes, while Asp92 becomes more destabilized in Ras•PLCε and Ras•GAP complexes.

Arg97

We could not find any previous reference regarding Arg97 position in Ras protein, even we found it very relevant for overall stability of all different structures. In Ras G12P, Q61V, and Q61L mutants in active state and in the G12V, and K117R mutant structures in inactive state this position show gain of stability around 2-kcal mol^{-1} than the active state conformation of wild type Ras while in active state mutant G12R Arg97 show loss of stability. In double mutant structure $Ras_{Y32C/C118S}$ inactive state (diphosphate bound) conformation show gain of stability around 2-kcal mol^{-1}. Another interesting case regarding Arg97 is in Ras_{A59G} and Ras_{G60A} mutant structures, where mutant structures in either state (active and inactive state) show gain of stability. Similarly when we compare Ras_{Q61G} structures with Ras wild type active state conformation we found Ras_{Q61G} active state conformation thermodynamically more stable than the wild type and inactive state of the mutant protein.

We then proceed to compare the stability pattern in wild type $Ras_{Active\ State}$ vs. Ras-partner bound state, remarkable Arg97 shows gain of stability in GEF bound complex and Ras_{Wild}**:**SOS_{cat} complex show overall more stability than the Ras_{Mutant}**:**SOS_{cat} structures. Among Ras•Effector complexes, we

found Ras•PLCε complex show gain of stability at Arg97 after complex formation.

4.6.10 Beta sheet 5 (β 5)

Asn116

The side chain amide of Asn116, which has been proposed to be involved in the binding of the O6-oxygen, is seen to make strong hydrogen bonds to the side chain of Thr144 and to the main chain oxygen atom of Val14 (Pai *et al*., 1990). Thus the main function of Asn116 is to tie together the three elements which are involved in nucleotide binding that are the phosphate binding loop $_{10}$GXXXXGKS, the $_{116}$NKXD and the $_{146}$SAK motifs (Pai *et al*., 1990). In this study, Asn116 show loss of stability around 2-kcal mol^{-1} in Ras•Byr2RBD complex than the wild type Ras active state structure.

Lys117

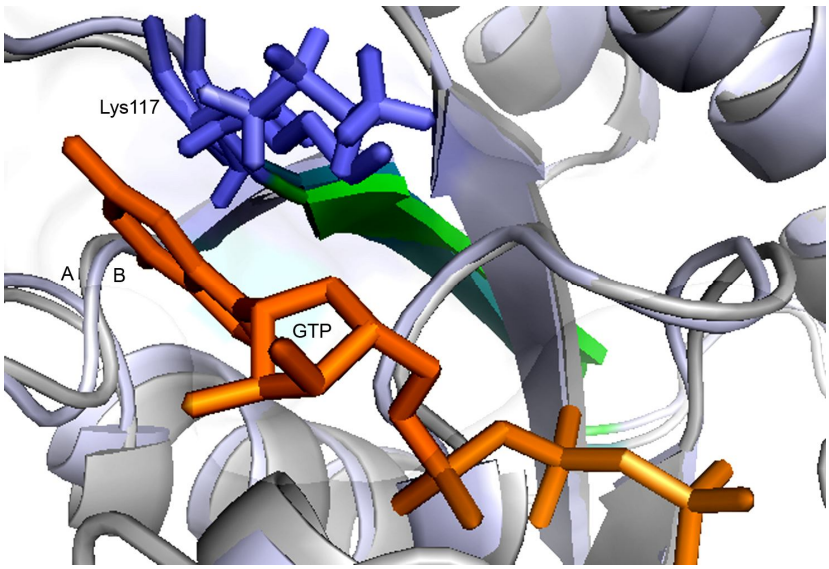


Figure 36: Asp117

Lys117 make hydrophobic interactions with the guanine base (Pai *et al*., 1990). Crystal structure shows that Lys117 links the P-loop and the $_{116}$NKXD motif by binding to the main chain carbonyl of Gly13. A further albeit weak contact of this amino acid is with O4' of the ribose (Pai *et al*., 1990).

Our calculations of Lys117 show that Lys117$_{Ras}$ is thermodynamically more stable in the triphosphate active state than inactive state (diphosphate bound state), although there is also a clear tendency to estability the residue by GAP, suggesting a role in reactivity. In Ras G12R, G12P, D38E, Q61K, and Q61L mutants in active state and in the G12V, and K117R mutant structures in inactive state this position show loss of stability around 2-kcal mol^{-1} than

the active state conformation of wild type Ras while active state G12D and Q61I mutants show gain of stability. In double mutant structure $Ras_{G12P, A59T}$ active state (triphosphate bound) conformation show loss of stability around 2-kcal mol^{-1}. Another interesting case is in Ras_{A59G} and Ras_{G60A} mutant structures, where both mutant structures show loss of stability. Ras_{G60A} mutant structures show estability phenomenon both in active and inactive state while Ras_{A59G} mutant only inactive state shows estability than the wild type Ras and the active state of the same mutation.

We then proceed to compare the stability pattern in wild type $Ras_{Active\ State}$ vs. Ras-partner bound state, remarkable Lys117 shows loss of stability in all Ras•GEF bound complexes. Among Ras•Effector complexes, we found Ras•PLCε, Ras•RalGDS, and Ras•Byr2RBD complexes show loss of stability at Lys117 after complex formation.

4.6.11 Flexible turn 8 (L 8)

Asp119

The carboxylate group of Asp119 makes four hydrogen bonds. One oxygen interacts with the exocyclic amino group and Wat292, the other oxygen binds to the endocyclic nitrogen N1 and to the hydroxyl of Ser145 (Pai *et al*.,1990). Ras_{D119A} mutation results in 20-fold reduction in affinity of binding to both GDP and GTP (Sigal *et al.,*1986). Mutating D138N (equivalent to Asp-119 in ras p21) of the ET-Tu-GDP complex structure results a change in the nucleotide specificity of EF-Tu from GDP to xanthosine 5'-diphosphate (Hwang & Miller 1987). Our calculations show a similar trend that for Lys118, suggesting a role in reactivity for this residue. In Ras G12P, Q61L, Q61I, and Q61V mutants in active state and in the G12V, and K117R mutant structures in inactive state this position show loss of stability around 2-kcal mol^{-1} than the active state conformation of wild type Ras while active state Q61H mutant show gain of stability. In double mutant structure $Ras_{Y32C/C118S}$ active state (triphosphate bound) conformation show loss of stability around 2-kcal mol^{-1}. Other interesting cases are in Ras_{G60A} and Ras_{Q61G} mutant structures, where both mutant structures show loss of stability. Ras_{G60A} mutant structures show instability phenomenon both in active and inactive state, yet the active conformation shows more instability than inactive conformation while Ras_{Q61G} mutant only inactive state shows estability than the wild type Ras and the active state of the same mutation.

We then proceed to compare the stability pattern in wild type $Ras_{Active\ State}$ vs. Ras-partner bound state. Remarkably, Asp119 show loss of stability in all GEF bound complexes. Among Ras•Effector complexes, we found Ras•PLC ε, Ras•RalGDS, and Ras•Byr2RBD complexes show loss of stability at Asp119 after complex formation. The instability was also confirmed in $Ras_{Active\ State}$ vs. Ras•GAP complex structure.

Arg123

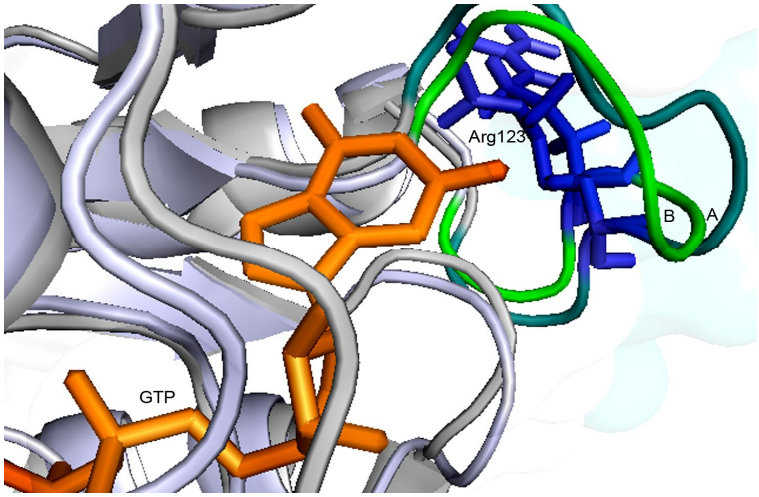


Figure 37: Arg123

Arg123 shows an overall stability (it appears to contribute to the stability of all conformations and bound modes) but the strength of stability varies depending on the nature of the conformation that adopted. Ee found that $Arg123_{Ras}$ is thermodynamically more stable in G12P mutant in active state while G12V (inactive state), and Q61K (active state mutant) show loss of stability around 2-kcal mol^{-1} than the active state conformation of wild type Ras. Moreover, Ras_{G60A} mutant structures show instability phenomenon in active state and on contrary stability in inactive state.

We then proceed to compare the stability pattern in wild type $Ras_{Active\ State}$ vs. Ras-partner bound state, remarkable $Arg123_{Ras}$ shows gain of extra stability in Ras•GEF (wild type) complex and also with Ras•GAP complex while Ras_{G12V} •PI3K-γ -complex show loss of stability.

4.6.12 Beta sheet 6 (β 6)

Glu143

Glu143 shows a strong overall stability but the strength of stability varies depending on the nature of the conformation that adopted. Upon comparison we found $Glu143_{Ras}$ is thermodynamically more stable in triphosphate bound

conformation than the diphosphate bound conformation. $Glu143_{Ras}$ in K117R mutant shows more stability in inactive state while Q61K (active state mutant) show loss of stability around 2-kcal mol^{-1} than the active state conformation of wild type Ras. Moreover, Ras_{G60A} mutant structures show instability (around 4-kcal mol^{-1}) in active state and on contrary stability in inactive state.

We then compare the stability pattern in wild type $Ras_{Active\ State}$ vs. Ras-partner bound state, $Glu143_{Ras}$ shows loss of stability in Ras•GEF (wild type and Ras_{A59G} mutant) complex and in Ras•GAP complex while Ras•Byr2RBD complex show gain of stability.

4.6.13 lexible turn 10 (L 10)

Ala146

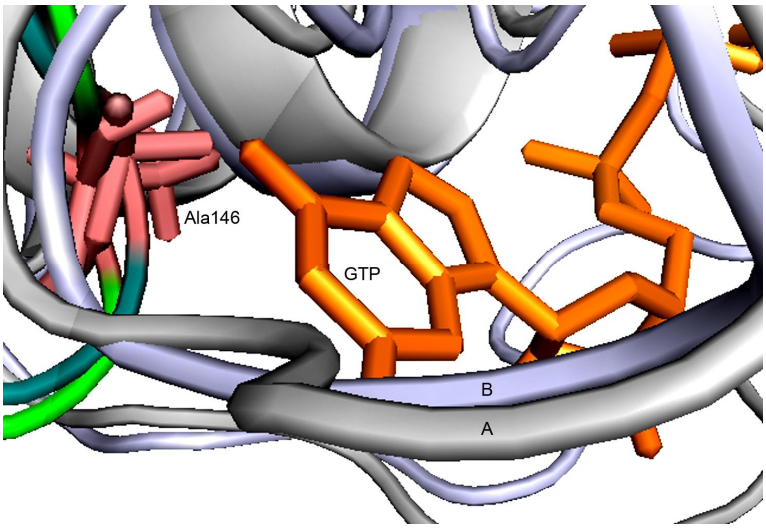


Figure 38: Ala146

The keto group at position-6 of the guanine base makes a hydrogen bond with the main chain NH of Ala146 (Pai *et al*., 1990). Ras_{K117R} show loss of stability around 2-kcal mol^{-1} than the active state conformation of wild type Ras, although no estability trends are observed.

Lys147

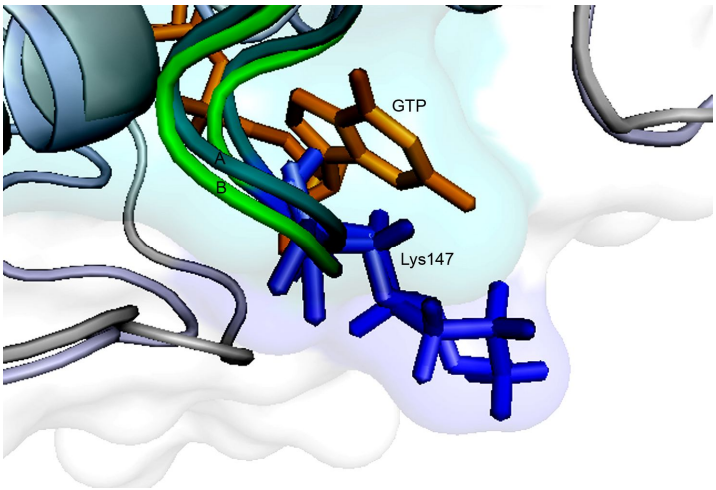


Figure 39: Lys147

$Lys147_{Ras}$ in triphosphate bound active state is stable around 2-kcal mol^{-1} than diphosphate bound conformation. Our comparative analysis shows an overall instability pattern when comparing the thermodynamic values with Ras wild type active state vs. other conformations and thus level of instability varies depending on the nature of the mutation and conformation that adopted and that in general triphosphate bound conformation of $Lys147_{Ras}$ is more estability than the diphosphate bound conformation.

$Lys147_{Ras}$ in G12D, G12P, G12R, G12P, D38E, Q61H, and Q61L active state mutants and G12V, and K117R inactive state mutants, solvation free energy values show loss of stability. Moreover, Ras_{G60A}, and Ras_{Q61G} mutant structures show instability in both conformations yet the active state mutants are more destabilized than the inactive state conformation (estability was more than 9-kcal mol^{-1} and 4-kcal mol^{-1} respectively). On contrary Ras_{A59G} mutant structures show more estability in inactive state than the active state of the mutant structure, the same situation was observed in double mutant structure ($Ras_{Y32C/C118S}$).

We then compare the stability pattern in wild type $Ras_{Active\ State}$ vs. Ras-partner bound state and found that $Lys147_{Ras}$ shows loss of stability in Ras•GEF bound complexes. Similarly we also found that Ras•GAP, Ras•Byr2RBD, Ras•PLC-ε, Ras•PI3K-γ, and Ras•RalGDS complexes show loss of stability.

Arg149

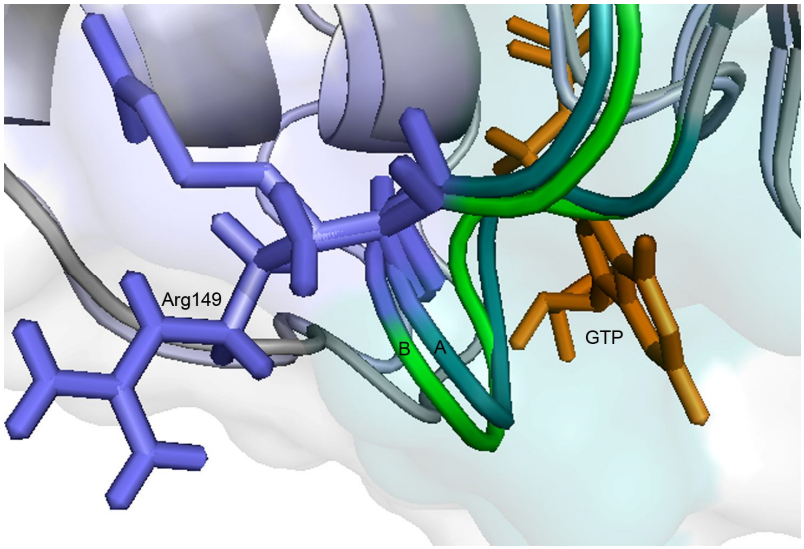


Figure 40: Arg149

$Arg149_{Ras}$ makes hydrophobic interactions with the guanine base (Pai *et al*.,1990). $Arg149_{Ras}$ in triphosphate bound active state is around 2-kcal mol^{-1} more stable than the diphosphate bound conformation. Our comparative analysis show more an instability pattern when comparing the thermodynamic values with Ras wild type active state vs. other conformations. We found thermodynamic instability in Ras Q61I, Q61V active state and K117R inactive state mutant structures and free energy stability in Ras G12P active state mutant. Moreover, Ras_{G60A}, and Ras_{Q61G} mutant structures show instability in both conformations although in Ras_{G60A} mutant structure inactive state mutant is more destabilized than the active state conformation (estability was more than 4-kcal mol^{-1}).

Similarly, Ras_{A59G} mutant structures show over 5-kcal mol^{-1} destability in inactive state than the active state of the mutant structure and wild type Ras, and Ras double mutant structure ($Ras_{Y32C/C118S}$) was also portraying similar scenario. On the whole in Ras•GEF bound complexes (active state: Ras_{A59G}:SOS_{cat} 3.2Å, Ras_{Wild}:SOS_{cat} 2.7Å, and Ras_{Y64A}:SOS_{cat} 2.18Å inactive state: Ras_{Y64A}:SOS_{cat}) we found free energy instability than Ras wild type active state, intriguing Ras_{Y64A} •GEF mutant in either GTP or GDP bound state were showing around 4-kcal mol^{-1} thermodynamic estability. Albeit, we found thermodynamic stability in Ras_{G12V} •PLC-ε effector bound complex than Ras wild type active state conformation.

Gln150

Our study shows that Ras Gln150 may not be playing any important role in the enzyme function, stability or conformation coupled changes. Few structures that were displaying significant changes were Ras G12P, and Q61L active state mutant, showing thermodynamic estability around 2-kcal mol^{-1}. Moreover, we also found that Ras_{G12V} •PLC-ε (Ras effector complex) and Ras•GEF inactive state complex (Ras_{Wild}:SOS_{cat}:Ras_{Y64A} •GDP) show thermodynamic estability around 2-kcal mol^{-1}.

4.6.14 Helix 5 (α 5)

Glu153

We found that Glu153 in general, tend to be stable and few structures showing importance of position. Our calculations show thermodynamic estability in Ras G12P, and Q61V active state mutants, while thermodynamic stability in Ras G12V mutant structures.

Furthermore, we found Ras_{G12V}·PI3K-γ -complex also shows thermodynamic instability when compared to Ras wild type active state while on the other hand we found that Glu153 is very stable in all Ras•GEF active and inactive state complexes.

Asp154

We could not find any previous reference regarding Asp154 position in Ras protein, yet thermodynamic values of this position grab our attention. In most of the cases that we studied Asp154 shows thermodynamic estability than the wild type Ras active state protein.

We found that Ras Asp154 in G12P, G12R, Q61H, Q61I, Q61K, and Q61V active state mutants and K117R inactive state mutant, solvation free energy values show loss of stability.

Moreover, Ras_{A59G}, and Ras_{Q61G} mutant structures show instability in both conformations likewise, Ras double mutant structures ($Ras_{Y32C/C118S}$) were also portraying similar scenario. Ras_{G60A} mutant structures show loss of stability, though inactive state was more unstable than the active state of the mutant structure and wild type Ras.

Furthermore, we also compare the stability of $Asp154_{Ras}$ between the Ras alone vs. partner bound state (effectors and regulators) and found that Ras_{G12V} •PLC-ε (over 5-kcal mol^{-1}), Ras_{E31K} •RalGDS, and Ras_{G12V}·PI3K-γ

complex structures show loss of stability than the wild type Ras active state. Here we also report that Asp154 becomes destabilize when Ras protein binds to GEF protein and remarkably, RasGDP•GEF complex show estability around 6-kcal mol^{-1}.

Thr158

We could not find any previous literature citing role of this residue in stability, or conformational changes. Our calculations show that this position becomes gain stability around 2-kcal mol^{-1} when Ras protein mutated to G12R. Moreover, Ras_{Q61G} mutant active state structure shows gain of stability than wild type active state Ras and inactive state mutant structure. More, we also found that Thr158 in Ras_{GDP}•GEF complex becomes destabilize.

4.7 *Protein-Protein interaction (PPIs) and conformational changes in phosphate hydrolyzing proteins*

In our previous work (Bonet *et al.,* 2006), we explore the possibility to find a relationship between reactivity and stability of residues upon formation of the active complex in specific PPIs models (CDK2/Cyclin A and Ras/GAP systems). We differentiated protein–protein interaction process in to three different states of the partners (see Glennon *et al.,* 2000) for the similar labeling scheme). We labeled the unbound partners with letters A and B (see Figure 41). The figure schematically represents the metastability hypothesis during the interaction between the regulatory proteins with a catalytic chain. In the figure, A represents the structure of the unbound protein and A' the structure of the bound protein complexed with B' (or with B in the situation, in the examples that the conformational changes in B are small during the formation of the complex). Both interfacing and active site regions are shown as irregularities on the surface. The interaction process can be in principle divided into two sub-processes, not necessarily real but correct for a thermodynamic treatment. On one hand, a conformational change is suffered by the partners in order to arrive to a ready-to-interact structure. During this process some regions of the two partners (darker regions in the figure) become instable. The second sub-process consists of the actual interaction between residues in the two partners, where the destabilized residues become stable mainly by electrostatic interactions. The dark circle in some of the pictures represents the substrate attached to the catalytic subunit of the complex. Following this hypothesis, the free energy barrier that needs to be surmounted to create an obligate or strong complex would

be smaller than the barrier to create a transient complex, while the reaction in the first case would be more exothermic. Hammond's principle suggests that the transition state for the complex formation will be closer to the reactants for strong complexes, and thus the conformational change that needs to take place will also be smaller in obligated complexes than in transient complexes.

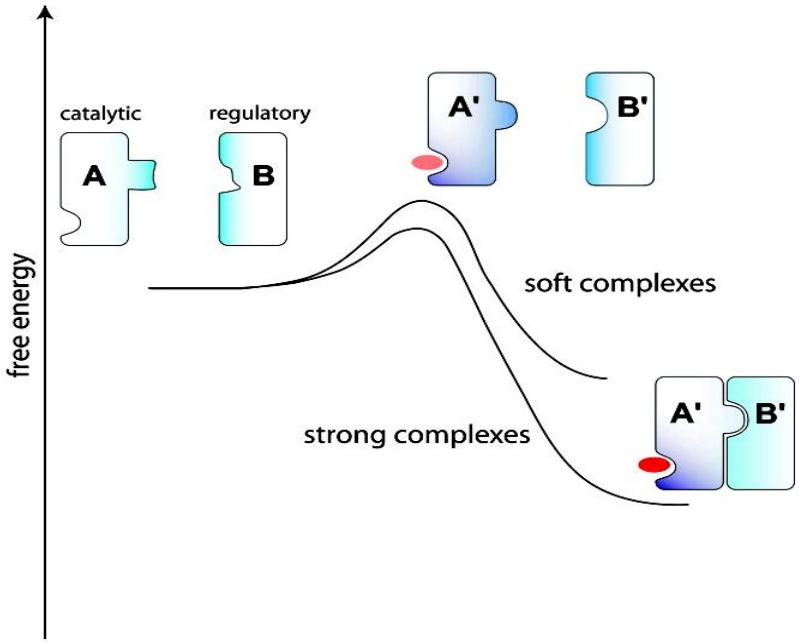


Figure 41: Metastability Hypothesis
Metastability hypothesis during the interaction between a regulatory protein with a catalytic chain (see the text for explanation) Adapted from Bonet *et al.*, 2006

Figure 42**a** shows the differential stability in the protein and water environments for all the residues in the unbound Ras chain (PDB code 5P21: Pai *et al.,* 1990). (Figure 42**b**, **c**) shows this differential stability for the Ras chain in its conformation when bound to GAP (PDB code 1WQ1: Scheffzek *et al.,* 1997), in the absence (Figure 42**b**) and presence (Figure 42**c**) of the partner GAP.

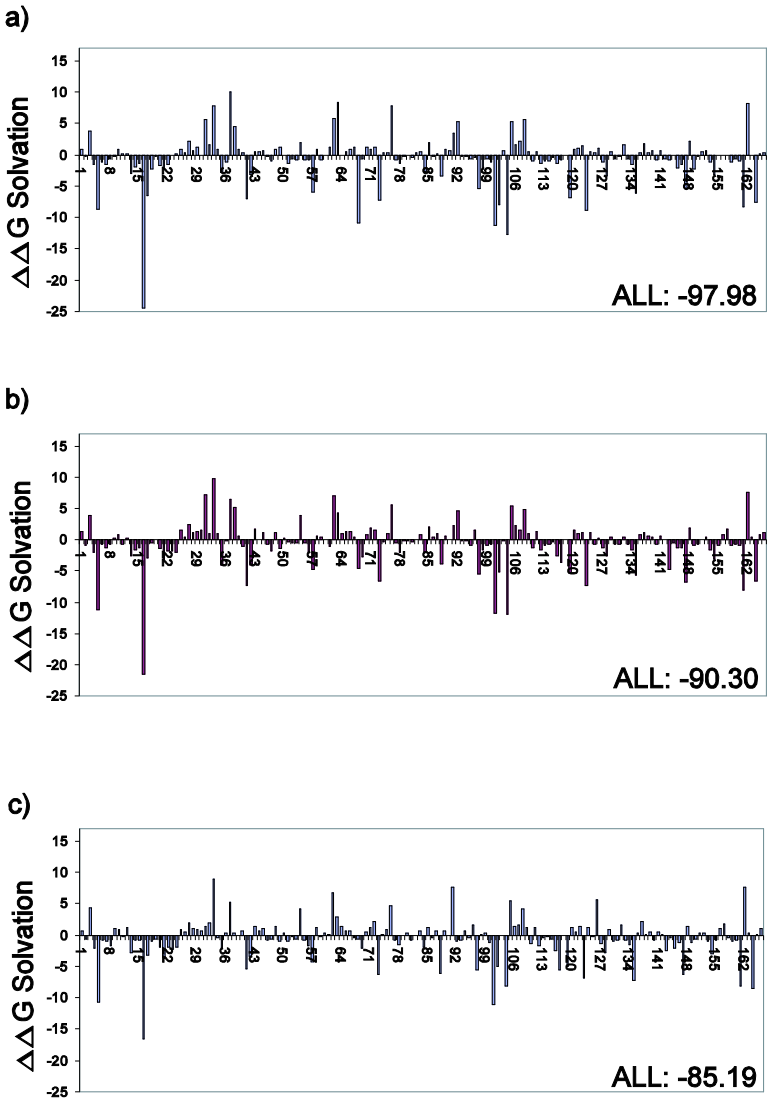


Figure 42: Bar plots showing the stabilization for all residues in Ras in three different conformational states:
(**a**) unbound Ras; (**b**) bound Ras conformation, with calculations performed in the absence of GAP; and (**c**) bound Ras. The "ALL" label refers to the sum of all stabilities.

4.8 General discussion

The work done and presented in this thesis includes the simultaneous discussion of sequence, structure, and function of the protein. It is widely accepted that those residues/positions that play a determinant role in the structure and functional analysis are assumed to be conserved. However, in a similar way to what occurs between the lower sequence conservation with respect to the higher structure conservation, it is also likely that functional conservation involves the need for a more rigorous definition based on

physicochemical properties that go beyond typical classifications of residues (*e.g.* polar-non polar).

Although is this work we were not able to establish a proper quantitative measure of functional classification of residues, the use of PDLD/S-LRA simulations as a compromise between rigor and efficacy has been demonstrated by our group (Bonet *et al.,* 2006, Schepper *et al.,* 2008) and others (Roca *et al.,* 2008).

In this work we evaluated the solvation free energy of individual resideus and compared it to prior knowledge of the ligand and partner status. We searched for the position that candidly play major role in the conformational changes and scrutinize statistically any such role. We performed Wilcoxon Test (see methodology for further information) and compared the Group-A structures with those of Group-B structures (see Table 3). In this comparative study we found number of positions (with respect to PDB: 5P21) that show statistical difference in their median. In a recent RMSD based study of Ras subfamily proteins, it was found that the conformational transition of Ras from the inactive to the active state is triggered by the displacement of switch-II region followed by a rapid movement of switch-I region (Noe *et al.,* 2005; Gorfe *et al.,* 2008) and suggested that the conformational properties of the active and inactive states of Ras protein are not the individual behavior of these switch regions, rather defined by the relative configuration of switch-I and switchII regions (Gorfe *et al.,* 2008). It would be interested to relate our findings to the global dynamics of Ras, using the tools developed in the group (Johnston *et al* 2005; Alcántara and Villà-Freixa, in preparation)

Latter, I proceeded to next layer of complexity of the project, where as per our assumption there must be some residues/positions that determine the fate of conformational changes when interacting with partner(s) protein(s) and stable complex formation. To do so, we merged down the Group-A with Group-C and name the newly formed group as Group-A extended (abbreviated as Group-A ext), similarly merged the Group-B with Group-D and name the newly formed group as Group-B extended. Redo the statistical tests and found the same position (positions 10, 30, 36, 37, 96, 123, 146, 147, and 149) significant with p-value > 0.05. More, the physico-chemical nature of residues at these positions was also diverse (but none of the polar side chain residue) and all belong to different region of the active site (except the switch-II region residues). At this stage we found it difficult to rationalize the complex observed results, but a possible solution would be to unbias the

samples size and to deep into certain mutations at particular points with multiple switch conformations that have been reported (Spoerner *et al.,* 2005; Spoerner *et al.,* 2007).

5 Conclusions

General methodological conclusions:

1. PDLD/S-LRA has proven to provide a semiquantitative picture of the electrostatic fingerprint of the Ras active site. Further work is needed to a) simplify the protocol to run shortest simulations and thus us the protocol in a high-throughput bioinformatics context, and b) while simplifying ensuring that the quality of the results improves.
2. The quality improvement of the PDLD/S-LRA protocol involves the improvement of the PDLD charges by global optimization techniques, adjustments of the Langevin dipoles protocol and LRA simulations using improved for fields. Adun is the key tool in this process, and some advances have already been done in this direction.
3. Improvements in the search for the conformational space for LRA simulations may involve the use of new algorithms for protein global dynamics, also currently being implemented by other members of the group into Adun.

Specific conslusions regarding the Ras active site electrostatic fingerprint:

1. A general trend is found that identifies a global negative electrostatic fingerprint for the overall active site in Ras. Such overall stabilization does not preclude the existence of positive solvation free energy values that are related to functionally relevant residues.
2. There appear to be a list of residues that present a statistically signifficant difference in free energy of solvation (i.e. stabilization) between GDP and GTP bound proteins, indicating a functional role in the GDP/GTP exchange or the protein's phosphate hydrolysis reactivity.
3. The presence of regulators/effectors forming complexes with Ras affects the stability of active site residues. This effect is more apparent during GAP activation for residues critical for catalysis or in Ras interaction with GEFs, foir residues participating in the GDP/GTP exchange.

4. The use of the PDLD/S-LRA based protocol can be extended to protein protein interactions, and Bonet el al 2006 shows that the preorganization hypothesis can be generalized to the study of interaction patches between protein structures. In this way, the concept of active site is related to the concept of hot spots for protein interaction.
5. Despite the encouraging data for the use of electrostatic calculations in assessing functional motifs in proteins, this thesis did not reach the objective of setting up a final protocol useful for bioinformatics studies. More work is needed to provide a general tool for the detection of functional patches in proteins.

6 Appendix

This section contains different subsections with extended material on the Ras protein primary to tertiary structure and its posttranslational modifications that has some relation to the thesis and that will serve to open the possibility of extensions to this work. It also contains sections including scripts and tools used through the thesis, with the aim to serve as inspiration for future research.

6.1 Ras Protein: Post-translational Modification & Membrane Anchorage

There are four mammalian Ras proteins, encoded by three ras genes: H-Ras, N-Ras, K-Ras4A and K-Ras4B. The K-Ras4B is more abundant in mammalian cells, and in the literature is usually referred as K-Ras. It was experimentally shown that both H-*ras* and N-*ras* (Ras isoforms) knockout mice are viable, whereas inactivation of the K-*ras* gene is embryonically lethal (Koera *et al.,* 1997; Esteban *et al.,* 2001).

All Ras isoforms are highly homologous; their G-domain (residues 1–165, which binds guanine nucleotides and is required for the switch function and for effector binding) is nearly identical, display approximately 85% amino acid sequence identity (Shields *et al.,* 2000). On the other hand, their C-termini (last ≈ 24 amino acids, termed as hypervariable region “HVR”) are highly varied between the different Ras isoforms (Lowy and Willumsen 1993). These differences in the HVR of Ras isoforms correlate with functional differences in cells and organisms, leading to the assumption that the individual Ras isoforms have special functions. In principle, the HVR region is further divided into 1) the linker domain and 2) lipid anchor region (the most C-terminal region) (see Figure 43), which contains the sequences responsible for the membrane anchoring of the various Ras isoforms, as well as for intracellular trafficking of Ras protein (Rotblat *et al.,* 2004; Hancock and Parton 2005; Quatela and Philips 2006; Ashery *et al.,* 2006; Matallanas *et al.,* 2006). Membrane affinity of the Ras anchor depends on hydrophobicity of the palmitate and the prenyl groups and the spacing between palmitate and prenyl groups (Gorfe and McCammon 2008). The spatio-temporal organization in cells is largely dependent on both the nature and the dynamics of the association of proteins with specific membranes and

on their targeting to defined domains or regions within these membranes. Ras proteins interactions with the plasma membrane leaflet are essential for their normal signaling and transforming activities (Hancock *et al.,* 1989; Hancock *et al.,* 1990; Choy *et al.,* 1999). Different biochemical and cell biological studies showed that H-Ras and K-Ras are localized in distinct microdomains in the plasma membrane and in endomembranes (Prior *et al.,* 2003).

Despite similarity between different Ras isoforms in their response to hormonal stimuli and signaling, yet there are differences between the isoforms in their activation of specific downstream signaling pathways (Yan *et al.,* 1998; Voice *et al.,* 1999; Walsh and Bar-Sagi 2001), thus it was concluded that different Ras isoforms are functionally not redundant and individual Ras isoforms have special functions (Johnson *et al.,* 1997; Koera *et al.,* 1997). More, data from experiments in which cholesterol was depleted from membrane domains indicate that H-Ras resides in lipid rafts, whereas K-Ras4B is excluded from lipid rafts and ysteine to the disordered plasma membrane (Roy *et al.,* 1999; Prior *et al.,*, 2001). Thus, Ras isoforms specificity is dictated by the C-terminus HVR (Kim *et al.,* 2002).

Different point mutation(s) of Ras proteins results either inactivation or overactivation of Ras protein in different human tumors (Bos 1989; Rodenhuis 1992). A potential mechanism that could give rise to some of these differences is distinct membrane anchorage of each Ras isoform, resulting in a dissimilar spatial organization in the plasma membrane and/or different subcellular localization (Hancock and Parton 2005; Quatela and Philips 2006). Oncogenic Ras can transform cells both in vitro and in vivo (Barbacid 1987), and it was shown that Ras transformed cells contribute to oncogenesis by over activation of several pathways (Bos 1995; Bar-Sagi and Hall 2000; Shields *et al.,* 2000; Downward 2003) and transformation of Ras specific activated pathway(s) is species dependent (Hamad *et al.,* 2002; Lim *et al.,* 2005).

Ras proteins and isoforms have some distinct but common motifs that facilitate their interactions with the membrane *e.g.* conserved C-terminal Prenylation signal '**CAAX**' motif (where 'C' is Cys, 'A' is Ala and 'X' is terminal aminoacid). The '**CAAX**' motif undergoes post-translational prenylation (covalent attachment of farnesyl fatty acid via thio-ether linkage to ysteine residues at or near the C-terminus) to generate S-farnesyl ysteine thioester, followed by proteolytic cleavage of the '**C-A**' peptide bond and methyl esterification of the resulting C-terminal isoprenylated

ysteine in the endoplasmic reticulum (ER) (Casey *et al.,* 1989; Gutierrez *et al.,* 1989; Casey and Seabra 1996; Zhang and Casey 1996; Choy *et al.,* 1999; Colicelli 2004) before transport to the plasma membrane. Although, isoprenylation is regarded more important than palmitoylation for membrane binding and transforming activity of Ras protein (Jackson *et al.,* 1990).

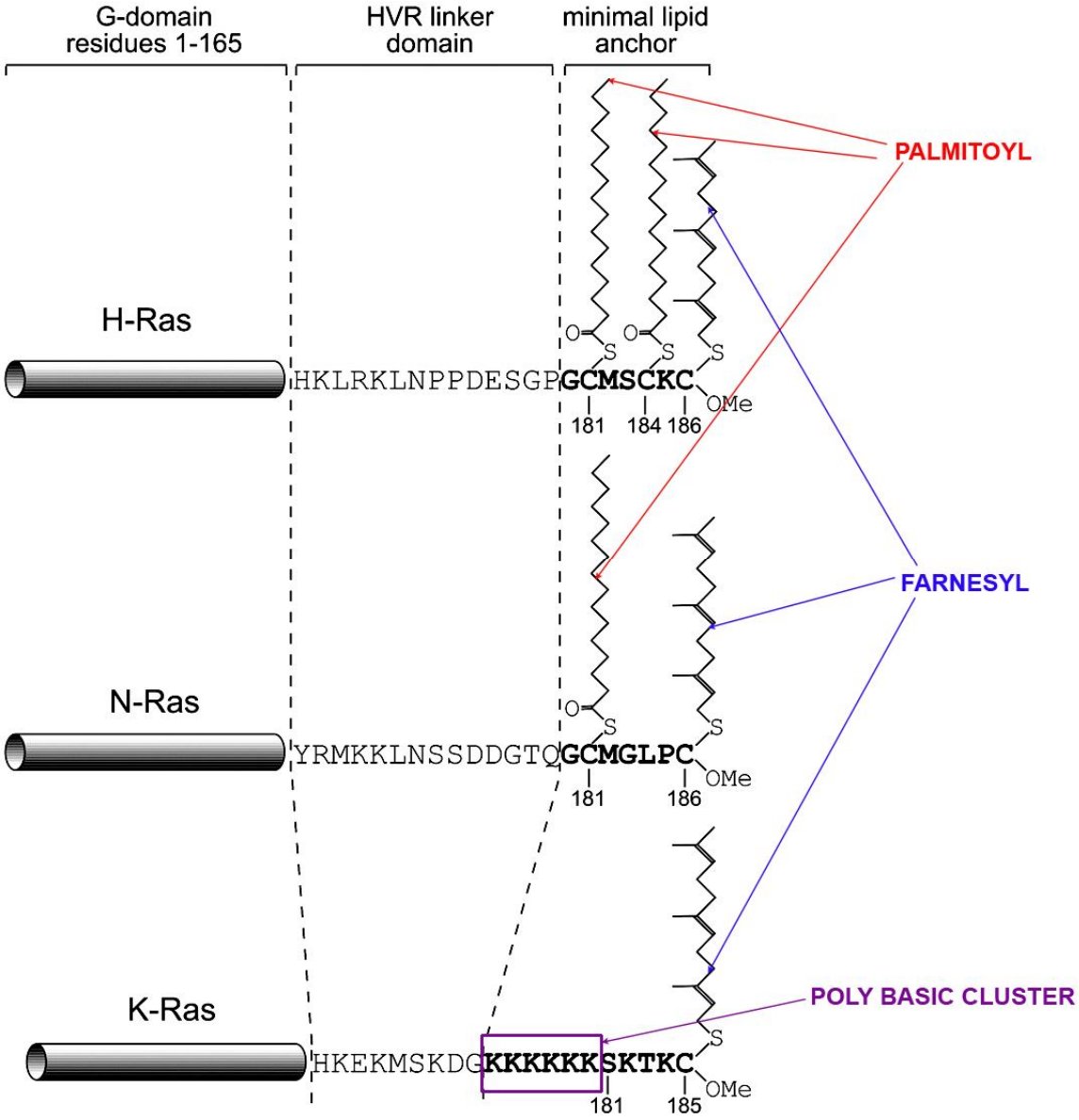


Figure 43: Domain structure of Ras isoforms
Domain structure of different Ras isoforms. The minimal lipid anchor (bold) is at the C-termini, containing a farnesyl moiety (all Ras isoforms), one (N-Ras) or two (H-Ras) palmitoyl moieties, or a six-lysine polybasic cluster (K-Ras). Complete HVR domain comprise the lipid anchor and HVR linker domain. The HVR linker is preceded by the G-domain (residues 1–165), which is highly homologous in all Ras isoforms. Adapted for simplicity from Eisenberg and Henis 2008.

Since S-farnesyl-carboxymethyl ester is not sufficient for effective membrane association, hence further modification is required. The second post-translational signals are palmitoylation signals that are located in the HVR region, and differ between Ras isoforms (see Figure 43). The palmitoylation signals are required to transit through the Golgi compartment (Apolloni *et al.,* 2000). Thus, the HVR regions of H-Ras, N-Ras and K-Ras4A contain one or two cysteines that undergo palmitoylation (C_{181} and C_{184} in H-Ras, C_{181} in N-Ras, and C_{180} in K-Ras4A), while in K-Ras4B (K-Ras) the second signal is provided by a six Lysine (K) polybasic motif (K_{175}–K_{180}) that interacts electrostatically with negatively-charged phospholipids in the internal membrane leaflet (Hancock *et al.,* 1989; Hancock *et al.,* 1990; Hancock *et al.,* 1991; Silvius 2002; Linder and Deschenes 2003; Smotrys and Linder 2004; Magee and Seabra 2005; Heo *et al.,* 2006) (see fig 1 above). The C-terminal domains up to the S-palmitoyl residue (C_{181} or C_{180}), or up to K_{175} in K-Ras, are termed **minimal lipid/membrane anchors**, and grafting of minimal lipid/membrane anchors onto other foreign proteins is sufficient to promote traffic and anchorage to the plasma membrane (Choy *et al.,* 1999; Apolloni *et al.,* 2000; Prior *et al.,* 2003). Charge and a lipid residue in the C-terminal HVR are sufficient for plasma membrane localization and function of H-Ras and K-Ras proteins (Welman *et al.,* 2000; Abankawa *et al*., 2008). Although three factor in combination contribute for plasma membrane localization (structural aspect) and functioning of the Ras protein that are the amount of net charge, charge distribution and the length of the anchoring domain. Increasing the net charge and charge concentration close to the C-terminus, increases both the percentage of membrane bound protein and also shifts the distribution from internal membranes to the plasma membrane (Welman *et al.,* 2000). Very recently for N-RAS and Ki-RAS4A it was shown that posttranslational mono-palmitoylation and farnesylation are not sufficient for specifying stable cell-surface localization of these membrane anchoring proteins. A third motif that comprises acidic residues, is present within the linker domain of all palmitoylated Ras HVRs, is necessary for stabilizing their localization to the plasma membrane. These acidic residues electrostatically stabilize palmitoylation and basic amino acids that are likely to interact (electrostatically) with acidic phospholipids enriched at the cell surface (Laude and Prior 2008). The farnesyl moiety of H-Ras and C-terminus residues contribute to H-Ras:Raf-CRD interactions, thereby increasing the affinity of H-Ras for the Raf-CRD. Moreover, farnesylation of full-length H-

Ras at C_{186} in an NMR based study does not result in any detectable conformational changes in H-Ras (Thapar *et al.,* 2004). However depalmitoylation rate in various forms of H-Ras (H-Ras_{WT}, dominant negative H-Ras_{S17N}, and oncogenic H-Ras_{Q61L} or H-$Ras_{R12/T59}$) in NIH-3T3 cells were analyzed and experimentally found that oncogenic H-Ras mutants results rapid removal of palmitoyl moiety. Based on this experimental study it was suggested that acyl-thioesterases access oncogenic H-Ras more easily because H-Ras C-terminus conformation against the membrane is altered (Baker *et al.,* 2003; Huster *et al.,* 2003). Mutating residues adjacent to the palmitoylated ysteine did not abolish H-Ras palmitoylation, but mutant proteins failed to bind to plasma membranes. Since palmitate is not attached without prior farnesyl addition, by mutational studies it was found that farnesyl is not needed as a signal for palmitate attachment or removal, and a combination of transient palmitate modification and basic residues (Lysine) can support H-Ras membrane binding and biological activity. Therefore, it was found that roles of palmitate is independent and distinct from those of farnesyl moities. More, Ras protein and plasma membrane association can be sustained largely through palmitates, farnesyl is freed to interact with other proteins (Booden *et al*., 1999). It was reported that H-Ras loses 90% of its transforming activity when mutated to serine (Ras_{C181S} and Ras_{C184S}), implying that palmitates make important contributions to oncogenicity (Baker *et al*., 2000). The distinct nature of the lipid modification determines the strength of membrane interaction of the modified protein as well as the specificity of membrane targeting (Resh 2004).

H-Ras and N-Ras are targeted to the plasma membrane via the exocytotic pathway through the Golgi apparatus (Apolloni *et al.,*, 2000; Silvius, 2002). In contrast, K-Ras4B reaches the plasma membrane by an as yet unknown, probably microtubule-dependent mechanism omitting the Golgi (Thissen *et al.,* 1997). The HVR linker region (Prior *et al.,* 2003) in H-Ras modulates membrane association and targeting. The interactions of the HVR linker domain with the plasma membrane are essential for stable membrane binding of H-Ras (Prior *et al.,* 2001; Niv *et al.,* 2002; Prior *et al.,* 2003; Eisenberg *et al.,* 2006).

Lipid anchor region contributes positive interactions with raft domains favoring H-Ras association with raft domains and due to the palmitoyl residues in the lipid anchor region (C_{181}), H-Ras prefer cholesterol-sensitive (raft-like) assemblies. H-Ras Guanine binding domain has a negative contribution to its membrane association (acts as repulsive force and

weakening the membrane association), this negative contribution depends on the activation state of H-Ras and is stronger in the active state (Ras•GTP bound conformation) (Rotblat *et al.,* 2004).

The two palmitates in H-Ras serve distinct biological roles. Monopalmitoylation of Cys181 is required and sufficient for efficient trafficking of H-ras to the plasma membrane, whereas monopalmitoylation of Cys184 does not permit efficient trafficking beyond the Golgi apparatus (Roy *et al.,* 2005). H-Ras palmitoylation at C_{181} is sufficient to stabilize the membrane interactions of H-Ras_{wt} (Roy *et al.,* 2005), but not at C_{184}. Thus C_{184} is presumably involved in the functional role of Ras proteins (Gorfe and McCammon 2008). Importantly, C_{181} is also the palmitoylation site in N-Ras. Concomitantly, electron microscopy (EM) studies of H-Ras membrane association revealed that the Ras_{C184S} mutation (monopalmitoylation at C_{181}) drives H-Ras_{G12V}•GTP into cholesterol-sensitive sites (Roy *et al.,* 2005), which in H-Ras_{wt} has a weaker affinity to raft domains than Ras•GDP. Thus, palmitoylation of C_{181} is not only important for anchoring, trafficking, , and signaling of Ras protein (Roy *et al.,* 2005) but also preferentially targets H-Ras to cholesterol-sensitive domains/clusters. Since K-Ras has poly-lysine strong polybasic cluster in the HVR region that interacts with negatively-charged lipids in the internal membrane leaflet (Cadwallader *et al.,* 1994; Ghomashchi *et al.,* 1995; McLaughlin and Murray 2005), appears at nonraft regions. Recently, it was demonstrated that two negatively-charged inositol phospholipids (phosphatidylinositol 4,5-bisphosphate and phosphatidylinositol 3,4,5-trisphosphate) target K-Ras to the plasma membrane (Heo *et al.,* 2006).

The different targeting of Ras isoforms to cholesterol dependent assemblies may affect their responsiveness to extracellular signals and there is enough evidence that H- and K-Ras may differ in their association with raft domains *e.g.* differential effects of external clustering of raft-associated influenza hemagglutinin (HA) variants on the lateral diffusion and signaling of H-Ras was observed, but not in case of K-Ras (Eisenberg *et al.,* 2006). Affinity of H- vs. K-Ras isoforms to cholesterol-sensitive assemblies (rafts) is highest for H-Ras_{wt}•GDP bound state, decreases (but still exists) for constitutively active H-Ras_{G12V}•GTP bound state, and is negligible for either K-Ras_{wt} or K-Ras_{G12V} (Prior *et al.,* 2001; Niv *et al.,* 2002; Prior *et al.,* 2003; Rotblat *et al.,* 2004; Eisenberg *et al.,* 2006).

6.2 Complete table of solvation free energy values

Average results for differential solvation free energy for all Ras protein residues described in the work. Blue and red correspond to residues with less than -2 kcal mol-1 stability or more than 2 kcal mol^{-1} unstability with respect to water

<A>	<B >	<C>	<D>	sec_str	A.A	Position
-12.64	-9.96	-14.00	-10.50	β1	K	5
-0.82	-0.69	-0.82	-1.31	β1	L	6
-1.01	-1.03	-0.86	-0.32	β1	V	7
-0.86	-0.82	-1.25	-2.17	β1	V	8
0.09	0.16	0.64	0.59	β1	V	9
-0.05	0.43	0.06	0.36	L1	G	10
-1.28	-0.71	-1.16	-2.01	L1	A	11
0.15	0.19	1.25	0.77	L1	G	12
0.40	0.29	0.60	0.73	L1	G	13
0.80	0.81	1.16	0.87	L1	V	14
0.91	0.89	1.69	1.61	α1	G	15
-1.92	-3.18	-0.19	-4.29	α1	K	16
-0.14	0.19	-0.03	0.07	α1	S	17
0.24	0.33	0.59	0.67	α1	A	18
-0.47	-0.43	-0.20	-0.15	α1	L	19
-1.37	-0.98	-1.02	-1.52	α1	T	20
-2.35	-2.03	-2.39	-2.48	α1	I	21
-1.19	-0.71	-1.09	-2.75	α1	Q	22
-0.84	-0.66	-0.90	-1.91	α1	L	23
-0.41	-0.27	-0.66	-0.37	α1	I	24
0.09	0.38	0.93	1.18	α1	Q	25
-0.71	-0.97	0.36	-0.70	α1	N	26
1.64	1.50	2.24	2.99	L2	H	27
0.28	0.41	1.24	1.24	L2	F	28
0.43	0.47	0.69	0.81	L2	V	29
-4.37	-1.44	-3.12	-5.54	L2	D	30
-1.88	-0.83	-1.96	0.31	L2	E	31
0.25	0.71	1.08	0.72	L2	Y	32
-0.13	-0.19	0.19	-4.34	L2	D	33
0.33	0.26	1.28	1.57	L2	P	34
0.37	0.26	0.36	0.53	L2	T	35
0.00	0.72	0.22	0.42	L2	I	36
3.79	1.42	-0.75	-4.27	L2	E	37
1.63	4.94	0.34	0.15	β2	D	38
0.42	0.49	-0.07	1.23	β2	S	39
-0.28	0.01	-0.26	0.08	β2	Y	40
-6.11	-4.68	-10.03	-14.07	β2	R	41
-2.96	-2.90	-7.78	-7.35	β2	K	42
-0.45	-0.34	-0.80	-0.96	β3	L	53

0.34	0.51	0.66	2.71	β3	D	54
-0.85	-0.88	-0.98	-0.41	β3	I	55
-1.45	-1.05	-1.62	-2.03	β3	L	56
1.82	0.42	1.27	0.01	β3	D	57
-0.83	-0.57	-0.92	-1.84	β3	T	58
-0.20	0.07	0.11	-0.22	L4	A	59
0.32	0.12	1.34	-0.02	L4	G	60
0.01	-0.04	0.63	0.62	L4	Q	61
0.25	0.98	-3.52	-6.12	L4	E	62
2.09	2.94	-4.72	-6.26	L4	E	63
1.02	0.87	1.40	0.57	L4	Y	64
1.16	0.63	0.76	1.63	L4	S	65
0.98	0.72	1.39	0.60	L4	A	66
0.85	0.59	1.23	0.42	α2	M	67
-2.77	-4.71	1.10	4.80	α2	R	68
1.73	1.68	-2.07	-5.97	α2	D	69
0.39	-0.76	-1.26	-0.27	α2	Q	70
0.78	-0.36	1.46	0.47	α2	Y	71
1.25	1.49	1.29	1.39	α2	M	72
0.79	0.20	0.92	1.42	α2	T	74
0.09	-0.09	-0.06	-0.57	α2	G	75
-0.65	-0.59	-0.92	-0.71	β4	G	77
-1.26	-0.87	-0.73	-1.30	β4	F	78
-0.63	-0.58	-0.62	-0.70	β4	L	79
-0.24	-0.58	-0.75	-1.98	β4	C	80
-0.27	-0.15	-0.33	-0.50	β4	V	81
-0.40	-0.17	-0.05	-0.13	β4	F	82
0.64	0.53	0.95	0.89	β4	A	83
-1.92	-1.38	-2.24	-2.49	β4	I	84
0.52	0.31	1.41	1.56	L6	N	85
-0.70	-0.53	0.16	0.82	L6	N	86
0.29	0.71	0.71	0.59	α3	T	87
-4.95	-5.35	-5.70	1.02	α3	K	88
0.51	0.60	-0.12	-0.33	α3	S	89
0.34	0.42	0.11	0.53	α3	F	90
-1.42	-1.91	-1.14	-1.74	α3	E	91
-2.03	-1.51	-3.62	-4.46	α3	D	92
-0.52	-0.24	-0.50	-0.49	α3	I	93
0.44	0.52	0.71	0.95	α3	H	94
-0.26	-0.61	-0.29	-0.42	α3	Q	95
-0.33	-0.67	-0.18	0.03	α3	Y	96
-6.51	-6.76	-7.81	-8.83	α3	R	97
-0.46	0.07	-0.73	-2.45	α3	Q	99
-1.08	-0.88	-1.07	-0.70	α3	I	100
0.78	0.83	0.56	-0.01	β5	M	111
-1.16	-0.92	-0.79	-0.89	β5	V	112
-0.80	-0.67	-1.03	-1.03	β5	L	113
-0.86	-0.83	-0.84	-1.00	β5	V	114
-0.69	-0.75	-0.31	-0.61	β5	G	115

-1.26	-1.09	-1.35	-1.00	β5	N	116
-2.42	-1.56	-0.93	0.07	β5	K	117
0.24	0.06	0.48	1.27	L8	C	118
-4.22	-4.45	-1.77	-0.67	L8	D	119
0.17	0.15	1.07	1.66	L8	L	120
0.35	0.28	0.81	0.76	L8	A	121
0.36	0.06	1.13	1.31	L8	A	122
-4.06	-2.78	-5.58	-6.35	L8	R	123
1.20	0.85	0.75	0.80	L8	T	124
-0.22	0.07	-0.25	0.56	L8	V	125
-0.64	-0.37	-1.15	-1.66	α4	S	127
0.41	0.45	0.49	0.85	α4	Q	129
-0.58	-0.23	-0.69	-1.07	α4	A	130
-0.58	-0.67	-0.85	-1.18	α4	Q	131
-0.27	-0.34	-0.46	-0.21	α4	L	133
0.71	0.71	0.59	0.60	L9	I	139
-0.55	-0.38	-0.39	-0.42	β6	P	140
0.65	0.77	0.51	0.00	β6	Y	141
-0.90	-0.83	-0.99	-0.71	β6	I	142
-4.52	-5.00	-4.56	-3.61	β6	E	143
-0.50	-0.46	-0.03	0.17	β6	T	144
-0.43	-0.11	-0.28	-1.24	L10	S	145
0.18	0.74	0.45	0.45	L10	A	146
-5.10	-3.77	-1.31	-0.67	L10	K	147
0.75	0.79	1.82	2.07	L10	T	148
-1.70	-0.09	-0.65	1.99	L10	R	149
-0.86	-0.52	-0.24	0.29	L10	Q	150
0.57	0.48	0.45	0.70	α5	G	151
0.70	0.59	0.85	0.34	α5	V	152
-1.82	-2.41	-3.63	-6.91	α5	E	153
-2.24	-1.81	-1.36	1.07	α5	D	154
0.28	0.00	0.34	0.37	α5	A	155
0.26	-0.02	0.28	0.56	α5	F	156
0.47	0.37	0.08	0.49	α5	Y	157
-1.05	-0.62	-0.63	-2.00	α5	T	158
-0.40	-0.35	-0.21	0.15	α5	L	159

Table 5: Complete table of solvation energy

7 References

Ahmadian, M. R., P. Stege, et al. (1997). "Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras." Nat Struct Biol **4**: 686-689.

Allen, M. P. and D. J. Tildesley (1987). Computer Simulation of Liquids, Oxford University Press.

Allin, C., M. R. Ahmadian, et al. (2001). "Monitoring the GAP catalyzed H-Ras GTPase reaction at atomic resolution in real time." Proc Natl Acad Sci U S A **98**: 7754-7759.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol **215**: 403-410.

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**: 3389-3402.

Apolloni, A., I. A. Prior, et al. (2000). "H-ras but not K-ras traffics to the plasma membrane through the exocytic pathway." Mol Cell Biol **20**: 2475–2487.

Aqvist, J., K. Kolmodin, et al. (1999). "Mechanistic alternatives in phosphate monoester hydrolysis: what conclusions can be drawn from available experimental data?" Chem Biol **6**: R71-R80.

Ashery, U., O. Yizhar, et al. (2006). "Nonconventional trafficking of Ras associated with Ras signal organization." Traffic **7**: 119-126.

Baker, T. L., M. A. Booden, et al. (2000). "S-Nitrosocysteine increases palmitate turnover on Ha-Ras in NIH 3T3 cells." J Biol Chem **275**: 22037-22047.

Baker, T. L., H. Zheng, et al. (2003). "Distinct rates of palmitate turnover on membrane-bound cellular and oncogenic H-ras." J Biol Chem **278**: 19292-19300.

Barbacid, M. (1987). "Ras genes." Annu Rev Biochem **56**: 779-827.

Bar-Sagi, D. and A. Hall (2000). "Ras and Rho GTPases: a family reunion." Cell **103**: 227-238.

Berman, H., K. Henrick, et al. (2003). "Announcing the worldwide Protein Data Bank." Nat Struct Biol **10**: 980.

Berman, H. M., J. Westbrook, et al. (2000). "The Protein Data Bank." Nucleic Acids Research **28**: 235-242.

Bollag, G. and F. McCormick (1991). "Differential regulation of rasGAP and neurofibromatosis gene product activities." Nature **351**: 576-579.

Bonet, J., G. Caltabiano, et al. (2006). "The role of residue stability in transient protein-protein interactions involved in enzymatic phosphate hydrolysis. A computational study." Proteins **63**(1): 65-77.

Booden, M. A., T. L. Baker, et al. (1999). "A non-farnesylated Ha-Ras protein can be palmitoylated and trigger potent differentiation and transformation." J Biol Chem **274**: 1423-1431.

Bos, J. L. (1989). "Ras oncogenes in human cancer: A review." Cancer Res **49**: 4682-4689.

Bos, J. L. (1995). "p21ras: an oncoprotein functioning in growth factor-induced signal transduction." Eur J Cancer **31A**: 1051-1054.

Bourne, H. R., D. A. Sanders, et al. (1991). "The GTPase superfamily: conserved structure and molecular mechanism." Nature **349**: 117-127.

Buhrman, G., G. Wink, et al. (2007). "Transformation efficiency of RasQ61 mutants linked to structural features of the switch regions in the presence of Raf." Structure **15** 1618–1629.

Bunney, T. D., R. Harris, et al. (2006). "Structural and mechanistic insights into Ras association domains of Phospholipase-C Epsilon." Molecular Cell **21**: 495-507.

Cadwallader, K. A., H. Paterson, et al. (1994). "N-terminally myristoylated Ras proteins require palmitoylation or a polybasic domain for plasma membrane localization." Mol Cell Biol **14**: 4722-4730.

Campbell, S. L., R. Khosravi-Far, et al. (1998). "Increasing complexity of Ras signaling." Oncogene **17**: 1395-1413.

Casey, P. J. and M. C. Seabra (1996). "Protein prenyltransferases." J Biol Chem **271**: 5289-5292.

Casey, P. J., P. A. Solski, et al. (1989). "p21ras is modified by a farnesyl isoprenoid." Proc Natl Acad Sci U S A **86**: 8323-8327.

Cepus, V., A. J. Scheidig, et al. (1998). "Time-resolved FTIR studies of the GTPase reaction of H-ras p21 reveal a key role for the beta-phosphate." Biochemistry **37**: 10263-10271.

Choy, E., V. K. Chiu, et al. (1999). "Endomembrane trafficking of ras: the CAAX motif targets proteins to the ER and Golgi." Cell **98**: 69-80.

Chu, Z. T., J. Villà-Freixa, et al. (2003). MOLARIS version alpha9.06.01.

Clamp, M., J. Cuff, et al. (2004). "The Jalview Java alignment editor." Bioinformatics **20**: 426-427.

Colicelli, J. (2004). "Human RAS superfamily proteins and related GTPases." Sci STKE **2004**: RE13.

Cox, A. D. and C. J. Der (2003). "The dark side of Ras: regulation of apoptosis." Oncogene **22**: 8999-9006.

Dessailly, B. H., M. F. Lensink, et al. (2007). "Relating destabilizing regions to known functional sites in proteins." BMC Bioinformatics **8**: 141.

Downward, J. (2003). "Targeting RAS signalling pathways in cancer therapy." Nat Rev Cancer **3**: 11-22.

Du, X., G. E. Black, et al. (2004). "Kinetic isotope effects in Ras-catalyzed GTP hydrolysis: evidence for a loose transition state." Proc Natl Acad Sci U S A **101**: 8858-8863.

Du, X., H. Frei, et al. (2000). "The mechanism of GTP hydrolysis by Ras probed by Fourier transform infrared spectroscopy." J Biol Chem **275**: 8492-8500.

Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Res **32**: 1792-1797.

Eisenberg, S. and Y. I. Henis (2008). "Interactions of Ras proteins with the plasma membrane and their roles in signaling." Cell Signal **20**: 31-39.

Eisenberg, S., D. E. Shvartsman, et al. (2006). "Clustering of raft-associated proteins in the external membrane leaflet modulates internal leaflet H-ras diffusion and signaling." Mol Cell Biol **26**: 7190-7200.

Elcock, A. H. (2001). "Prediction of functionally important residues based solely on the computed energetics of protein structure." J Mol Biol **312**: 885-896.

Esteban, L. M., C. Vicario-Abejon, et al. (2001). "Targeted genomic disruption of H-ras and N-ras, individually or in combination, reveals the dispensability of both loci for mouse growth and development." Mol Cell Biol **21**: 1444-1452.

Farnsworth, C. L. and L. A. Feig (1991). "Dominant inhibitory mutations in the Mg(2+)-binding site of RasH prevent its activation by GTP." Mol. Cell. Biol. **11**: 4822-4829.

Fasano, O., T. Aldrich, et al. (1984). "Analysis of the transforming potential of the human H-ras gene by random mutagenesis." Proc. Natl. Acad. Sci. USA **81**: 4008-4012.

Fidyk, N. J. and R. A. Cerione (2002). "Understanding the catalytic mechanism of GTPase-activating proteins: demonstration of the importance of switch domain stabilization in the stimulation of GTP hydrolysis." Biochemistry **41**: 15644-15653.

Ford, B., V. Hornak, et al. (2006). "Structure of a transient intermediate for GTP hydrolysis by ras." Structure **14**: 427-436.

Franken, S. M., A. J. Scheidig, et al. (1993). "Three-dimensional structures and properties of a transforming and a nontransforming glycine-12 mutant of p21H-ras." Biochemistry **32**: 8411-8420.

Ghomashchi, F., X. Zhang, et al. (1995). "Binding of prenylated and polybasic peptides to membranes: affinities and intervesicle exchange." Biochemistry **34**: 11910-11918.

Giehl, K. (2005). "Oncogenic Ras in tumour progression and metastasis." Biol Chem **386**: 193-205.

Glennon, T. M., J. Villà, et al. (2000). "How does GAP catalyze the GTPase reaction of Ras? A computer simulation study." Biochemistry **39**: 9641-9651.

Gorfe, A. A. and J. A. McCammon (2008). "Similar membrane affinity of mono- and Di-S-acylated ras membrane anchors: a new twist in the role of protein lipidation." J Am Chem Soc **130**: 12624-12625.

Grigorenko, B. L., A. V. Nemukhin, et al. (2005). "QM/MM modeling the Ras-GAP catalyzed hydrolysis of guanosine triphosphate." Proteins **60**: 495-503.

Gutierrez, L., A. I. Magee, et al. (1989). "Post-translational processing of p21ras is two-step and involves carboxyl-methylation and carboxy-terminal proteolysis." Embo J **8**: 1093-1098.

Hall, B. E., D. Bar-Sagi, et al. (2002). "The structural basis for the transition from Ras-GTP to Ras-GDP." Proc. Natl. Acad. Sci. USA **99**: 12138-12142.

Hall, B. E., S. S. Yang, et al. (2001). "Structure-based mutagenesis reveals distinct functions for Ras switch-I and switch-II in Sos-catalyzed guanine nucleotide exchange." J. Biol. Chem. **276**: 27629-27637.

Hamad, N. M., J. H. Elconin, et al. (2002). "Distinct requirements for Ras oncogenesis in human versus mouse cells." Genes Dev **16**: 2045-2057.

Hancock, J. F., K. Cadwallader, et al. (1991). "A CAAX or a CAAL motif and a second signal are sufficient for plasma membrane targeting of ras proteins." Embo J **10**: 4033-4039.

Hancock, J. F., A. I. Magee, et al. (1989). "All ras proteins are polyisoprenylated but only some are palmitoylated." Cell **57**: 1167-1177.

Hancock, J. F. and R. G. Parton (2005). "Ras plasma membrane signalling platforms." Biochem J **389**: 1-11.

Hancock, J. F., H. Paterson, et al. (1990). "A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane." Cell **63**: 133-139.

Heo, W. D., T. Inoue, et al. (2006). "PI(3,4,5)P3 and PI(4,5)P2 lipids target proteins with polybasic clusters to the plasma membrane." Science **314**: 1458-1461.

Herrmann, C., G. Horn, et al. (1996). "Differential interaction of the ras family GTP-binding proteins H-Ras, Rap1A, and R-Ras with the putative effector molecules Raf kinase and Ral-guanine nucleotide exchange factor." J. Biol. Chem. **271**: 6794-6800.

Huang, L., F. Hofer, et al. (1998). "Structural basis for the interaction of Ras with RalGDS." Nat. Str. Biol. **5**: 422-426.

Huster, D., A. Vogel, et al. (2003). "Membrane insertion of a lipidated ras peptide studied by FTIR, solid-state NMR, and neutron diffraction spectroscopy." J Am Chem Soc **125**: 4070-4079.

Hwang, Y. W. and D. L. Miller (1987). "A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein." J. Biol. Chem. **262**: 13081-13085.

Jackson, J. H., C. G. Cochrane, et al. (1990). "Farnesol modification of Kirsten-ras exon 4B protein is essential for transformation." Proc Natl Acad Sci U S A **87**: 3042-3046.

John, J., M. Frech, et al. (1988). "Biochemical properties of Ha-ras encoded p21 mutants and mechanism of the autophosphorylation reaction." J. Biol. Chem. **263**: 11792-11799.

Johnson, L., D. Greenbaum, et al. (1997). "K-ras is an essential gene in the mouse with partial functional overlap with N-ras." Genes Dev **11**: 2468-2481.

Johnston, M. A., I. F. Galván, et al. (2005). "Framework-based design of a new all-purpose molecular simulation application: the Adun simulator." J Comput Chem **26**: 1647-1659.

Kabsch, W. and C. Sander (1983). "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features." Biopolymers **22**: 2577-2637.

Karlovich, C. A., L. Bonfini, et al. (1995). "In vivo functional analysis of the Ras exchange factor son of sevenless." Science **268**: 576-579.

Koera, K., K. Nakamura, et al. (1997). "K-ras is essential for the development of the mouse embryo." Oncogene **15**: 1151-1159.

Kosloff, M. and Z. Selinger (2001). "Substrate assisted catalysis -- application to G proteins." Trends Biochem Sci **26**: 161-166.

Kraulis, P. J., P. J. Domaille, et al. (1994). "Solution structure and dynamics of Ras p2 bGDP determined by heteronuclear Three- and Four-Dimensional NMR spectroscopy." Biochemistry **33**: 3515-3531.

Krengel, U., I. Schlichting, et al. (1990). "Three-dimensional structures of H-ras p21 mutants: molecular basis for their inability to function as signal switch molecules." Cell **62**: 539-548.

Laude, A. J. and I. A. Prior (2008). "Palmitoylation and localisation of RAS isoforms are modulated by the hypervariable linker domain." J Cell Sci **121**: 421-7.

Lee, F. S., Z. T. Chu, et al. (1993). "Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs." J Comput Chem **14**: 161-185.

Lee, F. S. and A. Warshel (1992). "A Local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations." J. Chem Phys. **97**: 3100-3107.

Levitt, M. (2001). "The birth of computational structural biology." Nat Struct Biol **8**: 392-393.

Lim, K. H., A. T. Baines, et al. (2005). "Activation of RalA is critical for Ras-induced tumorigenesis of human cells." Cancer Cell **7**: 533-545.

Lowy, D. R. and B. M. Willumsen (1993). "Function and regulation of ras." Annu Rev Biochem **62**: 851-891.

Ma, J. and M. Karplus (1997). "Molecular switch in signal transduction: Reaction paths of the conformational changes in ras p21." Proc. Natl. Acad. Sci. USA **94** 11905–11910.

Maegley, K. A., S. J. Admiraal, et al. (1996). "Ras-catalyzed hydrolysis of GTP: a new perspective from model studies." Proc. Natl. Acad. Sci. USA **93**: 8160–8166.

Magee, T. and M. C. Seabra (2005). "Fatty acylation and prenylation of proteins: what's hot in fat." Curr Opin Cell Biol **17**: 190-196.

Margarit, S. M., H. Sondermann, et al. (2003). "Structural evidence for feedback activation by Ras.GTP of the Ras-specific nucleotide exchange factor SOS." Cell **112**: 685-95.

Marshall, M. S. and L. A. Hettich (1993). "Characterization of Ras effector mutant interactions with the NF1-GAP related domain." Oncogene **8**: 425.

Matallanas, D., V. Sanz-Moreno, et al. (2006). "Distinct utilization of effectors and biological outcomes resulting from site-specific Ras activation: Ras functions in lipid rafts and Golgi complex are dispensable for proliferation and transformation." Mol Cell Biol **26**: 100-116.

McLaughlin, S. and D. Murray (2005). "Plasma membrane phosphoinositide organization by protein electrostatics." Nature **438**: 605-611.

Milburn, M. V., L. Tong, et al. (1990). "Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins." Science **247**: 939-45.

Mitin, N., K. L. Rossman, et al. (2005). "Signaling interplay in Ras superfamily function." Curr Biol **15**: R563-R574.

Muegge, I., T. Schweins, et al. (1996). "Electrostatic control of GTP and GDP binding in the oncoprotein p21ras." Structure **4**: 475-489.

Muegge, I., T. Schweins, et al. (1998). "Electrostatic contributions to protein-protein binding affinities: application to Rap/Raf interaction." Proteins **30**: 407-423.

Nassar, N., G. Horn, et al. (1996). "Ras/Rap effector specificity determined by charge reversal." Nat. Struct. Biol. **3**: 723-729.

Nassar, N., G. Horn, et al. (1995). "The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue." Nature **375**: 554-560.

Neal, S. E., J. F. Eccleston, et al. (1990). "Hydrolysis of GTP by p21NRAS, the NRAS protooncogene product, is accompanied by a conformational change in the wild-type protein: use of a single fluorescent probe at the catalytic site." Proc. Natl. Acad. Sci. USA **87**: 3562-3565.

Niv, H., O. Gutman, et al. (2002). "Activated K-Ras and H-Ras display different interactions with saturable nonraft sites at the surface of live cells." J Cell Biol **157**: 865-872.

Pacold, M. E., S. Suire, et al. (2000). "Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma." Cell **103**: 931-943.

Paduch, M., F. Jelen, et al. (2001). "Structure of small G proteins and their regulators." Acta Biochem Pol **48**: 829–850.

Pai, E. F., U. Krengel, et al. (1990). "Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 A resolution: implications for the mechanism of GTP hydrolysis." EMBO J. **9**(8): 2351-2359.

Pasqualato, S. and J. Cherfils (2005). "Crystallographic evidence for substrate-assisted GTP hydrolysis by a small GTP binding protein." Structure **13**: 533-540.

Peeper, D. S., T. M. Upton, et al. (1997). "Ras signalling linked to the cell-cycle machinery by the retinoblastoma protein." Nature **386**: 177-1781.

Polakis, P. and F. McCormick (1993). "Structural requirements for the interaction of p21ras with GAP, exchange factors, and its biological effector target." J. Biol. Chem. **268**: 9157–9160.

Prior, I. A., A. Harding, et al. (2001). "GTP-dependent segregation of H-ras from lipid rafts is required for biological activity." Nat Cell Biol **3**: 368-375.

Prior, I. A., C. Muncke, et al. (2003). "Direct visualization of Ras proteins in spatially distinct cell surface microdomains." J Cell Biol **160**: 165-170.

Prive, G. G., M. V. Milburn, et al. (1992). "X-ray crystal structures of transforming p21 ras mutants suggest a transition-state stabilization mechanism for GTP hydrolysis." Proc Natl Acad Sci U S A **89**: 3649-53.

Quatela, S. E. and M. R. Philips (2006). "Ras signaling on the Golgi." Curr Opin Cell Biol **18**: 162-167.

Rensland, H., J. John, et al. (1995). "Substrate and product structural requirements for binding of nucleotides to H-ras p21: the mechanism of discrimination between guanosine and adenosine nucleotides." Biochemistry **34**: 593-599.

Resh, M. D. (2004). "Membrane targeting of lipid modified signal transduction proteins." Subcell Biochem **37**: 217-232.

Rittinger, K., P. A. Walker, et al. (1997). "Structure at 1.65 A of RhoA and its GTPase-activating protein in complex with a transition-state analogue." Nature **389**: 758-762.

Roca, M., B. Messer, et al. (2008). "On the relationship between folding and chemical landscapes in enzyme catalysis." Proc Natl Acad Sci U S A **105**: 13877-13882.

Rodenhuis, S. (1992). "Ras and human tumors." Semin Cancer Biol **3**: 241-247.

Rodriguez-Viciana, P., C. Sabatier, et al. (2004). "Signaling specificity by Ras family GTPases is determined by the full spectrum of effectors they regulate." Mol Cell Biol **24**: 4943-4954.

Rotblat, B., I. A. Prior, et al. (2004). "Three separable domains regulate GTP-dependent association of H-ras with the plasma membrane." Mol Cell Biol **24**: 6799-6810.

Rouhi, M. (2000). "Most proficient enzyme works traditionally." Chem Eng News **78**: 49-50.

Roy, S., R. Luetterforst, et al. (1999). "Dominant-negative caveolin inhibits H-Ras function by disrupting cholesterol-rich plasma membrane domains." Nat Cell Biol **1**: 98-105.

Roy, S., S. Plowman, et al. (2005). "Individual palmitoyl residues serve distinct roles in H-ras trafficking, microlocalization, and signaling." Mol Cell Biol **25**: 6722-6733.

Scheffzek, K. and M. R. Ahmadian (2005). "GTPase activating proteins: structural and functional insights 18 years after discovery." Cell Mol Life Sci **62**: 3014-3038.

Scheffzek, K., M. R. Ahmadian, et al. (1997). "The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants." Science **277**: 333-338.

Scheffzek, K., P. Grunewald, et al. (2001). "The Ras-Byr2RBD complex: structural basis for Ras effector recognition in yeast." Structure **9**: 1043-50.

Scheper, J., B. Oliva, et al. (2009). "Analysis of electrostatic contributions to the selectivity of interactions between RING-finger domains and ubiquitin-conjugating enzymes." Proteins **74**: 92-103.

Schlichting, I., S. C. Almo, et al. (1990). "Time-resolved X-ray crystallographic study of the conformational change in Ha-Ras p21 protein on GTP hydrolysis." Nature **345**: 309-315.

Schutz, C. N. and A. Warshel (2001). "What are the dielectric "constants" of proteins and how to validate electrostatic models?" PROTEINS: Structure, Function, and Genetics **44**: 400-417.

Schweins, T., M. Geyer, et al. (1995). "Substrate-assisted catalysis as a mechanism for GTP hydrolysis of p21ras and other GTP-binding proteins." Nat Struct Biol **2**: 36-44.

Schweins, T., R. Langen, et al. (1994). "Why have mutagenesis studies not located the general base in ras p21." Nat Struct Biol **1**: 476-484.

Schweins, T., K. Scheffzek, et al. (1997). "The Role of the Metal Ion in the p21ras Catalysed GTP-hydrolysis: Mn2+ versus Mg2+." J. Mol. Biol. **266**: 847-856.

Sham, Y. Y., Z. T. Chu, et al. (2000). "Examining methods for calculations of binding free energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA calculations of ligands binding to an HIV protease." Proteins **39**: 393-407.

Shields, J. M., K. Pruitt, et al. (2000). "Understanding Ras: 'it ain't over 'til it's over'." Trends Cell Biol **10**: 147-54.

Shurki, A. and A. Warshel (2004). "Why does the Ras switch "break" by oncogenic mutations?" Proteins **55**: 1-10.

Sigal, I. S., J. B. Gibbs, et al. (1986). "Mutant ras-encoded proteins with altered nucleotide binding exert dominant biological effects." Proc. Natl. Acad. Sci. USA **83**: 952-956.

Silvius, J. R. (2002). "Mechanisms of Ras protein targeting in mammalian cells." J Membr Biol **190**: 83-92.

Smotrys, J. E. and M. E. Linder (2004). "Palmitoylation of intracellular signaling proteins: regulation and function." Annu. Rev. Biochem. **73**: 559-587.

Sondermann, H., S. M. Soisson, et al. (2004). "Structural analysis of autoinhibition in the Ras activator Son of sevenless." Cell **119**: 393-405.

Spoerner, M., C. Herrmann, et al. (2001). "Dynamic properties of the Ras switch I region and its importance for binding to effectors." Proc. Natl. Acad. Sci. USA **98**: 4944-4949.

Sprang, S. R. (1997). "G protein mechanisms: insights from structural analysis." Annu Rev Biochem **66**: 639-678.

Thapar, R., J. G. Williams, et al. (2004). "NMR characterization of full-length farnesylated and non-farnesylated H-Ras and its implications for Raf activation." J Mol Biol **343**: 1391-1408.

Thissen, J. A., J. M. Gross, et al. (1997). "Prenylation-dependent association of Ki-Ras with microtubules. Evidence for a role in subcellular trafficking." J Biol Chem **272**: 30362-30370.

Topol, I. A., R. E. Cachau, et al. (2004). "Quantum chemical modeling of the GTP hydrolysis by the RAS-GAP protein complex." Biochim Biophys Acta **1700**(1): 125-136.

Valleau, J. P. and G. M. Torrie (1977). A Guide to Monte Carlo for Statistical Mechanics. New York, Plenum.

Villà, J. and A. Warshel (2001). "Energetics and dynamics of enzymatic reactions." J Phys Chem B **105**: 7887-7907.

Voice, J. K., R. L. Klemke, et al. (1999). "Four human ras homologs differ in their abilities to activate Raf-1, induce transformation, and stimulate cell motility." J Biol Chem **274**: 17164-17170.

Walsh, A. B. and D. Bar-Sagi (2001). "Differential activation of the Rac pathway by Ha-Ras and K-Ras." J Biol Chem **276**: 15609-15615.

Warshel, A. (1978). "Energetics of enzyme catalysis." Proc Natl Acad Sci U S A **11**: 5250-5254.

Warshel, A. (1991). Computer modeling of chemical reactions in enzymes and solutions. New York, John Wiley & Sons.

Warshel, A. and J. Florian (1998). "Computer simulations of enzyme catalysis: finding out what has been optimized by evolution." Proc Natl Acad Sci U S A **95**: 5950-5955.

Warshel, A., J. Florián, et al. (2001). "Circe effect versus enzyme preorganization: what can be learned from the structure of the most proficient enzyme?" Chembiochem **2**: 109-111.

Warshel, A. and G. King (1985). "Polarization constraints in molecular dynamics simulation of aqueous solutions: the surface constraint all atom solvent (SCAAS) model." Chem. Phys. Lett. **121**: 124-129.

Warshel, A., M. Strajbl, et al. (2000). "Remarkable rate enhancement of orotidine 5'-monophosphate decarboxylase is due to transition-state stabilization rather than to ground-state destabilization." Biochemistry **39**: 14728-14738.

Welman, A., M. M. Burger, et al. (2000). "Structure and function of the C-terminal hypervariable region of K-Ras4B in plasma membrane targetting and transformation." Oncogene **19**: 4582-4591.

Wennerberg, K., K. L. Rossman, et al. (2005). "The Ras superfamily at a glance." J Cell Sci **118**: 843-846.

White, M. A., C. Nicolette, et al. (1995). "Multiple ras functions can contribute to mammalian-cell transformation." Cell **80**: 533–541.

Wittinghofer, A. (2006). "Phosphoryl transfer in Ras proteins, conclusive or elusive?" Trends Biochem Sci **31**: 20-23.

Wittinghofer, A. and N. Nassar (1996). "How Ras-related proteins talk to their effectors." Trends Biochem Sci. **21**: 488-491.

Wittinghofer, A., K. Scheffzek, et al. (1997). "The interaction of Ras with GTPase-activating proteins." FEBS Lett **410**: 63–67.

Yan, J., S. Roy, et al. (1998). "Ras isoforms vary in their ability to activate Raf-1 and phosphoinositide 3-kinase." J Biol Chem **273**: 24052-24056.

Zhang, F. L. and P. J. Casey (1996). "Protein prenylation: molecular mechanisms and functional consequences." Annu Rev Biochem **65**: 241-269.