STRUCTURAL ANALYSIS OF ONCOGENIC H-RAS MUTANTS G12A AND G13A

by

Ning Wu

A thesis submitted in conformity with the requirements for the degree of Master of Science

Graduate Department of Biochemistry

University of Toronto
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-45880-6
Structural Analysis of Oncogenic H-Ras Mutants G12A and G13A

Ning Wu

In conformity with the requirements for the degree of Master of Science, 1999
Department of Biochemistry, University of Toronto

ABSTRACT

About 30% of human cancers are due to point mutations in one of the H-, K-, and N-Ras genes. These small GTPases act as early switch molecules turning on and off cell proliferation processes in response to exocellular stimuli. Common mutations at Gly12, Gly13 and Gln61 prevent Ras from hydrolyzing GTP into GDP in the presence of a GAP (GTPase activating protein), prolonging the lifetime of the active state, therefore rendering the protein oncogenic. From previous mutational and crystallographic studies, we know that too big a side chain at position 12 would sterically hinder the correct positioning of the Gln61 side chain for catalysis. To answer the questions why G12A is oncogenic while G12P is not and why a small change from a hydrogen to a methyl group at Gly13 also makes the protein transforming, we studied crystal structures of G12A and G13A H-Ras. Free G12A/GDP was crystallized and subsequent modeling of its structure into the complex of H-Ras and p120GAPGRD showed that Cβ of Ala12 occupied the same position as that of Pro12. The G13A/GDP structure revealed no sign of involvement of the 13th residue with any other part of the molecule or with GAP. Therefore, it appears that a full understanding of the influences of amino acid changes at these positions will require complex structures of H-Ras mutants and the catalytic domain of GAP.
ACKNOWLEDGEMENTS

First, I would like to thank my supervisor Dr. Emil F. Pai, for all the freedom I had in trying out various ideas, and continuous support and encouragement when things repeatedly turned the wrong way.

Not only my committee members themselves Drs. D. Rose and R. Reithmeier, but their labs as well provided me with much appreciated advice and assistance.

I am very grateful to have had Dr. Young-Hwa Song constantly by my side when I took my first steps in the world of protein biochemistry, and to have Wanda Gillon taking care of all things in the lab that I took for granted.

In the two and half years of my study, many people, past and present members of the Pai lab have helped me on a great number of occasions. Particularly, I am thankful for the extensive computer knowledge of Dr. Rosi Hynes and Piotr Sliz, for Matthew Kimber's deep understanding of crystallography, and for Robert Garces’ help in creating the beautiful slides and figures.

Of course, I just have to mention my parents, the truly dearest people in the whole world.
# TABLE OF CONTENTS

## Chapter 1: Introduction

1. **Biological Studies on Ras Proteins**

   1.1. **Ras superfamily of small GTPases**
   
   1.1.1a **Mammalian H-, K- and N-Ras**
   
   1.1.2 **Ras Signaling Pathway**
   
   1.1.2a **Overview**
   
   1.1.2b **Ras activators/exchange factors**
   
   1.1.2c **Ras effectors**
   
   1.1.2d **Negative Ras regulators**
   
   1.2 **Structural Studies of Ras and Ras-interacting Proteins**
   
   1.2.1 **X-ray crystallography as a tool in understanding biological interactions**
   
   1.2.2 **Structure of H-Ras from x-ray crystallography**
   
   1.2.3 **Interactions of Ras with its regulators and effectors**
   
   1.2.3a **Complex structure of H-Ras and the hSosl CDC25H**
domain

1.2.3b Rap1A/c-Raf1 complex 26
1.2.3c H-Ras/p120^GAP^ GRD complex 28

1.3 Mutation studies of H-, K-, and N-Ras 32
1.3.1 Glycine 12 34
1.3.2 Glycine 13 35
1.3.3 Aim of this thesis 35

Chapter 2: Methods and Results 39

2.1 Mutation of the H-ras Expression Vector 39
2.1.1 G12A 39
2.1.2 G13A 40

2.2 Expression and Purification of G12A and G13A 41

2.3 Nucleotide Exchange 45

2.4 Crystallization 45
2.4.1 G12A 45
2.4.2 G13A 47

2.5 Data Collection, Processing, Molecular Replacement and Refinement 47

Chapter 3: Analysis and Future Direction 55

3.1 G12A 55
3.2 G13A 56

3.3 Conclusion 57
## Appendix A

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Crystallization of the tubulin-binding domain and catalytic domain of neurofibromin</td>
<td>64</td>
</tr>
<tr>
<td>I.1</td>
<td>Background</td>
<td>64</td>
</tr>
<tr>
<td>I.2</td>
<td>NF1-GRD purification and crystallization</td>
<td>64</td>
</tr>
<tr>
<td>I.3</td>
<td>Molecular replacement</td>
<td>66</td>
</tr>
<tr>
<td>I.4</td>
<td>Future work</td>
<td>66</td>
</tr>
<tr>
<td>II</td>
<td>HtpG</td>
<td>70</td>
</tr>
<tr>
<td>II.1</td>
<td>Experimental attempts</td>
<td>70</td>
</tr>
<tr>
<td>II.2</td>
<td>Construct</td>
<td>71</td>
</tr>
<tr>
<td>II.3</td>
<td>Results</td>
<td>71</td>
</tr>
<tr>
<td>III</td>
<td>X-Ras</td>
<td>77</td>
</tr>
<tr>
<td>III.1</td>
<td>Background</td>
<td>77</td>
</tr>
<tr>
<td>III.2</td>
<td>Protein purification and crystallization</td>
<td>77</td>
</tr>
<tr>
<td>III.3</td>
<td>Results</td>
<td>78</td>
</tr>
</tbody>
</table>

## Appendix B

Correct sequence of ptacRas plasmid

## Bibliography
LIST OF FIGURES AND TABLES

Chapter 1

Fig. 1.1 Ras activation cycle 4
Fig. 1.2 Ras effector pathways 16
Fig. 1.3 1.35Å crystal structure of H-Ras (1-166) bound to GppNp 19
Fig. 1.4 Schematic drawing of H-Ras and GppNp interactions 21
Fig. 1.5 Coordination of Mg^{2+} in the H-Ras/nucleotide complexes 22
Fig. 1.6 Active site of H-Ras 24
Fig. 1.7 Backbone comparison of H-Ras/GDP and H-Ras/GTP structures 25
Fig. 1.8 H-Ras and hSos1 catalytic domain complex structure 27
Fig. 1.9 Complex structure of p120^{GAP}-GRD with H-Ras 29
Fig. 1.10 Catalytic residues in H-Ras/p120^{GAP}-GRD complex 30
Fig. 1.11 Backbone comparison of the ground state and the transition state complexes of Rho and rhoGAP 33
Table 1.1 Transforming Ras mutants 37
Table 1.2 Transformation properties of Ras Gly12 mutants 38

Chapter 2

Fig. 2.1 Purification of Ras mutant proteins 43
Fig. 2.2 G12A/GDP crystals 47
Fig. 2.3 G12A/GppNp crystals 47
Fig. 2.4 G13A/GDP crystals 48
Fig. 2.5 G13A/GppNp crystals 48
Table 2.1 Molecular replacement solutions from AMoRe 51
Table 2.2 Refinement statistics 52
Fig. 2.6 Ramachandran plot of G12A/GDP structure 53
Fig. 2.7 Ramachandran plot of G13A/GDP structure 54

Chapter 3

Fig. 3.1 Electron density map of G12A/GDP 58
Fig. 3.2 Electron density map of G13A/GDP 59
Fig. 3.3 Backbone superposition of mutants/GDP, WT H-Ras/GppNp and H-Ras/GDP 60
Fig. 3.4 Modeled active site of G12A, G12P with p120GAP-GRD 61
Fig. 3.5 Interaction of G13A with p120GAP-GRD 62

Appendix A

Fig. I.1 Purification of NF1-GRD 67
Fig. I.2 NF1-GRD crystals 69
Fig. II.1 NR30 induction assay 73
Fig. II.2 Purification of NR30 protein 74
Fig. II.3 HtpG crystals 76
Fig. III.1 Protein sequence alignment of H-Ras and X-Ras using the program GCG 80
Fig. III.2 “X-Ras” induction assay 81
Fig. III.3 Purification of “X-Ras”/CAT1 82
Fig. III.4 CAT1 crystals 84
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akt</td>
<td>protein kinase B = PKB</td>
</tr>
<tr>
<td>bFGF</td>
<td>basic fibroblast growth factor</td>
</tr>
<tr>
<td>β-OG</td>
<td>octyl-β-D-glucopyranoside</td>
</tr>
<tr>
<td>cAMP</td>
<td>cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CDC25H domain</td>
<td>CDC25 homologous domain</td>
</tr>
<tr>
<td>C_v</td>
<td>column volume</td>
</tr>
<tr>
<td>DH domain</td>
<td>Dbl homology domain</td>
</tr>
<tr>
<td>DTT</td>
<td>D,L-dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetraacetate</td>
</tr>
<tr>
<td>EGF</td>
<td>epidermal growth factor</td>
</tr>
<tr>
<td>ERK</td>
<td>extracellular signal regulated kinase</td>
</tr>
<tr>
<td>GAP</td>
<td>GTPase-activating protein</td>
</tr>
<tr>
<td>GDP</td>
<td>guanosine 5’-diphosphate</td>
</tr>
<tr>
<td>GEF</td>
<td>guanine nucleotide exchange factor</td>
</tr>
<tr>
<td>GppNp</td>
<td>guanosine-5’-(β,γ-imido)triphosphate</td>
</tr>
<tr>
<td>GRD</td>
<td>GAP related domain</td>
</tr>
<tr>
<td>GSK3</td>
<td>glycogen synthase kinase 3</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione-S-transferase</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine 5’-triphosphate</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-2-hydroxyethylpiperazine-N’-2-ethanesulfonic acid</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl thiogalactoside</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>JNK</td>
<td>c-jun N-terminus kinase</td>
</tr>
<tr>
<td>JNKK</td>
<td>JNK kinase</td>
</tr>
<tr>
<td>$K_m$</td>
<td>Michaelis constant</td>
</tr>
<tr>
<td>MAPK</td>
<td>MAP kinase = mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MEK</td>
<td>MAPK/ERK kinase</td>
</tr>
<tr>
<td>Mnk</td>
<td>MAP kinase interacting kinase</td>
</tr>
<tr>
<td>MPD</td>
<td>2-methyl-2,4-pentanediol</td>
</tr>
<tr>
<td>NF1</td>
<td>neurofibromin</td>
</tr>
<tr>
<td>NGF</td>
<td>nerve growth factor</td>
</tr>
<tr>
<td>O.D.(_{600})</td>
<td>optical density at 600nm</td>
</tr>
<tr>
<td>PA</td>
<td>phosphatidic acid</td>
</tr>
<tr>
<td>PC</td>
<td>phosphatidylcholine</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PDGF</td>
<td>platelet-derived growth factor</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PH domain</td>
<td>pleckstrin homology domain</td>
</tr>
<tr>
<td>PI3-K</td>
<td>phosphoinositide 3-kinase</td>
</tr>
<tr>
<td>PIP(_2)</td>
<td>phosphatidylinositol-3,4-bisphosphate</td>
</tr>
<tr>
<td>PIP(_3)</td>
<td>phosphatidylinositol-3,4,5-triphosphate</td>
</tr>
<tr>
<td>PKB</td>
<td>protein kinase B = Akt</td>
</tr>
<tr>
<td>PKC</td>
<td>protein kinase C</td>
</tr>
<tr>
<td>PLD</td>
<td>phospholipase D</td>
</tr>
<tr>
<td>PP2A</td>
<td>protein phosphatase 2A</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Rsk</td>
<td>ribosomal S6 kinase</td>
</tr>
<tr>
<td>SAPK</td>
<td>stress activated protein kinase</td>
</tr>
<tr>
<td>SEK</td>
<td>SAPK/ERK kinase</td>
</tr>
<tr>
<td>SH2 domain</td>
<td>Src homology 2 domain</td>
</tr>
<tr>
<td>SH3 domain</td>
<td>Src homology 3 domain</td>
</tr>
<tr>
<td>TGF-α</td>
<td>transforming growth factor-α</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)aminomethane</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

1.1 Biochemical studies on Ras proteins

1.1.1 Ras superfamily of small GTPases

Members of the 21kDa Ras family of small GTPases act as switch molecules in relaying signals from the cell surface receptors to transcription factors and regulatory proteins. In a variety of cell types, activation of Ras is essential for growth factor-induced cell proliferation. Such growth factors include epidermal growth factor (EGF), transforming growth factor-α (TGF-α), platelet-derived growth factor (PDGF) and basic fibroblast growth factor (bFGF) (Denhardt, 1996).

So far, the family is categorized into 4 distinct branches according to sequence: Ras, Rho, Rab and Ran/TC4. Compared to the Ras branch, there is a 12-amino acid insertion in all proteins of the Rho branch between positions 122 and 123. The crystal structure of human Rac1, a member of the Rho branch, has been solved. The overall structure is very similar to that of H-ras except for an extra short helix formed by the inserted residues (Hirshberg et al., 1997). Several minor insertions and deletions are found in the Rab and Ran branches. The Ras and Rho branches are the most homogeneous with proteins always sharing more than 50% amino acid sequence identity, while the Rab branch includes the largest number of proteins as well as the highest degree of heterogeneity (Chardin, 1993). There is about 30% identity between the
members of different branches but the nucleotide-binding motif GXXXXGKS/T is completely conserved.

While activation of Ras proteins leads to general cell proliferation, Rho proteins are involved in regulation of actin cytoskeleton reorganization, and filopodia and lamellipodia formation. Rab proteins play a role in vesicle transport inside the cell and in secretion while Ran proteins are involved in the transport of RNA and proteins across the nuclear membrane (Hall, 1990). Sequence and functional differences imply that a separate set of activators and effectors are used for each branch and this has been confirmed experimentally.

1.1.1a Mammalian H-, K- and N-ras

In this study, we focus on the human H(arvey)-, K(irsten)- and N(euroblastoma)-Ras, where both H- and K-Ras genes were originally identified from murine sarcoma viruses. The isoforms have distinct expression patterns in different tissues, and there is recent evidence that they show preferences in effector binding (refer to section 1.3). About 30% of human cancers involve mutations of these genes, though the tissue distributions are nonrandom (Bos, 1988).

Of the 188 amino acids (or 189 in the case of the alternatively spliced K-ras-2 gene), residues 5-164 (catalytic domain) are highly conserved, while the C-terminal tail, the hypervariable region, exhibits a low degree of sequence conservation. The C-terminal consensus sequence in Ras CAAX (C, Cys; A, aliphatic amino acid; X, any amino acid) is a membrane targeting signal. During the first set of post-translational modifications, the cysteine is farnesylated as a thioester, and AAX is digested away before the terminal
cysteine carboxyl is methylated. Proper processing of the first set allows the palmitoylation on the cysteine residues contained within the hypervariable domain of H-Ras, N-Ras and K-RasA, a splice variant of K-Ras. The other variant K-RasB, instead of a modified cysteine residue, has a stretch of 6 lysines that allows membrane association. Both signals are required for the correct plasma membrane targeting.

\[
\begin{align*}
\text{H-Ras} & \quad \text{PDESGPGCMSCKCVLS} \\
\text{N-Ras} & \quad \text{SDDGTQGCMLPCVVM} \\
\text{K-RasA} & \quad \text{DDKTPGCVKKKCVIM} \\
\text{K-RasB} & \quad \text{EGKKKKKSKTKCVIM}
\end{align*}
\]

* The CAAX motifs are in bold. Palmitoylated, potentially palmitoylated cysteines, and polybasic domains are underlined (Marshall et al., 1993).

1.1.2 Ras Signaling Pathway

1.1.2a Overview

Signal transduction is a vast field that is far from being completely understood. The well-known Ras pathway is but a tiny corner of the web. Hence a very generic picture will be given here for the purpose of this study (Fig. 1.5).

The cell surface tyrosine kinase receptors, or receptors that are directly linked to tyrosine kinases (e.g. T-cell receptor), can activate Ras. Upon growth factor binding, the tyrosine kinase receptor dimerizes as a heterodimer in the case of PDGFR or as a homodimer in the case of EGFR, and phosphorylates itself in the cytoplasmic portion, recruiting other kinase substrates (e.g. Src) and adapter proteins (e.g. Grb2) through binding of their SH2 domains to the phosphorylated tyrosine on the receptor. Grb2 is
constitutively bound to Sos and the Grb2-Sos complex is localized to the membrane either through direct receptor interaction (Rozakis-Adcock et al., 1993; Li et al., 1993; Gale et al., 1993) or indirectly by forming a larger complex with another adapter protein such as Shc (Egan et al., 1993a). Sos (son of sevenless), a soluble cytosolic guanine nucleotide exchange factor (GEF) that catalyzes GDP to GTP exchange of Ras, is translocated to the membrane where Ras resides. Sos forms a complex with Ras/GDP and forces the release of the diphosphate nucleotide. The dimeric complex of Sos and the nucleotide-free form of Ras immediately recruits GTP from the surrounding solution ([GTP]/[GDP] ~ 10). The process is then reversed, resulting in the dissociation of Sos and the active form of Ras, its GTP complex. This form then interacts with its effectors, leading eventually to gene transcription. The signal is turned off when GTP is hydrolyzed to GDP with the help of GTPase activating proteins (GAPs), which only recognize the triphosphate nucleotide bound form of Ras (Trahey et al., 1987). Ras cycles between these two states acting as a switch (Fig. 1.1). Microinjection of Ras antibodies at different times of the cell cycle reveals that Ras is essential during the first 8 hours of G1 of the cell cycle, after which the cell is committed to complete G1 and enter S phase (Yu et al., 1988).

\[
\begin{align*}
\text{GEF} & \\
\text{Ras/GDP} & \iff \text{Ras/GTP} \rightarrow \text{Effectors} \\
\text{GAP} & \\
\end{align*}
\]

Fig. 1.1 Ras cycle
1.1.2b Ras activators/exchange factors

Following the identification of the son-of-sevenless gene (Sos), a *Drosophila* homologue of yeast CDC25, two murine (mSos1 and 2) and one human (hSos) have been isolated (Bonfini *et al*., 1992; Bowtell *et al*., 1992; Chardin *et al*., 1993), and they all have Ras nucleotide exchange activity. Other Ras-GEFs (guanine nucleotide exchange factors), RasGRP and RasGRF, have been discovered recently. Both with expression limited to neural tissue, and their activation controlled by local Ca\(^{2+}\) levels (Farnsworth *et al*., 1995; Ebinu *et al*., 1998).

i. Sos

Mammalian Sos has three conserved regions, the N-terminal DH-PH domains followed by the CDC25H (CDC25 homologue) catalytic domain and the C-terminal proline-rich sequence that binds the SH3 domains of Grb2, a human homologue of *Drosophila* Sem-5 adapter protein. Grb2-Sos complexes can be detected in non-activated cells, and both SH3 domains of Grb2 are required for high affinity binding to Sos. In mouse, the intrinsic CDC25H activity is autoinhibited by the DH and Grb2-binding domains (Kim *et al*., 1998), a regulatory measure.

In Src-transformed cells, the Shc adapter protein, a substrate of Src, replaces the EGF receptor as the phosphotyrosine target of the Grb2-Sos complex (Egan *et al*., 1993). Shc contains a protein tyrosine binding (PTB) domain, a Src SH2 domain, and three tyrosine autophosphorylation acceptor sites (Tarnawski *et al*., 1998). The Grb2 SH2 domain has the highest affinity (nanomolar range) for the Shc phosphotyrosine peptide, followed by peptides mimicking Grb2 binding sites on EGF and HGF receptors. Shc is
phosphorylated by the kinase receptor (NGF) upon binding, and serves as a docking protein for the Grb2-Sos complex (Salcini et al., 1994). Binding to the phosphotyrosine motif does not change the affinity of Grb2 SH3 domains for Sos proline-rich motifs (Cussac et al., 1994), but the Grb2-mSos1 complex has higher affinity for phosphotyrosine peptides than Grb2 by itself (Chook et al., 1996).

The DH (DbI homology) domain is the catalytic domain of the DbI family of proteins, the RhoGEFs. The PH domain is invariably located immediately C-terminal to the DH domain. The PH domain, approximately 120 amino acids, was first identified as one of a series of internal repeats of pleckstrin, the major substrate for protein kinase C in platelets. PH domains comprise seven \( \beta \)-strands plus a C-terminal \( \alpha \)-helix. The loops between the strands, variable in length and sequence, dictate the ligand specificities of the domains. The proposed common ligands for the PH domains are inositol 1,4,5-trisphosphate, PIP\(_2\), the \( \beta\gamma \)-subunits of heterotrimeric G proteins and protein kinase C. It has been proposed that binding to inositol phosphates serves to recruit proteins with PH domains to the cell membrane (Hirata et al., 1998). In the study of DbI proteins, it was found that truncation of the PH domain results in a loss of their transforming capability, which can be restored by adding the plasma membrane targeting sequence of H-Ras, suggesting PH domains role in membrane localization (Whitehead et al., 1995).

The DH-PH domain does not seem to be involved with the Ras nucleotide exchange activity of Sos. However, a recent study has shown that this fragment may couple activated Ras to Rac, hence to the Rac effectors and cytoskeleton reorganization (Nimnual et al., 1998).
A recent kinetic study of the interaction between H-Ras and the catalytic domain of a mouse GEF, Cdc25<sup>Mm285</sup> (Lenzen et al., 1998), indicates that the acceleration of GDP release from H-Ras by Cdc25<sup>Mm285</sup> is more than 10<sup>5</sup>-fold. However, the stimulation does not discriminate between the nature of the bound nucleotide, whether it is triphosphate or diphosphate. The nucleotide-release reaction consists of a fast binding step of the GEF to H-Ras and a rate-limiting step that is thought to be a conformational change from a high- to a low-affinity nucleotide binding conformation in H-Ras. The affinity between the nucleotide-free H-Ras and Cdc25<sup>Mm285</sup> (dissociation constant 4.6nM) is 500-fold lower than the Ras-GDP interaction. The reverse reaction then begins with the formation of the ternary complex of Ras-nucleotide-Cdc25<sup>Mm285</sup> followed by the rate-limiting conformational change of Ras. Whether Ras/GDP or Ras/GTP is generated reflects the cellular concentrations of GDP vs. GTP, and also depends on the activities of other regulatory proteins such as GAPs. Of course, the localization of GEFs to the cell membrane after activation may play a role in the outcome. This kinetic view is further confirmed by the crystal structure of the H-Ras/hSos1 complex (Boriack-Sjodin et al., 1998; see section 1.2.3a).

1.1.2c Ras effectors

Ras/GTP is capable of interacting directly with a number of proteins: Raf-1, PI3-K, p120GAP, RalGDS and possibly certain MEKKs. The extent of activation of these parallel paths depends on many factors, such as the type and number of receptors involved, or the cell types, which dictate the preferential expression of certain isoforms over others. In addition, these paths may come to reinforce each other. All the proposed
Effectors here can be activated through other means that do not involve Ras activation. A very simplified picture is provided here. Note that some studies were in vitro binding assays, some were in vivo co-immunoprecipitation upon over-expression of one or more proteins involved. Raf is the only true Ras target protein that has been verified by genetic studies (Vojtek et al., 1998).

i. Raf/MEK/ERK1/2 pathway

In addition to recruitment to the plasma membrane by Ras/GTP, full activation of Raf-1 serine/threonine kinase requires phosphorylation of certain amino acids. Activated Raf-1 then phosphorylates MAP kinase kinases (MEK1 and MEK2) which in turn activate p42 and p44 MAP kinases also referred to as ERKs (extracellular signal regulated kinases). ERKs are responsible for phosphorylating cytoplasmic ribosomal S6 kinase (Rsk) (Sturgill et al., 1988) and MAP kinase interacting kinase (Mnk) (Wang et al., 1998; Waskiewicz et al., 1997) as well as stimulating various transcription factors in the nucleus.

Raf:

There are three conserved regions (CRs) in Raf-1. The N-terminal CR1 has the RBD (Ras-binding domain 51-131) followed by CRD (cysteine-rich domain 139-184) (Morrison et al., 1997). A single R89L mutation in Raf prevents Ras-mediated activation in Sf9 insect cells. RBD acts as a potent competitive inhibitor of p120GAP and neurofibromin in Ras binding (Warne et al., 1993). The crystal structure of Rap1A/GppNp (overall 50% identity to Ras, the same effector region though) complexed with Raf1-RBD confirms that the complex interaction also spans the switch I region of
Rap1A (Nassar et al., 1995). CRD also preferentially binds to the active form of Ras (Brta et al., 1995; Drugan et al., 1996). In fact, Ras binding to both the RBD and the CRD appears to be required for Raf-1 activation. In the full-length Raf-1 molecule, CRD is inaccessible for Ras binding unless RBD-mediated binding has occurred. However, CRD associates with different residues (amino acids 26-28 and 40-49) of Ras than does RBD (Hu et al., 1995).

CR2 (residues 254-269) is rich in serine/threonine residues, some of which are phosphorylation sites. The C-terminal CR3 is the kinase domain. It has not only serine phosphorylation sites, but also tyrosine phosphorylation sites (Morrison et al., 1997).

Regulatory aspects of Raf-1 activation are far from fully worked out. It is thought that first, Raf-1 translocation to the cell membrane is essential for its activation since addition of the Ras C-terminal plasma membrane targeting sequence (CAAX) to Raf-1 converts it to a potent transforming protein (Stokoe et al., 1994; Leevers et al., 1994). Second, phosphorylation of Tyr 340 and 341 enhances the catalytic activity of Raf-1. Possible kinases are Src (Marias et al., 1995) and Janus kinase (Xia et al., 1996). While there are indications that some Ser/Thr phosphorylations lead to Raf-1 inactivation (Samuels et al., 1993, Hafner et al., 1994), the interpretation remains controversial, since two serines, 43 and 621, are both constitutively phosphorylated. The conformational changes induced in Raf-1 activate its kinase domain which has specificity for a proline-rich region in MEK1 and MEK2 only (Denhardt, 1996). However, there still are many questions to be answered. In vivo, Raf-1 is constitutively associated with a 14-3-3 protein as well as with heat shock proteins hsp90 and hsp50. The precise function of these interactions remains to be studied. In addition, the three mammalian isoforms Raf-1, A-
Raf and B-Raf may be involved in unique regulatory events even though they share some common steps (Morrison et al., 1997).

Raf may well be another intersecting point of many signalling events. The activation of Raf-1 in both the receptor tyrosine kinase- and protein kinase C-mediated cases requires the binding of Raf-1 to Ras/GTP. But they differ in that the latter is not blocked by dominant negative N17Ras (Marias et al., 1998). Activation of Raf-1 through non-receptor tyrosine kinases (Src) during mitosis is Ras-independent. These activated Raf-1 molecules are predominantly located in the cytoplasm and they do not lead to stimulation of the MEK/ERK pathway (Ziogas et al., 1998).

MAPK:

MAP kinases (MAPKs) comprise a distinct group of Ser/Thr kinases. They have been classified into three subfamilies: mitogen-responsive ERKs, including ERK1 and ERK2; stress-responsive SAPKs (stress-activated protein kinase), also called c-jun N-terminus kinases (JNKs); and p38 kinase. They are characterized by the dual phosphorylation motif Thr-X-Tyr, where X is Glu for ERKs, Pro for JNKs and Gly for p38. Phosphorylation on both threonine and tyrosine residues is essential for MAPK activation. Activated MAPKs translocate to the nucleus where they phosphorylate and activate transcription factors and other target proteins. The activated transcription factors, in turn, bind to DNA elements and modulate gene expression (Efimova et al., 1998; Su et al., 1996).

Different subfamilies of MAPKs belong to distinct and only partially overlapping kinase cascades. While the ERKs are phosphorylated by MEK, one of a subset of protein kinases that possess unique dual specificity for MAPKs, another protein of the same
subset SEK/JNKK acts on the JNK/SAPKs. Whereas Raf activates MEK, SEK/JNKK is preferentially activated by MEKKs, not by Raf proteins (Minden et al., 1994a). Two pathways seem to involve Ras/GTP: Ras/Raf/MEK/ERK and Ras/MEKK/JNK/JNK. Although activated Ras is sufficient for full ERK response it only produces a weak JNK activation (Minden et al., 1994b), which is more strongly induced by various stress factors and cytokines in a largely Ras independent manner. In some cases, members of the Rho branch act synergistically with Ras in activating JNK. Racl and the Cdc42 active forms can stimulate SAPK/JNK more potently than Ras, but neither initiates the ERK cascade (Lopez-Ilasaca et al. 1998).

ii. PI-3 kinase (PI3-K)

PI3-K is another one of the extensively studied Ras effectors. It is a lipid kinase with specificity for the D3 position of the inositol ring, generating the second messenger PIP,₃. In vivo, dominant negative S17N Ras inhibits growth factor induced PIP₃ production (Rodriguez-Viciana et al., 1994). PI3-K consists of two subunits, the N-terminal regulatory subunit p85 and the catalytic subunit p110. P85 possesses one SH3 and two SH2 domains, and a domain that may have GAP activity towards an unidentified target. Ras/GTP interacts directly with p110 through the effector region (32-40) (Rodriguez-Viciana et al., 1994).

If detached from the extracellular matrix, epithelial cells enter programmed cell death. However, activated Ras can rescue the cells from apoptosis through the PI3-K/PKB pathway, not the Raf pathway (Khwaja et al., 1997). This is demonstrated by expressing constitutively activated PI3-K (p110K227E) or PKB; also by studying Ras
effector mutants (Kauffmann-Zeh et al., 1997). Each Ras effector mutant resides in a constitutively activated V12 Ras background. The partial loss-of-function mutants selectively interact with effectors: V12S35Ras with Raf, V12C40Ras with PI3-K and V12G37Ras with RalGDS. Expression of V12C40Ras protected cells from c-Myc-induced death as effectively as a constitutively active PI3-K, the p110K227E mutant. Interestingly, activation through the Raf pathway promotes apoptosis.

Phosphorylated inositol lipids interact directly with the serine/threonine protein kinase PKB/Akt. Binding of the N-terminal PH domain of PKB to PIP₂ or PIP₃ localizes PKB to the membrane where it is phosphorylated both in its kinase domain and its regulatory domain by PKB kinase (Hemmings, 1997; Stephens et al., 1998). Membrane localization of PKB by addition of a myristoylation sequence causes activation due to constitutive phosphorylation (Khon et al., 1996). This active PKB moves on to phosphorylate down-stream targets: glycogen synthase kinase 3 (GSK3) among others, until inactivated with dephosphorylation by protein phosphatase 2A (PP2A). Phosphorylation inactivates GSK3, leading to stimulation of glycogen synthesis.

The Sos PH domain is another target of PIP₃ and/or PIP₂ (see later this section). PIP₃ is also capable of binding with high affinity to the SH2 domains of proteins such as Src and the p85 subunit of PI3-K. It can compete with tyrosine-phosphorylated proteins for binding to these sites (Rameh et al., 1995). This perhaps serves to amplify the signal.

### iii. Crossing over to the Rho branch GTPase pathways

Instead of using the Raf-1/ERK pathway, Rac and Cdc42, members of the Rho family, activate MEKK (MEK kinase), JNKK/SEK (JNK kinase or SAPK/ERK kinase),
and JNK/SAPK, resulting in phosphorylation of the c-Jun transcription factor. This leads to increased transactivation activity of the protein, hence increased transcription from AP-1 promoter sequences. Besides gene expression, Rac also regulates actin filament accumulation at the plasma membrane to produce lamellipodia and membrane ruffles; and Cdc42 stimulates the formation of filopodia. Rho controls the assembly of actin stress fibers and focal adhesion complexes (Olson et al., 1995). Cell division requires not only synthesis of new proteins, but also cytoskeleton reorganization. It is not surprising that Ras and Rho pathways can be coupled.

The fact that Ras, but not Raf-CAAX stimulates JNK1 implies some downstream effector of Ras is connected to Rac, and this is required for Ras-induced transformation (Olson et al., 1995). Two candidates for such an effector are p120GAP coupled to p190 and Sos.

**P120GAP and p190**

P120GAP is a negative regulator of Ras, but it may also be an effector. It has a SH2-SH3-SH2 region N-terminal to its catalytic domain, and is found to directly bind to autophosphorylated β-PDGF receptors *in vivo*.

A protein p190 is found to associate constitutively with the N-terminal domain (1-445 amino acids) of p120GAP in the absence of serum stimulation (McGlade et al., 1993). The N-terminus domain of p190 contains the GTP binding motif (GXXGXXGKS) that is found in all known GTPases. It is followed by a 778aa segment highly identical (95%) to human transcriptional repressor GRF1, including 5 CXXC and 2 C/HXXC/H Zn²⁺ finger DNA-binding domains. The C-terminal domain sequence is very similar to RhoGAP. Indeed, p190 displays GAP activity on the Rho branch proteins *in vitro* (Settleman et al.,
1992b; McGlade et al., 1993). It was observed that approximately 25% of p190 in normal rat fibroblasts are found in the nuclear fraction. It has been hypothesized that binding of Ras/GTP alters p120GAP conformation, exposing the domain required for p190 interaction, thus localizing p190 in the cytoplasm (Settleman et al., 1992a). However, expression of the dominant negative mutant Ras S17N does not interfere with the association, suggesting that the interaction is Ras/GTP independent (Pronk et al., 1993). Therefore the precise reason and consequence of this interaction is still unknown.

However G3BP, a GAP-SH3 domain binding protein, shows sequence similarity to RNA-binding proteins, and interacts with P120GAP in a Ras/GTP dependent manner, supporting a connection between p120GAP and RNA metabolism (Parker et al., 1996). Maybe p120GAP is a true effector of Ras after all.

Sos

The DH-PH domain of hSos can activate Rac, but not CDC42 in vivo if activated Ras is co-expressed and the downstream PI3-K is activated as well (Nimnual et al., 1998). The Sos PH domain is shown to bind tightly and specifically to phosphatidylinositol derivatives but not to Gβγ of G-proteins (Chen et al., 1997; Mahadevan et al., 1995). In addition, the level of activation of PI3-K by Ras mutants correlates with actin rearrangement, acting through Rac (Rodriguez-Viciana et al., 1997).

iv. RalGDS

Two Ral proteins, RalA and RalB (85% sequence identity), constitute a subfamily of the Ras branch of GTPases. Unlike Ras, most of the Ral proteins are in intracellular vesicles. They appear to regulate the activity of exocytic and endocytic vesicles. RalGDS
is a Ral-GEF identified from yeast, and is found to preferentially bind the active form of Ras. In transient expression assays in COS cells, transfected Ras enhances RalGDS nucleotide exchange activity in vivo (Urano et al. 1996). In these assays, the inhibition of RalGDS nucleotide exchange activity by dominant-negative Ral mutants suppresses oncogenic transformation by Ras ~50-75%. Ral regulates CDC42 activity by interacting with RalBP1 (Ral-binding protein 1), a CDC42 GAP. The details of interaction remain to be established. This may be yet another link from Ras to actin cytoskeleton organization.

In both normal and Src transformed cells, phospholipase D (PLD) can be immunoprecipitated with Ral by anti-RalA antisera (Jiang et al., 1995). And the first 11 amino acids of Ral seem to be important for this interaction. PLD catalyses the hydrolysis of phosphatidylcholine (PC) to phosphatic acid (PA) and choline. Further conversion of PA into diacylglycerol (DAG) may activate protein kinase C (Feig et al. 1996).

v. Other possible effectors

From bovine brain membrane extract, a GST-H-Ras/GTPγS but not a GST-H-Ras/GDP affinity column is able to select out a 180kDa protein that is almost identical to human AF-6, the fusion partner of the ALL-1 protein (Kuriyama et al., 1996).

Another human protein, identified as Rin1, preferentially binds to the GTP form of H-Ras and is not able to interact with either an effector mutant or a dominant negative mutant of H-Ras. It also competes directly with Raf-1 for binding to H-Ras in vitro (Han et al., 1995).
### Ras/GTP

<table>
<thead>
<tr>
<th>c-Raf1</th>
<th>MEKK</th>
<th>p120GAP</th>
<th>PI3-K</th>
<th>RalGDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEK</td>
<td>SEK</td>
<td>p190</td>
<td>PIP3</td>
<td>Ral</td>
</tr>
<tr>
<td>ERK1/2</td>
<td>JNK</td>
<td>Rac</td>
<td>Sos</td>
<td>RBP1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rac</td>
</tr>
</tbody>
</table>

---

Gene expression

Actin cytoskeleton organization

---

**Fig. 1.2** Ras effector pathways
1.1.2d Negative regulators of Ras

The activation signal generated by Ras/GTP has to be turned off promptly when the extracellular stimuli no longer exist, otherwise it will result in tumorous growth as in the case of Ras mutants. To return to the inactive GDP-bound form, GTP has to be hydrolyzed. Since Ras itself is a poor GTPase, with a $k_{cat}$ of 3.4s$^{-1}$ for N-Ras (Neal et al. 1988), help from GTPase activating proteins (GAPs) is needed. Binding of GAP to Ras increases the hydrolysis rate of GTP by five orders of magnitude (Neal et al. 1988).

i. Functional differences between p120$^{\text{GAP}}$ and NF1

The two most studied GAPs in humans are p120$^{\text{GAP}}$ and neurofibromin (NF1). Both proteins are ubiquitously expressed. Subcellular fractionation shows that while p120$^{\text{GAP}}$ is predominantly cytoplasmic, NF1 binds to microtubules (DeClue et al, 1991). Mutations in the NF1 gene manifest themselves in various phenotypes, such as café-au-lait spots, benign neurofibromas and increased incidence of malignant tumors. Sequence alignment identified a conserved region GRD (GAP related domain) that has GAP activity. Malignant schwannoma cell lines have a low level of NF1 expression which causes elevated levels of activated Ras even though both p120$^{\text{GAP}}$ and Ras expressions are normal (DeClue et al, 1992). NF1 Ras activation is inhibited non-competitively by certain lipids such as phosphatidate and phosphatidylinositol-4,5-bisphosphate in micromolar concentrations while p120$^{\text{GAP}}$ is unaffected. In addition, NF1 binds to H-Ras/GTP with higher affinity than p120$^{\text{GAP}}$, and it is a more efficient catalyst (Ahmadian et al, 1996).
1.2 Structural Studies of Ras and Ras-interacting Proteins

1.2.1 X-ray crystallography as a tool in understanding biological interactions

To understand mechanisms of cellular processes, information on the three-dimensional structure of enzymes and other macromolecules is essential. Two major techniques exist for the structural determination of macromolecules at atomic resolution: X-ray diffraction of crystals and nuclear magnetic resonance spectroscopy (NMR). While NMR does not require crystals and provides more information on the dynamics of the molecule, X-ray crystallography can be applied to compounds with molecular weight larger than 30,000, up to at least $10^6$, with no theoretical limit.

1.2.2 Structure of H-Ras from X-ray crystallography

All crystallographic studies done on Ras have utilized the truncated version of the protein where the flexible and highly variable C-terminal 20 amino acids have been eliminated. As mentioned in section 1.1.1a, this tail is essential in membrane localization, but does not play a role in catalysis. The truncated form is more compact and therefore generates better diffracting crystals.
Fig. 1.3 Crystal structure of H-Ras (1-166) with GppNp (Pai et al., 1990)

The core of the molecule (6 beta sheets and 5 alpha helices) are shown in red, while the loops are shown in greenish-blue. GppNp and residues involved in GTP hydrolysis (Gln61, Thr35) are shown in stick models. The catalytic water is drawn in magenta.

Loop 1 (L1) is involved in phosphate binding. L2 belongs to the Switch I region where Ras interacts with its effectors. L4 of the Switch II region is the most mobile part of the whole molecule. It also contains the catalytic Gln61.
Five α-helices and a central six-stranded β-sheet form the rigid protein core and are connected by 10 loops (Fig. 1.3; Pai et al., 1990). Three of these loops are most important for Ras function.

Loop L1 is the phosphate-binding loop. It contains the sequence $^{10}\text{GAGGVGKS}^{17}$ that fits the general guanosine nucleotide-binding motif GXXXXGKS/T (Chardin 1993). Substitutions of the glycines at positions 12 and 13 by other amino acids are found in naturally occurring tumors (Bos 1988). The side chain of the conserved lysine 16 forms an ion-pair with the γ-phosphate and hydrogen bonds with the main-chain oxygens of residues 10 and 11 (Fig. 1.4). Serine 17 is vital for coordinating the $\text{Mg}^{2+}$ ion (Fig. 1.5). Its mutation to asparagine leads to loss of Ras function, becoming the dominant negative inhibitor of Ras due to higher affinity for GEFs (Feig et al., 1988; Powers et al., 1989). The rest of the nucleotide binds to Ras through more hydrogen bonds involving the hydroxyl group of the ribose and the usual H-bond donors and acceptor groups on the base (Fig. 1.4). In addition, the Phe28 ring and the hydrocarbon chain in Lys117 sandwich the base in a hydrophobic pocket.

Amino acids 32-40, containing half of loop L2 and parts of β-strand 2 are also called the “effector loop” due to their involvement in the binding of GAP and other Ras downstream effectors. Mutation studies have found that single-residue changes in this region can cause the loss of the ability to activate one pathway while leaving the other pathways undisturbed (Winkler et al., 1997). Thr35 in this region is involved in H-bonding to the γ-phosphate oxygen in the GTP bound form.

Loop L4 (amino acids 59-64) contains the catalytic residue Gln61. If properly positioned, the Gln61 side chain is hydrogen-bonded to the catalytic water molecule 175,
Fig. 1.4  Schematic drawing of Ras and GppNp interactions.
open arrows = H-bonds
solid arrows = bonds between Mg$^{2+}$ and its ligands
(adapted from Pai et al., 1989)
Fig. 1.5 Scheme for the hexa-coordination of Mg2+ in the complex with Ras/GppNp and GDP as determined by x-ray crystallography (adapted from John et al., 1993)

In both cases, Asp57 and Ser17 contribute to Mg2+ coordination. However, Thr35 is only complexed to Mg2+ in the triphosphate nucleotide conformation.
orientating it and activating it for nucleophilic attack on the γ-phosphate of GTP. With the additional water ligand, the γ-phosphate forms a penta-coordinated transition state. Subsequent hydrolysis occurs when an inorganic phosphate dissociates (Fig. 1.6). Mutations of this residue have been found in a number of tumors. Of the whole structure, electron density for residues 60-65 is least well defined. The high mobility of this loop explains why H-Ras is a poor GTPase by itself. Binding of p120GAP increases the hydrolysis reaction by 5 orders of magnitude by fixing the Gln61 side chain in the optimal orientation and by stabilizing the developing negative charge in the transition state through an arginine side chain (refer to section 1.2.3c). Gln61 interacts with the catalytic water molecule and γ-phosphate in the Ras/GAP complex.

H-Ras binds to GDP or GTP with similar affinities, $6.1 \times 10^{11} \text{M}^{-1}$ and $1.8 \times 10^{12} \text{M}^{-1}$ respectively at 5°C (John et al., 1990). But the addition of the γ-phosphate induces significant conformational changes in the backbone structures of switch I (30-38) and switch II (60-76) regions (Fig. 1.7). By comparing structures from different crystal packings, thus ruling out packing contacts, this difference is identified to be the cause of effector binding to the GTP form, but not GDP form of H-Ras (Milburn et al., 1990).

1.2.3 Interactions of Ras with its regulators and effectors

1.2.3a Complex structure of H-Ras and the hSos1 CDC25H domain

The crystal structure of H-Ras complexed with the catalytic domain of hSos1 in the absence of nucleotides was solved recently (Boriack-Sjodin et al., 1998). The interface is primarily hydrophilic and involves the phosphate-binding loop (P-loop),
Fig. 1.6 Structure of the H-Ras active site (Pai et al., 1990)

Mg$^{2+}$ is coordinated by 2 phosphate oxygens, the hydroxyl groups of Ser17 and Thr35 and two water molecules (not shown here). The catalytic water molecule (orange) is activated by the Gln61 side chain and attacks the gamma phosphate as a nucleophile, causing the hydrolysis of GTP to GDP.
Fig. 1.7  Backbone comparison of H-Ras/GDP and H-Ras/GppNp crystal structures

Red:  H-Ras/GppNp, Pai et al., 1990
Blue:  H-Ras/GDP, Tong et al., 1991

Loop 4 (amino acids 59-64) is the least well ordered region of the whole structure in both the GDP and GTP forms. However, by comparing structures from different crystal forms, thus eliminating packing artifacts, the shifts in both switch regions are significant and are the source of recognition by interacting proteins.
Switch I (residues 25-40) and Switch II (residues 57-75) regions of H-Ras. The Switch I region, which normally sandwiches the nucleotide with the P-loop, is completely removed from the nucleotide-binding site. The usually disordered Loop 4 of the Switch II region is held very tightly by Sos and therefore is well defined.

The catalytic domain of hSos1 stabilizes H-ras in the nucleotide-free state by inserting an α-helix into H-Ras (Fig. 1.8), thus disrupting the original ionic interactions with and H-bonds to the phosphate groups and Mg$^{2+}$ of GDP/Mg$^{2+}$, while leaving the base- and ribose-binding sites unchanged. Therefore it allows nucleotide binding as well as its release, showing no preference for GTP over GDP. Since the cellular concentration of GTP is 100-fold higher than that of GDP, GTP instead of GDP loads to H-Ras (Boriack-Sjodin et al., 1998).

1.2.3b Rap1A/c-Raf1 complex

Rap1A belongs to the Ras branch of the small GTPases superfamily. It has an effector region identical to Ras (residues 32-40) and an overall sequence identity of 50%. Its GppNp-bound form is complexed to the Ras-binding domain (RBD) of the serine/threonine kinase c-Raf1 in the crystal structure (Nassar et al., 1995). The Rap1A structure is very similar to that of H-Ras/GppNp with a water molecule at the same position as the catalytic Wat175 in the uncomplexed H-Ras structure. The 81-amino acid RBD has the ubiquitin superfold.

The interaction is mostly hydrophilic and is mediated by formation of an antiparallel β-sheet by strands B1-B2 from the RBD and strands β2-β3 (a.a. 36-58) from Rap1A. Residues 33-39 from the effector region of Rap1A are in close contact with the
hSos1 stabilizes the nucleotide-free H-Ras by inserting an alpha-helix into the nucleotide binding site, forcing the Switch I region outwards while fixing the usually mobile Switch II in place.
RBD as expected from biochemical studies. The Switch II region, although it shows the most significant change from the GDP-to GTP-bound form, is not directly involved in binding. The RBD has a 1300-fold lower affinity for the GDP-bound Ras, therefore the complex dissociates upon GTP hydrolysis. But the structure does not explain how the high-affinity to low-affinity complex change occurs.

1.2.3c H-Ras/p120\textsuperscript{GAP}-GRD complex

One year into our study of G12A and G13A H-Ras, the crystal structure of the transition state complex of Ras-p120\textsuperscript{GAP}-GRD was published (Scheffzek et al., 1997) (Fig. 1.9 and 1.10). The overall structures are similar to the isolated H-Ras and p120\textsuperscript{GAP}-GRD molecules (Scheffzek et al., 1996).

In the complex structure, the γ-phosphate leaving group is mimicked by AlF\textsubscript{3}. Aluminum can have anywhere from 1 to 6 ligands, so it has the potential to assume a penta-coordinated structure like the phosphate transition state. In addition, F has greater capacity than O for forming H-bonds due to its electronegativity. An Al-F bond has the same length as a P-O bond in phosphate (Chabre, 1990). Ras on its own does not bind or become activated by AlF\textsubscript{4}⁻, unlike the α-subunit of heterotrimeric G proteins (Sondek et al., 1994; Higashijima et al., 1991). However, Ras/GDP interacts with AlF\textsubscript{4}⁻ in the presence of stoichiometric amounts of either p120\textsuperscript{GAP} or NF1. Neither oncogenic Q61LRas nor a non-functional GAP mutant produces such a complex because both Gln61 and the conserved Arg789 are required to stabilize the complex (Mittal et al., 1996) as described above.
Fig. 1.9  Crystal structure of p120GAP-GRD and H-Ras complex (Scheffzek et al., 1997)

The phosphate-binding loop as well as both switch regions are involved in the interaction. The usually disordered 60's region in H-Ras is now fixed in space by GAP.
The side chain of Arg789 (GRD) stabilizes the developing negative charge in the transition state. This Arg finger is fixed in place by the interaction between Arg903 of GRD and the main chain carbonyls around Arg789. The catalytic Gln61 is also fixed in the right orientation by the main chain carbonyl of Arg789. Loop 4 (containing Gln61) is no longer mobile in the complex.

The distance between the carbonyl oxygen of Gln61 side-chain and the catalytic water is 2.93Å. And the shortest distance between the amide N of Gln61 and a F atom of AlF₃ is also 2.93Å.
The interface of the complex consists of weak van der Waals interactions and polar interactions involving residues of the phosphate-binding loop (amino acids 10-16), Switch I (amino acids 30-37), and Switch II (amino acids 60-76) on H-Ras.

Five acidic residues of Switch I (D30, E31, D33, E37 and D38) form a negatively charged patch for interaction with Lys949 of p120\textsuperscript{GAP}-GRD. This region also interacts with Ras effectors.

Unlike in the isolated Ras structure, the Switch II region is well defined in the complex. Residues 61 to 63 are forced to form a short $3_1\alpha$ helix. Arginine 789, conserved in all Ras-GAPs, neutralizes the developing negative charge on the leaving $\gamma$-phosphate with its side chain, while its main chain carbonyl group is H-bonded to the amide of the catalytic Gln61 of H-Ras. This fixes Gln61 in the right orientation to interact with the catalytic water molecule, thus activating it for nucleophilic attack on the $\gamma$-phosphate. This arginine finger is stabilized by the ionic interaction of Arg903 with the main chain carboxyls of residues 788-790.

The phosphate-binding loop is part of the active site. Therefore it is in close contact with p120\textsuperscript{GAP}-GRD. From the complex structure, it appears that any side chain larger than a hydrogen atom at position 12 of H-Ras will be within van der Waals distance from both the carbonyl of Arg789 and the side chain of Gln61. This closeness only seems to interfere with the catalysis in the transition state not the ground state, since the glycine 12 mutants of Ras bind to GAP with similar affinity as the wild type (Gideon et al., 1992) without GTP hydrolysis. However, the structure did not provide an explanation as to why the G12P (Gly12 to Pro) mutant is not oncogenic even though proline is sterically bulkier than glycine.
Since the oncogenic Ras mutants still bind to GAP, several laboratories, including this one, have tried to crystallize the ground state complexes of the H-Ras with different GRD domains, though without much success due to the low affinity at the ground state. However, the structures of both the ground state and the transition state complexes of Rho, another small G protein, with its rhoGAP have been solved (Rittinger et al., 1997a, b). Even though the Ras- and Rho-specific GAPs share no significant sequence or tertiary structure similarity (other than both being highly helical), both accelerate GTP hydrolysis by contributing a catalytic arginine residue and stabilizing the Switch I and II regions of the G protein. Comparison of the two complexes show that there is a 20° rigid body rotation of the two proteins relative to each other around an axis close to the phenolic hydroxyl of Tyr66Rho (Switch II) upon transition state formation (Fig. 1.11). In addition, the protein-protein interface becomes more extensive, supporting the 100-fold enhancement in binding at the transition state. This large rearrangement brings the catalytic Arg85rhoGAP side chain into the right position to interact with β- and γ-phosphates and its main chain carbonyl with the side chain of catalytic Gln63Rho. One would expect a similar rearrangement for Ras-RasGAP complex from the ground to the transition state.

1.3 Mutation studies of H-, K- and N-Ras

The N-terminal 85 residues of all Ras isoforms, containing the two switch regions are identical; the next 80 residues are 95% conserved. When activated, all three interact with the same set of effectors: Raf, PI3-K, RalGDS and AF6, with indistinguishable
Fig. 1.1  Backbone comparison of the ground state and the transition state complexes of a rho and rhoGAP (Rittinger et al., 1997a and b)

light blue:  Cdc42 (ground state)
ligh pink:  RhoA (transition state)
Blue:  rhoGAP in ground state
Cyan:  rhoGAP in transition state
Dark blue:  GppNp (ground state)
Sky-blue:  GDP (transition state)

The Switch II region remains at the same place in both complexes due to interaction with rhoGAP. The Rho molecules undergo a 20 degree rotation around an axis (normal to the paper) in going from the ground to the transition state where there is a more extensive protein-protein interface. The rhoGAPs are aligned with LSQ command in O, with r.m.s. fit of 0.928. The transformation matrix is then applied to the whole complexes.
binding affinities. Residues 166 to 185 are the hypervariable region (HVR) that has only 15% sequence identity among the isoforms. HVR dictates different post translation processing and means of membrane localization. It appears that the varying length of the HVR leads to a different efficiency of different isoforms in activating Raf-1 and PI3-K (Yan et al, 1998).

Specific Ras isoforms are mutated in different tumors: 50% of colon cancers and 90% of pancreatic cancers have Ki-ras mutations; N-ras mutations occur in 25% of acute leukemias (Yan et al, 1998). The most frequent mutations occur at codons 12, 13 and 61 (Table 1.1).

Since all Gly12 and Gly13 mutants bind to GAP with almost wild-type affinity, it appears that the Ras/RasGAP ground state complex tolerates large side chains at these positions.

1.3.1 Glycine 12

Gly12 mutations have been studied extensively. 19 mutants were constructed and inserted into the c-H-Ras1 gene of rat fibroblasts through in vitro recombination, and assayed for their ability to induce malignant transformation (Table 1.2). All but mutation to proline (G12P) proved to be oncogenic (Seeburg et al, 1984). Crystal structures of three Gly12 mutants G12R, G12V and G12D bound to GppNp have been determined (Krengel et al, 1990). It was found that the large side chains all prevented the proper binding of the catalytic water molecule, thus interfering with GTP hydrolysis. The triphosphate nucleotide GppNp, a non-hydrolyzable analogue of GTP, is still bound.
Since the large side chains may diminish GTP-binding, but not abolish it, the mutant Ras remains in the active form.

Franken et al. (1993) studied the G12D and G12P mutants extensively and found that G12P has a slightly higher intrinsic GTPase activity than the wild type (~2-fold). In addition, high concentrations of p120\textsuperscript{GAP} give G12P a 5-fold stimulation of the GTPase activity, but not to G12D which also has a 4.5-fold reduction in the intrinsic GTPase activity. Both mutants still bind to p120\textsuperscript{GAP} with almost wild-type affinity. X-ray crystal structures show that G12P/GppNp has a very similar active site to the wild type, including the position of the catalytic water.

1.3.2 Glycine 13

Gly13 mutations are less well characterized. Hybridization of DNA oligomer probes with genomic DNA detected G13D and G13V mutants of N-Ras in acute myeloid leukaemia (Bos et al, 1985). G13T and G13S mutants have intrinsic GTPase activity two or three times that of wild-type Ras, but they are not activated by GAP and they are transforming in a germinal vesicle breakdown assay with *Xenopus laevis* oocytes (Chung et al, 1993).

1.3.3 Aim of this thesis

Although no biochemical study has been done on the activation of G12A with very high concentrations of GAP, it is reasonable to conclude that the transforming ability of G12A comes from its resistance to GAP stimulation. Alanine has the next smallest side chain to glycine. If a methyl group at position 12 is too big for the catalytic
site, so should be the rather bulky 5-membered pyrrolidine ring in proline. But that does not seem to be the case. Through the study of the G12A mutant structure and modeling of the rigid proline ring, we would like to see whether a slight change in the C9 positions marks the line of oncogenicity.

By analyzing the structure of G13A, the smallest disturbance one can study with X-ray crystallography at this site, we set out to test whether mutation at the 13th position alters the loop L1 conformation relative to the wild-type Ras as proposed by Chung et al (1993), as the reason for rendering the molecule oncogenic.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Codon</th>
<th>Mutation</th>
<th>Amino acid change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. <em>in vivo</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H-Ras</td>
<td>12</td>
<td>GGC-GTC</td>
<td>Gly-Val*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGC-GAC</td>
<td>Gly-Asp*</td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>CAG-AAG</td>
<td>Gln-Lys*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAG-CGG</td>
<td>Gln-Arg*</td>
</tr>
<tr>
<td>K-Ras</td>
<td>12</td>
<td>GGT-CGT</td>
<td>Gly-Arg*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-AGT</td>
<td>Gly-Ser*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-TGT</td>
<td>Gly-Cys*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-GAT</td>
<td>Gly-Asp*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-GTT</td>
<td>Gly-Val*</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>GGC-GAC</td>
<td>Gly-Asp*</td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>CAA-CAT</td>
<td>Gln-His*</td>
</tr>
<tr>
<td>N-Ras</td>
<td>12</td>
<td>GGT-AGT</td>
<td>Gly-Ser*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-GTT</td>
<td>Gly-Val*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-GAT</td>
<td>Gly-Asp*</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>GGT-CGT</td>
<td>Gly-Arg*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-GTT</td>
<td>Gly-Val*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-GAT</td>
<td>Gly-Asp8*</td>
</tr>
<tr>
<td>B. <em>in vitro</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H-Ras</td>
<td>12</td>
<td>various</td>
<td>all except Pro*</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>GGC-GAC</td>
<td>Gly-Asp*</td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>various</td>
<td>all except Glu &amp; Pro*</td>
</tr>
<tr>
<td>?:Ras</td>
<td>13</td>
<td>GGT-GAT</td>
<td>Gly-Thr$^{5}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGT-GTT</td>
<td>Gly-Ser$^{5}$</td>
</tr>
</tbody>
</table>

**Table 1.1**  Transforming Ras mutants (*Bos* 1988*; Bos *et al.*, 1985; *Chung et al.*, 1993*)

?: unspecified
<table>
<thead>
<tr>
<th>Residue 12</th>
<th>Activated genes</th>
<th>Focus formation</th>
<th>Agar growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td>H-, K-, N-Ras</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ile</td>
<td>H-Ras</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Val</td>
<td>H-, K-Ras</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Leu</td>
<td>H-Ras</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Ser</td>
<td>K-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Thr</td>
<td>H-Ras</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Ala</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Met</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Cys</td>
<td>K-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Tyr</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Phe</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Trp</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Pro</td>
<td>H-Ras</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lys</td>
<td>v-BALB-ras</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Arg</td>
<td>H-Ras, Ra-Ras</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>His</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Asp</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Glu</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Asn</td>
<td>H-Ras</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Gln</td>
<td>H-Ras</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**Table 1.2**  
Transformation properties of Ras Gly12 mutants

(Focus formation was determined by the ability of ras-cotransfected Rat-1 cells to form foci. Agar growth was assessed by plating individual ras-cotransfected subclones, or cotransfected mass cultures into soft agar at two densities and scoring colonies of >30 cells after 2-3 weeks.)

Adapted from Table 1 in Seeburg et al., 1984. Also see the reference within.
Chapter 2

METHODS AND RESULTS

2.1 Mutation of the H-ras expression vector

The expression vector ptacRas, derived from pKM1, encodes the truncated form of H-ras (amino acids 1-166) (Tucker et al., 1986; John et al., 1988; John et al., 1989). However, some incomplete (tac promoter region) and incorrect (3' end of the ras gene) regions had to be identified by sequencing (Dorothe Vogt, unpublished results, see Appendix B).

2.1.1 G12A

G12A was generated with the QuickChange™ site-directed mutagenesis kit from Stratagene, and the supplier’s recommended procedure was followed. The primer was designed to contain the single amino acid substitution, as well as the only SacII endonuclease site in the mutated plasmid.

\[
\begin{align*}
5' & \quad GTG GTG GGC \textcolor{red}{\text{GCC GCG GGT}} GTG GGC AAG \\
& \text{SacII codon 12} \\
3' & \quad CAC CAC CCG \textcolor{red}{\text{CGG CGC CCA}} CAC CCG TTC \\
\end{align*}
\]

Pfu DNA polymerase was used to extend the mutagenic primers to cover the entire plasmid, resulting in nicked circular double stranded DNA. The mother plasmid was
digested away by DpnI that is specific for methylated DNA. Then the nicked circular dsDNA was transformed into XL2-Blue ultracompetent cells.

2.1.2 G13A

Similar primers were designed for the G13A mutation. However, the Stratagene kit did not work here because the primer designed according to the published sequence did not match the real plasmid. In addition, the ptacRas vector was contaminated with an unknown plasmid during the original construction (personal communication with Robert Cool). The conventional PCR and ligation strategy was employed instead. Even then, the miniprep plasmid DNA (Qiagen spin column) from the WT H-Ras expression cell line had to be re-transformed into XL Blue cells many times before a pure ptacRas DNA stock was obtained.

5' end

AGG AAA CAG AAT TCT ATG ACA GAA TAC AAG CTT GTT GTT

EcoRI

GTT GGC GCC GGC GCT GTG GGC

codon 13

3' end

CAC CAG CAC CAT GGG CAC GTC

NcoI

Taq DNA polymerase (NEB) was used to synthesize the 340bp fragment by PCR (Stratagene Robocycler 40), which was then digested with EcoRI and NcoI (NEB) for 1h at 37°C in EcoRI buffer supplied and agarose gel purified (Qiagen kit). The PCR gave a
workable yield only when the WT EcoRI/NcoI fragment was used as the template, another misbehavior of ptacRas. The clean PCR fragment was ligated into linearized (EcoRI/NcoI digested also) ptacRas vector with various ratios using T4 ligase (NEB). After overnight ligation at 16°C, the mixture was transformed into XL2-Blue cells with the standard heat shock procedure.

**PCR cycle:** (30sec at 92°C 1min at 55°C 1.5min at 72°C) x 25 cycles

**PCR condition:**
- Taq polymerase: 2.5 units
- Taq buffer: as specified by the manufacturer
- dsDNA template: 1ng (the EcoRI/NcoI fragment)
- Primers: 50pmol each
- dNTP: 0.8mM
- MgCl₂: 0.5mM
- Total volume: 50μL

**EcoRI/NcoI double digest:** 20 units of each enzyme used per 50μL of PCR product

### 2.2 Expression and purification of G12A and G13A

*E. coli* cells harboring the expression vectors for the respective mutant were grown in 2L of super-rich media at 37°C and induced overnight with 1mM IPTG after the O.D.₆₀₀ reached 0.7.

**Super-rich medium (1L):**
- Trypton: 12g
- Yeast extract: 24g
- Glucose: 5g
Glycerol 4mL

900mL, autoclave

then add 100mL sterile filtered potassium phosphate
buffer pH7.6

The cells were harvested by centrifuging at 4000rpm for 30min. The cell pellet was re-suspended in 64mM Tris pH7.5, 1mM DTT, 1mM PMSF and 1mM EDTA; then French pressed twice at 15,000psi. The lysate was spun at 17,000rpm for 1 hour and the supernatant was loaded onto a 150mL Q-sepharose column (Pharmacia) that had been equilibrated with 3 column volumes of buffer A (64mM Tris pH7.5, 1mM DTT, 10mM MgCl2). After washing the column with 1 column volume of buffer A, a linear gradient was run from 0% to 100% buffer B (64mM Tris pH7.5, 1mM DTT, 10mM MgCl2, 1M NaCl) over 3 column volumes. The fractions containing H-ras (at ~80% bufferB) were pooled and solid ammonium sulfate was added, first to 50% concentration to remove some unwanted proteins, then to 70% to concentrate the semi-purified protein mixture for the next column. The ammonium sulfate precipitate was re-dissolved in buffer C (64mM Tris pH7.5, 1mM DTT, 40mM MgCl2, 200mM NaCl and 50μM GDP), and loaded onto a Superdex 75 column (Pharmacia) that was pre-equilibrated with the same buffer (Fig. 2.1). The mutant Ras proteins eluted at the same positions as the WT protein from both columns. The clean Ras protein was pooled and concentrated with Biomax5K (Millipore) to about 20mg/mL.
(a). Elution profile of H-Ras from the Q-sepharose column

(b). SDS-PAGE of Q-sepharose fractions around the H-Ras peak. Each lane represents 10μL of a 6mL fraction.
(The two fractions circled in red, 6mL each, contain H-Ras protein and are pooled for the ammonium sulfate precipitation.)

Fig. 2.1 Purification of H-Ras protein
(c). Elution profile of H-Ras from the superdex-75 column

(d). SDS-PAGE of fractions containing H-Ras protein from the Superdex-75 column. Each lane represents 10µL of 1mL fraction.

Fig. 2.1 Purification of H-Ras protein
(See text for more detailed description.)
2.3 Nucleotide exchange

EDTA was added to 1mL of concentrated Ras solution (about 20mg/mL) to a final concentration of 50mM. A NAP-10 column (Pharmacia), pre-equilibrated with 50mM Tris pH8.5, 200mM ammonium sulfate, 5mM DTT and 10µM ZnSO₄, was used to exchange buffer so calf intestine alkaline phosphatase (CIAP) (Boehringer Mannheim) can hydrolyze GDP into GMP. The buffer exchanged Ras solution with 1mM GppNp and 0.2units CIAP/1mg Ras was incubated at room temperature for 40min. The reaction was stopped by a second buffer exchange (PD-10 column, Pharmacia) to the final buffer of 64mM Tris pH7.5, 10mM MgCl₂ and 2mM DTT. The manufacturer’s protocols were followed for both the NAP-10 and PD-10 buffer exchanges. Ras/GppNp was collected, concentrated to 16mg/mL and stored in liquid nitrogen.

2.4 Crystallization

2.4.1 G12A

G12A/GDP was crystallized using the batch method at room temperature. 3µL of 16mg/mL protein was mixed with 6µL of 43% PEG400 and 1µL of 1% β-OG in depression-well slides (Fig. 2.2). The best single crystals appeared after 3 to 5 days, continued to grow for about 2 weeks and reached a maximum size of 1x0.3x0.3mm³. They belong to space group P6₁ (2 molecules per asymmetric unit) with unit cell axes of a=126.3Å, b=126.3Å, c=45.1Å, and diffracted to 2.5Å resolution.

G12A/GppNp crystals could be grown from identical conditions, except the crystals were always small and highly twinned (Fig. 2.3). Various concentrations of
protein and PEG, different types of PEG and other precipitants, different temperatures and seeding were explored, but none gave significant improvement.

2.4.2 G13A

In an attempt to crystallize mutant-H-Ras/NF1-GRD complexes (Appendix A, part II), we obtained G13A/GDP crystals from hanging-drops at room temperature (Fig. 2.4). 17.5mg/mL G13A/GDP was mixed with purified NF1-GRD protein at a 1:1 molar ratio. 2μL of this mixture was combined with 2μL of the precipitant (100mM HEPES pH7.5, 1.65M Li₂SO₄, 2mM AIF₄⁻ (2mM AIF₃ with 20mM NaF)) to form the hanging-drop, and 0.5mL of the precipitant was used as the reservoir. Crystals appeared overnight and grew to full sizes in two weeks.

These crystals have a different packing compared to the G12A/GDP ones. Even though the high concentration of salt in the crystallization condition inhibited the formation of stable H-Ras/NF1-GRD complexes, NF1-GRD molecules did interact with H-Ras transiently. This effectively changed the solubilities and diffusion properties of both populations of proteins in the drop. In a way, NF1-GRD acted as a protein buffer used in crystallization. These crystals belong to space group R32, with unit cell axes of a=92.9Å, b=92.9Å, c=120.6Å, and diffracted to 1.8Å resolution.

G13A/GppNp crystals, grown from the same condition as G12A/GppNp, suffered the same fate as well (Fig. 2.5).
Fig. 2.2  G12A H-Ras/GDP crystals (1mm*0.3mm*0.3mm after 3 weeks)  
(Crystals obtained from depression well slides with ~25% PEG400 as the precipitant.)

Fig. 2.3  G12A H-Ras/GppNp crystals (0.05mm in the longest dimension.)  
(Crystals were grown from depression well slides with PEG400 as the precipitant.)
Fig. 2.4  G13A H-Ras/GDP crystals (0.5mm*0.5mm*0.3mm after 2 weeks)
Crystals were grown from hanging-drops with NF1-GRD as a protein buffer.
See the text for details.

Fig. 2.5  G13A H-Ras/GppNp crystals (0.03mm in the longest dimension)
These crystals were grown from depression well slides with PEG400 as the precipitant.
2.5 Data collection, processing, molecular replacement and refinement

A Gl2A/GDP diffraction data set (2.5Å) was collected at the CHESS synchrotron beam line F1 (Cornell University, Ithaca, NY) with a frozen crystal. The crystallization solution itself served as the cryoprotectant. A G13A/GDP data set (1.8Å) was collected at the BNL (Brookhaven National Laboratory, Long Island, NY) beam line X8C. The crystal was frozen with 100mM HEPES pH7.5, 1.8M Li$_2$SO$_4$ and 18% glycerol without soaking.

The data were processed with the program packages DENZO and Scalepack (Otwinowski et al., 1997). Molecular replacement solutions were found with AMoRe (automatic molecular replacement package, Navaza 1994), using the wild type H-Ras/GDP structure (PDB code 4Q21) as the search model. The solutions were refined with the help of the program CNS (Crystallography & NMR System, Brünger et al., 1998) following standard protocols (Table 2.1). The first step was torsion angle simulated annealing at 2500K with slowcool of 25K drop per cycle. The model was then displayed using O (Jones et al., 1991) and a sigmaA weighted 2Fo-Fc electron density map calculated by CNS. Some side-chain orientations were adjusted manually. Most of the peptide main chain agreed with the density except for the Switch II region (amino acids 61-67 for G13A, amino acids 61-65 and 260-266 (the second chain) for G12A). The new model then underwent several rounds of conjugate gradient refinement, with maps calculated and more manual adjustments each round until there was no obvious disagreement between the map and the model. Then another round of simulated annealing was performed to remove the side chains from possible traps of energetically unfavorable geometry. Water molecules were picked using the water-pick program in
CNS with default settings, and the list was edited according to the density map. In total, 26 water molecules were included in G12A and 160 in G13A. This was followed by group-B-factor refinement for G12A and individual B-factor-refinement for G13A and more energy minimization until both the R and $R_{\text{free}}$ factors converged.

Since the G12A crystal form had 2 molecules per asymmetric unit, restrained non-crystallographic symmetry was used during CNS refinement. The restrain weight was optimized by trial and error with the first annealing step. Weight = 100 was the one that gave the lowest $R_{\text{free}}$ and thus was used for the whole refinement.

The Ramachandran plots, calculated with Procheck (Fig. 2.6 and 2.7) show that most residues have the favored geometry except the ones that belong to the Switch II region (amino acids 61-65 and 260-266 (of the second molecule in the asymmetric unit) of G12A and amino acids 61-67 of G13A). These loops are highly mobile in both structures.
G12A

Rotation and translation to be applied to the input coordinates

\[ x_{\text{out}} = \text{rotation} \ast x_{\text{in}} + \text{translation} \]

solution with 2 fragments:

\[
\begin{align*}
\begin{array}{ccc}
-0.82638 & -0.40344 & -0.39285 \\
-0.09123 & -0.59251 & 0.80038 & \ast & \text{xin}\text{(mod=11)} + \\
-0.55567 & 0.69726 & 0.45283 \\
-0.42967 & -0.73024 & 0.53116 \\
-0.70782 & -0.09288 & -0.70026 & \ast & \text{xin}\text{(mod=11)} + \\
0.56070 & -0.67684 & -0.47697 \\
\end{array}
\end{align*}
\]

\[ \text{c-factor} = 66.6, \quad \text{r-factor} = 35.6 \]

G13A

Rotation and translation to be applied to the input coordinates

\[ x_{\text{out}} = \text{rotation} \ast x_{\text{in}} + \text{translation} \]

solution with 1 fragment:

\[
\begin{align*}
\begin{array}{ccc}
-0.30072 & 0.34958 & -0.88733 \\
0.28721 & -0.85401 & -0.43379 & \ast & \text{xin}\text{(mod=11)} + \\
-0.90944 & -0.38530 & 0.15641 \\
\end{array}
\end{align*}
\]

\[ \text{c-factor} = 53.4, \quad \text{r-factor} = 39.5 \]

Table 2.1 Molecular replacement solutions from AMoRe
<table>
<thead>
<tr>
<th></th>
<th>G12A</th>
<th>G13A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit cell Axes (Å)</td>
<td>a=126.3, b=126.3, c=45.1</td>
<td>a=92.9, b=92.9, c=120.6</td>
</tr>
<tr>
<td>Unit cell angles (°)</td>
<td>α=90, β=90, γ=120</td>
<td>α=90, β=90, γ=120</td>
</tr>
<tr>
<td>Space group</td>
<td>P6₃</td>
<td>R32</td>
</tr>
<tr>
<td>Resolution*</td>
<td>2.5Å (2.56Å-2.50Å)</td>
<td>1.8Å (1.84Å-1.80Å)</td>
</tr>
<tr>
<td>R_{sym} (overall)</td>
<td>6%</td>
<td>5%</td>
</tr>
<tr>
<td>R_{sym} (highest shell)</td>
<td>11.7%</td>
<td>20.6%</td>
</tr>
<tr>
<td>Completeness (over all)</td>
<td>96.1%</td>
<td>100%</td>
</tr>
<tr>
<td>Completeness (highest shell)</td>
<td>91.8%</td>
<td>99.9%</td>
</tr>
<tr>
<td>I/σ (highest shell)</td>
<td>1887/227</td>
<td>1460/160</td>
</tr>
<tr>
<td>Test set</td>
<td>1357*(9.4%)</td>
<td>1825*(9.7%)</td>
</tr>
<tr>
<td>R_{crys}</td>
<td>24%</td>
<td>21%</td>
</tr>
<tr>
<td>R_{free}</td>
<td>27%</td>
<td>25%</td>
</tr>
<tr>
<td>rmsd bond length</td>
<td>0.010Å</td>
<td>0.006Å</td>
</tr>
<tr>
<td>rmsd bond angle</td>
<td>1.44°</td>
<td>1.22°</td>
</tr>
<tr>
<td>B_{average}</td>
<td>48</td>
<td>28</td>
</tr>
</tbody>
</table>

* The resolution range in the bracket is the highest resolution shell.

* number of reflections

\[ R_{sym} = \frac{|I_{hi} - I_h|}{I_{hi}}, \] where \(I_{hi}\) is the scaled intensity of the \(i\)th symmetry-related observation of reflection \(h\) \((h, k, l)\) and \(I_h\) is the mean value.

\[ R_{crys} = \frac{\sum_h |F_{oh} - F_{ch}|}{\sum_h F_{oh}}, \] where \(F_{oh}\) and \(F_{ch}\) are the observed and calculated structure factor amplitudes for reflection \(h\).

\(R_{free}\) is calculated the same way as \(R_{crys}\) with the test set.

rsmd = root mean standard deviation

\(B_{average}\) = average temperature factor

Table 2.2 Refinement Statistics
Fig. 2.6 Ramachandran plot of G12A/GDP structure generated by Procheck

| Residues in most favoured regions [A,B,1] | 268  | 59.3% |
| Residues in additional allowed regions [a,b,1,p] | 30   | 10.9% |
| Residues in generously allowed regions [-a,-b,-1,-p] | 0    | 0.0%  |
| Residues in disallowed regions | 2    | 0.7%  |
| Number of non-glycine and non-proline residues | 300  | 100.0% |
| Number of end-residues (excl. Gly and Pro) | 4    |       |
| Number of glycine residues (shown as triangles) | 22   |       |
| Number of proline residues | 6    |       |
| Total number of residues | 332  |       |

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% of the most favoured regions.
Fig. 2.7  Ramachandran plot of G13A/GDP structure (Procheck).
Chapter 3

ANALYSIS AND FUTURE DIRECTIONS

3.1 G12A H-Ras

Fig. 3.1 shows the 2Fo-Fc electron density map of G12A H-Ras before the model Gly12 was mutated to alanine. The map was calculated without the contribution of the alanine side chain to avoid model bias. The positive density of the methyl group of Ala12 can clearly be seen but is not in close contact with other parts of the protein or nucleotide. From the alignment of the backbone of the mutant with those of wild type H-Ras/GDP and H-Ras/GppNp structures (Fig. 3.3), it is clear that the two GDP forms show no significant structural differences over all. The GppNp form differs from the other two only in the Switch I and II regions, where the change is due to the binding of the triphosphate nucleotide. Loop 1, where the mutation is, appears to be rigid and unaltered. While the average B factor for the whole molecule is 48, it is only 27 for Loop 1, indicating that the density is well defined for Loop1.

Loop 1 of the mutant is modeled to fit the same region of the wild type H-Ras from the complex structure of p120GAP-GRD and H-Ras with the lsq command in O (Fig. 3.4). The methyl side chain of Ala12 gets within van der Waals distance of both the main chain of GRD Arg789 (2.17Å) and the side chain of H-Ras Gln61 (2.37Å), as proposed by Scheffzek et al. (1997). Therefore it appears that G12A can bind to GRD in the ground state just like the wild type H-Ras. However, the complex is not able to generate the conformational change that is necessary to position the catalytic residues in their exact transition state positions due to steric hindrance from the extra methyl group. This
hypothesis is supported by the relative rotation of Rho and rhoGAP going from the ground to transition state (Rittinger et al., 1997a and b; refer to section 1.2.3c, Fig. 1.11).

When Loop 1 from another mutant G12P H-Ras/GDP (Scheidig et al., 1994) is also overlaid on top of the wild type and G12A loops (Fig. 3.4), Cβ of Pro occupies almost the same position as Cβ of Ala. Hence, the modeling does not differentiate the two mutants. To pursue the problem further, one needs to compare the real complex structures of each mutant with a GRD, especially the transition state complex of G12P with a GRD, provided such a complex can be crystallized.

3.2 G13A H-Ras

Again the Ca alignment shows no significant change in Loop 1 of the mutant (Fig. 3.3). The electron density map (Fig. 3.2), contoured at 1.3σ is generated before mutating Gly13 of the model to alanine, again to avoid model bias. The density for the methyl group is clearly pointing towards the solvent region (left side), away from the nucleotide. Therefore, if Gly13 is replaced by Glu, Val, Thr or Ser, the Cβ of the side chain will still be in that direction, positioning the rest of the side chain away from the catalytic site. It appears that the side chain of this position does not interact with anything other than solvent molecules, even in the modeled complex with GRD (Fig. 3.5). So by itself, the free Ras structure does not explain the importance of Gly13. Maybe bulky side chains at the 13th position hinder the proper positioning of the Gly12 backbone in the complex structure. This can only be verified by comparing the ground state and transition state complexes of the mutant with a GRD.
3.3 Conclusion

In this study, high resolution structures of H-RasG12A/GDP and H-RasG13A/GDP were obtained. Unfortunately, the triphosphate-nucleotide bound forms of the mutants did not give diffracting-grade crystals. The free mutant structures were modeled into the transition state complex of H-Ras with p120\textsuperscript{GAP}-GRD in order to understand the effects of the mutations on the active site. However, the structural basis of G12P and G13A oncogenicity remains unclear. Modeling alone cannot answer the question.

As mentioned before and in the Appendix, the crystallization of the H-Ras –GAP complex is very difficult, so far there is only one example (Scheffzek et al., 1997) at the transition state. If both Ala12 and Ala13 truly hinder the proper positioning of the active site residues, then it is most likely impossible to crystallize the transition state complexes of the mutants. On the other hand, the ground state has too low a binding affinity to be stable under crystallization conditions. Therefore, methods other than crystallography may be necessary.

NMR is a good way to study dynamics. In principle, one can label H-Ras and optimize the complex formation condition in solution by titrating with unlabelled GRD, thus obtaining an H-Ras spectrum in the complex form. By comparison with the free H-Ras spectrum, conformation changes in H-Ras should be detected as peak shifts. Of course, one has to take into account peak-broadening due to complex formation. However, it will be difficult to detect any changes in Arg789\textsuperscript{GAP} due to H-Ras mutation because the GRD domain is still too big for NMR now.
Fig. 3.1   Electron density map (2Fo-Fc) of G12A/GDP calculated before the Gly12 in the model was mutated to Ala. One can clearly see the density for the Ala methyl group.
Fig. 3.2 2Fo-Fc map of G13A/GDP  
The map was calculated before Gly13 was mutated to Ala in the model. The methyl group of Ala13 is obviously present and is pointing away from the nucleotide into the solvent region. One can also see the octahedral coordination of the Mg2+ ion by Ser17, beta phosphate and 4 water molecules.
Fig. 3.3  Backbone superposition of mutants (green), WT/GppNp (red) and WT/GDP (blue) H-Ras molecules. One can see that there is no significant difference between the WT/GDP and the mutants. Both switch regions move drastically when GTP is bound. However, no change is detectable at Loop 1 where the mutation is.
Fig. 3.4  Modeling of G12A, G12P and p120GAP-GRD catalytic site

The short peptide of G12A/GDP and G12P/GDP (residues 10-15) are modeled onto the same region of the wild-type H-Ras in the complex structure of Ras-GAP (Scheffzek et al., 1997). There is no significant difference between the Cβ positions of the Ala12 and Pro12. And Cβ is within van der Waals distance from both the side chain of Gln61 of Ras and main-chain carbonyl of Arg789 of GAP.

All components from the complex crystal structure are colored by atom-type. AlF₃ is in blue, G12A peptide is magenta, and G12P peptide is green.
Fig 3.5  The short peptide from residues 10-15 of G13A/GDP is modeled into the Ras/GAP complex (Scheffzek et al., 1997). As indicated by the free G13A structure, the methyl group of Ala13 is not involved in GAP binding.

Brown:  p120GAP-GRD
Grey:  H-Ras residues 10-15 from the complex structure
Green:  G13A peptide
Red:  Mg2+
Blue:  AlF3
Orange:  water
GDP is colored by atom-type.
In parallel to my experiments on H-Ras mutants, I also worked on a number of side projects, which are described in this appendix.
I. Crystallization of the tubulin-binding domain and catalytic domain of neurofibromin

I.1 Background

Neurofibromin (NFI), the product of the neurofibromatosis gene, is a GTPase activating protein. Mutation of this gene results in variable phenotypes such as benign neurofibromas, flat pigmented lesions and malignant peripheral nerve sheath tumors (Xu et al., 1990).

There is 30% sequence identity between the GRDs (GAP-related domains) of p120\textsuperscript{GAP} and NF1. NF1 binds to H-ras with 300-fold higher affinity than p120\textsuperscript{GAP}. Its Ras activation is inhibited non-competitively by certain lipids, arachidonate, phosphatidate and phosphatidylinositol-4,5-bisphosphate, in micromolar ranges where p120\textsuperscript{GAP} is relatively insensitive (Golubic et al., 1991; Bollag et al., 1991).

In addition, NF1 associates with cytoplasmic microtubules, and the 80 residues involved in this interaction are N-terminal to the minimum catalytic domain of NF1 (Gregory et al., 1993).

The structure of the minimum catalytic domain of NF1 was solved by X-ray crystallography (Scheffzek et al., 1998). The overall fold is very similar to that of p120\textsuperscript{GAP}.

I.2 NF1-GRD purification and crystallization

The NF1-GRD construct in the lab includes not only the catalytic domain of the NF1, but also the 80 amino acids tubulin-binding domain at its N-terminus (amino acids
1095 to 1569). It was inserted into the pSE-420 expression vector (Pharmacia) by Dr. Young-Hwa Song. The protein was over-expressed in *E. coli* XL1-Blue cells (overnight induction with 1mM IPTG at 37°C) and purified using a Q-sepharose column followed by a phenyl-sepharose column (Fig. 1.1), both from Pharmacia.

**Q-sepharose:**

| Buffer A | 50mM Tris pH8.0, 1mM EDTA, 1mM DTT, 1mM PMSF |
| Buffer B | 50mM Tris pH8.0, 1mM EDTA, 1mM DTT, 1mM PMSF, 1M NaCl |
| Gradient | A linear gradient of 0-40% buffer B in 2 C, (column volumes) followed by 40-100% buffer B in 1 C |
| Flow rate | 2mL/min |

**Phenyl-sepharose:**

| Buffer A | 5mM Tris pH8.0, 1mM DTT, 1mM EDTA, 1mM PMSF |
| Buffer B | 100mM Tris pH8.0, 1mM DTT, 1mM EDTA |
| Gradient | A linear gradient from 100% to 35% buffer B in 1.2 C, followed by a stationary phase of 0.6 C, at 35% buffer B, then another linear gradient from 35% to 0% buffer B in 1 C |
| Flow rate | 2mL/min |

The protein was then concentrated to about 10mg/ml for crystallization trials using the hanging-drop technique. Small crystals (Fig. 1.2) were grown from 100mM glycine pH10, 1.6M (NH₄)₂SO₄ and 10mM Zn(Ac)₂. A frozen data set to 2.55Å was
collected at Brookhaven National Laboratory beam line X8C and processed with DENZO. It was 98% complete and overall $R_{\text{sym}}$ was 11%. The space group was $P4_1 22$ or $P4_3 22$ with unit cell axes of $215.5 \times 215.5 \times 107.2 \text{Å}^3$.

I.3 Molecular replacement

Attempts to solve the structure using molecular replacement using the p120$^{\text{GAP}}$ GRD structure (Scheffzek et al., 1996) as the search model failed. Not only there are six molecules in one asymmetric unit (based on Matthew's coefficient) without any 6-fold or 3-fold symmetry and the search model had only 80% of the mass of NF1-GRD, but more importantly, as we know now, the model Ca does not agree well with that of NF1-GRD.

I.4 Future work

Since heavy metal derivative soaks have failed due to the high salt and high pH crystallization conditions and molecular replacement did not work either (see above), production of a Se-Met protein for MAD phasing is the obvious next step. Identifying all Se atoms will be a challenging project due to the large number of Se atoms (16 in one molecule without counting the N-terminal Met residue) that have to be identified from anomalous diffraction data.
(a). Elution profile from Q-sepharose
Amax = 2.0

(b). SDS-PAGE gel of fractions from Q-sepharose column around the NF1-GRD peak. Each lane represents 10μL of a 6mL fraction.
(*: fractions of NF1-GRD pooled for the phenyl-sepharose column)

Fig. 1.1 Purification of NF1-GRD
(c). Elution profile of NF1-GRD from the phenyl-sepharose column

(d). SDS-PAGE of fractions of the NF1-GRD peak from the phenyl-sepharose column. Each lane represents 10μL of a 4mL fraction.

Fig. 1.1 Purification of NF1-GRD
(See the text for more detailed description.)
Fig. 1.2  NF1-GRD crystals

These crystals were grown using the hanging-drop technique. 10mg/mL protein was mixed in 1:1 ratio with the reservoir solution of 100mM glycine pH10.0, 1.6M ammonium sulfate and 10mM zinc acetate. The crystals reached max. size of 0.15mm*0.05mm*0.05mm in 2-3 weeks.
II.  HtpG

II.1  Experimental attempts

Since we have the NF1-GRD construct in the lab, and information obtained from the mutant H-ras structures themselves was not enough to explain their transforming abilities, it was natural to try to crystallize GRD and H-Ras in their complex forms. As mentioned in chapter 1, the minimum catalytic domain of NF1 has a 20-fold higher affinity for H-Ras than p120GAP (Ahmadian et al., 1996). With this advantage and the larger size of our protein (the tubulin-binding domain and GRD), we hoped to overcome the problem experienced with complex formation.

The two proteins were purified separately and combined in a 1:1 molar ratio. In an attempt to generate the transition-state of the complex which is more stable than the ground state complex (Scheffzek et al., 1997), this mixture with additional 2mM AlF₄⁻ (2mM AlF₃ with 20mM NaF) was used for hanging-drop crystallization trials. However, the two proteins crystallized separately and often both crystal forms in the same drop, never as a complex. Once nucleation of one of the free proteins occurred, the equilibrium pulled away from the complex side, hence the complex concentration never reached saturation.

To prevent the complex from falling apart, we attempted to construct a fusion protein. The C-terminus of NF1-GRD was linked to the N-terminus of H-ras by a 30-amino-acid peptide: (SGGGG)₆, where Ser was chosen for solubility and Gly for flexibility. The linker was expected to increase the effective local concentrations of the
two proteins and at the same time, to allow the two proteins enough freedom of motion to be able to form the functional complex.

II.2 Construct

A BspEI site was engineered into the linker DNA sequence, and two separate PCRs were performed; one with NF1-GRD and half of the linker at its 3' end and the other with H-ras and the second half of the linker at its 5' end. Then 3-piece ligation with the pTrc99 plasmid (Pharmacia) was followed by transformation into E. coli BL21(DE3) cells. All this work was done in collaboration with Dean Sas, a former summer student in the laboratory.

A protein of the expected size of 72kDa was produced after overnight induction at 37°C (Fig. II.1) and purified the same way as NF1-GRD. The protein eluted at about the same position as NF1-GRD (Fig. II.2) from both columns. Crystals were obtained from 100mM Tris pH8.0, 2M (NH₄)₃PO₄ (Fig. II.3).

II.3 Result

Due to the lack of antibodies and easy functional assays, the purified protein was sent for N-terminal protein sequencing to confirm its identity. The first 10 amino acids had a perfect match to E. coli heat shock protein HtpG, which is a homologue of eukaryotic hsp83 (Bardwell et al., 1987). Synthesis of heat shock proteins in E. coli is induced by heat, viral infection, ethanol, azide and other toxic compounds. Since the NF1-GRD DNA sequence has many arginine codons that are not utilized at any significant rate by E. coli, forced production of the protein put bacteria cells under great
stress, triggering the synthesis of heat shock proteins. Overall the amino acid sequences of HtpG and the human hsp83 show 42% identity (Bardwell et al., 1987).

The structure of the N-terminal ATP-binding domain of the yeast homologue hsp90 was solved (Prodromou et al., 1997). Unfortunately, the full length E. coli HtpG protein crystal only diffracted to about 8Å resolution, even after exhaustive refinement of the crystallization conditions.
Fig. II.1 Induction of NR30

Lane 1 MW marker (kDa)
Lane 4 Before induction
Lane 5 After induction

Lanes 2 & 3 are the whole cell lysate before and after induction of another colony. In this case, there is no induction of protein around 70kDa.
(a). Elution profile of NR30 from the Q-sepharose column

(b). SDS-PAGE of fractions from the Q-sepharose column (Lanes 4 and 5 are collected for the next step.)

Fig. II.2 Purification of NR30
(c). Elution profile of NR30 from the phenyl-sepharose column

(d). SDS-PAGE of fractions containing NR30 from the phenyl-sepharose column

Fig. II.2 Purification of NR30 (details in the text)
These crystals were about 0.5mm in diameter, and were grown at room temperature from hanging-drop with 100mM Tris pH8.0 and 2M ammonium phosphate as the reservoir solution.
III. X-Ras

III.1 Background

X-Ras is a newly discovered murine Ras molecule. It has 57% sequence identity to human H-ras (Fig. III.1). According to our collaborator John W. Schrader of the University of British Columbia, this protein interacts with a greatly overlapping range of exchange factors and effectors as H-Ras. When immobilized on a column, X-Ras pulls tubulin from whole cell lysate. It would be interesting to see whether there is any significant structural difference between H-Ras and X-Ras.

III.2 Protein purification and crystallization

X-Ras DNA was inserted into ptacRas, the same expression vector as H-Ras by our collaborator (John W. Schrader), and the protein was over-expressed in the E. coli strain BL21(DE3)pLysS overnight at 37°C with 1mM IPTG. A protein of 25KD was strongly expressed, which was around the calculated molecular weight for X-Ras (Fig. III.2). The cells were lysed using a French Press and the protein was purified using Q-sepharose (150mL) followed by phenyl sepharose (20mL) (both from Pharmacia) (Fig. III.3).

**Lysis buffer:** 64mM Tris pH7.5, 1mM DTT, 1mM PMSF, 1mM EDTA

**Q-sepharose:**

Buffer A 64mM Tris pH7.5, 1mM DTT, 10mM MgCl₂

Buffer B 64mM Tris pH7.5, 1mM DTT, 10mM MgCl₂, 1M NaCl
Gradient: A linear gradient from 0% to 100% Buffer B in 3C.

Flow rate: 2mL/min

Phenyl-sepharose:

Buffer A: 20mM Tris pH7.5, 1mM DTT
Buffer B: 50mM Tris pH7.5, 1mM DTT, 300mM NaCl
Gradient: A linear gradient from 100% to 0% Buffer B in 3C.
Flow rate: 0.8mL/min

The protein in 20mM Tris pH7.5 and 1mM DTT was concentrated to 13mg/ml for hanging-drop crystallization trials. And crystals were obtained from 100mM Citrate pH5.0, 5% MPD and 12% PEG3350 (Fig. III.4).

III.3 Results

The purified protein was sent to our collaborators for functional assays. Mass spectrum of the tryptic digest of the protein identified two internal peptide sequences (M. Quadroni, UBC, personal communications):

KITGYTTVDISQWHR
MLNELQQYCDEWQGGA

Both peptides clearly matched to *E. coli* chloramphenicol acetyltransferase type I (CAT I), whose crystal structure was solved (Crane et al., 1997).

CAT I is encoded by the pLysS plasmid in the expression *E. coli* strain for antibiotic resistance selection. This plasmid supplies a low level of T7 lysozyme, which selectively inhibits T7 RNA polymerase; thus, effectively eliminating leaky expression of
the cloned protein before induction. This feature is particularly useful for induction of toxic proteins. Even though the pLysS plasmid is compatible with pET plasmids, chloramphenicol resistance selection is needed to maintain it.

The over-expression of CAT I protein is not directly linked to addition of IPTG. Colonies have been found to express the protein both before and after adding IPTG, and this expression was not consistent. Apparently, some unknown variable that differed in each preparation induced the expression of CAT I much like a heat shock protein, although no chloramphenicol was used.
Fig. III.1  Protein sequence alignment of H-Ras with x-Ras

Red = identical residues
Blue = residues with conserved biochemical properties
Fig. III.2 "X-Ras" induction (SDS-PAGE)

Lane 1: MW marker
Lane 2: After induction
Lane 3: Before induction

(Overnight culture was used to inoculate a 3mL culture, which was induced with 1mM IPTG after 3.5 hours. After 6 hours of induction, sample buffer was mixed with some culture and run on a denaturing (SDS) gel.)
(a). Elution profile of "X-Ras" from the Q-sepharose column

(b). SDS-PAGE of Q-sepharose fractions

Fig. III.3 Purification of "X-Ras"
(c). Elution profile of "X-Ras" from the phenyl-sepharose column

(d). SDS-PAGE of fractions containing "X-Ras" form the phenyl-sepharose column

Fig. III.3  Purification of "X-Ras" (see text for more detail)
Fig. III.4  CAT1 crystals

Crystals were grown at room temperature from hanging drops with 100mM citrate pH5.0, 5% MPD and 12% PEG3350 as the reservoir solution. The crystals reach 0.3mm in each dimension after 3 weeks.
Appendix B

Correct Sequence of the expression vector ptacRas

<table>
<thead>
<tr>
<th></th>
<th>Length: 5262</th>
<th>July 22, 1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>The bolded region codes for the truncated form of H-Ras (amino acids 1-166).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>AATTAATTCA</td>
<td>TGCTGTGGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCGACGGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CACCAATGCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCTGGCGTCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGCAGCCATC</td>
</tr>
<tr>
<td>51</td>
<td>GGAAGCTGTG</td>
<td>GTATGGCTGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCAGGTGTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AATCACTGCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TAATTCGTGT</td>
</tr>
<tr>
<td>101</td>
<td>CGCTCAAGGC</td>
<td>GCACCTCCGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCTGGATAAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GTTTTTTCG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCGACATCAT</td>
</tr>
<tr>
<td>151</td>
<td>AACGGTTCTG</td>
<td>GCAAATATTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGAAATGAGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGTTGACAAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TAATCATCGG</td>
</tr>
<tr>
<td>201</td>
<td>CTGCTATAAT</td>
<td>GTGTGGAATT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GTGAGCGGAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACAAAATTTCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CACAGGAAAC</td>
</tr>
<tr>
<td>251</td>
<td>AGAATTCTAT</td>
<td>GACAGAATAC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AAGCTTGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TTGTGGCGGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CGGCGGTGTGT</td>
</tr>
<tr>
<td>301</td>
<td>GGAAGAGTGT</td>
<td>CGCTGACCAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCAAGCTGAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAGAAACATT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TTGTGGACGA</td>
</tr>
<tr>
<td>351</td>
<td>ATACGACCCC</td>
<td>ACTATAGAGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ATTCCTACCG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAAGCAGGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GTCATTGTG</td>
</tr>
<tr>
<td>401</td>
<td>GGGAGACGTG</td>
<td>CCTGTGGGAC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ATCCTGGATA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CGGCCGCCCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGAGAGTAC</td>
</tr>
<tr>
<td>451</td>
<td>AGCGCCATGC</td>
<td>GGGACCAGTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CATGCGCACC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGGAGGGGCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCCTGTGTGT</td>
</tr>
<tr>
<td>501</td>
<td>GTTTGCCATC</td>
<td>AAACACACCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AGTCTTTTGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGACATCCAC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAGTACAGGG</td>
</tr>
<tr>
<td>551</td>
<td>AGCAGATCAA</td>
<td>ACGGGTGAGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GACTCGGATG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACGTGCCCCAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGTGTGGGTG</td>
</tr>
<tr>
<td>601</td>
<td>GGAACAAGT</td>
<td>GTGACCTGGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGCAACGCAC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GTGGAATCTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGCAGGTCTA</td>
</tr>
<tr>
<td>651</td>
<td>GGCACCCTGG</td>
<td>CGAAGCTACG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCAATCCCCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CATCGAGACC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGCGCCAAGA</td>
</tr>
<tr>
<td>701</td>
<td>CCCGCCAGGG</td>
<td>AGTGAGGAGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCCTTTTACA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CGTTGGTGCG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TGAGATCCGG</td>
</tr>
<tr>
<td>751</td>
<td>CAGCACTAGC</td>
<td>TGCGGAAGCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAAACCTCCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GATGAGAGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCCCGGCTG</td>
</tr>
<tr>
<td>801</td>
<td>CATGAGCTGC</td>
<td>AAGTGTTGTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCTCCTGACG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAGCACAAGGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCAGGACATG</td>
</tr>
<tr>
<td>851</td>
<td>GAGGTGCCGG</td>
<td>ATGCAGGAGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAGGTGCAGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CGGAAGGAGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AGGAAGGAAG</td>
</tr>
</tbody>
</table>

85
GACGGAAGCA AGGAAGGAAG GAAGGGCTGC TGGAGCCCAG TCACCCCGGG

ACCGTGGGCC GAGGTGACTG CAGTCAGTCA TGTCCTTTGT GCCCGTCCCG

GCACAGGCTC AGGACGAGGA GGTGCCGGAT GCAGGGAGGT AGGTACTCCT

TAAGGGAAAA TCTAAGGCCCA GAGAGACAGC CGGTAACAGC TAGGCTGCCT

CTCTGCTCGC CTTTCTGTCT CATAGACAAA GCTTACTCCC CATCCCCCGG

AAAGATTTTT TTACTATAAA ACGCTGATGG AAGCGTTTAT GCGGAAGAGG

TAAAGCCCTT CCCCAGTAAC AAAAAAACA CAGCATAAAT AACCCCGCTC

TTACACATTC CAGCCCTGAA AAAAGGCATC AAATTTACC ACACCTATGG

TGCCTTGTCTT TACACATTC AACAAAATCG CAATGCTTGG CAGCATAAAT

AAAGATTTTT TTAATGCTTC TTACACATTTC TATGGTTGGT ATATGGGAA

GACGGAAGCA AGGAAGGAAG GAAGGGCTGC TGGAGCCCAG TCACCCCGGG
2101  GCTTCGGCGG  CGTGGACATG  GTGATCAGCG  GCAATGTGCC  GACGGGTGCC  
2151  GGGTTAAGTT  CTTCCGCTTC  ACTGGAAGTC  GCGGTCGGAA  CCGTATTGCA  
2201  GCAGCTTTAT  CATCTGCCGC  TGGACGGCGC  ACAATCGCG  CTTTAACGGTC  
2251  AGGAAGCAGA  AAAACCAGTTT  GTAGGCTGTA  ACTGCAGGAT  CATGATCAG  
2301  CTAATTTCCG  CGCTCGGCAA  GAAAGATCAT  GCCTTGCTGA  TGCAATGCGG  
2351  CTCACTGGGG  ACCAAAGCGT  TTTCCATGCC  CAAAGTGCTG  GCTGTCGTCA  
2401  TCATCAACAG  TAACTTCAGA  CGTACCCCTGG  TTGGCAGCGA  ATACAACACC  
2451  CGTCGTGAAC  AGTGCAGAAG  CGGTGCAGGCT  TTCTCTCAGC  AGCCACCCCT  
2501  GCGTGATGTC  ACCATTTGAAG  AGTTCCAGCG  TGTTGTCGAT  GAAGTGGACC  
2551  CGATCGTGGGC  AAAACCCGCTG  CGTCATATAC  TGACTGAAAG  CGCCCCAGCC  
2601  GTTGAAGCTG  CCAGCGCGCT  GGAGCAAGTG  GACCTGAAAC  GTATGGGCCA  
2651  GTTGATGGCG  GAGTCTCATG  CCTCTATGCG  CGATGATTTC  GAAATACCGG  
2701  TGCCGCAAAT  TGACACTCTG  GTAGAATCTG  TCAAAGCTGT  GATTTGGCGA  
2751  AAAGTTGCGC  TACGCATGAC  CGGCGGCGGA  TTTGGCGGCT  GTATCGTCGC  
2801  GCTGATCCCG  GAAGAGCTGG  TGCCTGCCGC  ACAGCAAGCT  GTGCGCTGAA  
2851  AATATGAAGC  AAAACCATGT  ATTAAAGAGA  CTTTTTACGT  TGTGAAACCA  
2901  TCACAAGGAG  CAGACAGTG  CGTAACGAAA  CTCCCCGCACT  GCCACCGGAT  
2951  GGCAGCGGTA  CCCATCTTGC  TGCCTGCAGC  GTTTCGGGTA  TGACGGTGAA  
3001  AACCTCTGAC  ACATGCAGCT  CCCGGAGAGC  GTCACAGCCT  GTCTGTAAGC  
3051  GGATGCAGCGG  AGCAGACAGG  CCGTCAGGGG  CGCGTCAAGG  GGTTGTTGCC  
3101  GGTGTCGGGG  CGCAGCCATG  ACCCAGTCCC  GTAGAGATAG  CGGAGTGAT  
3151  ACTGCGTAAA  CTATCGGGCA  TCAGACGACA  TTGACTGAG  AGTGCAACAT  
3201  ATGCCGTTGTA  AAATACCCGA  CAGATCGCTA  AGGAGAAAAAT  ACCGCACTCAG  
3251  GCGCTCTTCC  GCTTCCCTGGC  TCACCTGACTC  GCTGCGCTCG  GTGCTTCGGC  

87
3301  TGGCGGAGGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA
3351  GAATCAGGGG ATACACGCAG AAAAAGACATG TGACAAAAAG GCCAGCAAAA
3401  GGCGAGGAAC CGTAAAAAGG CCGCTGGCTC GGCCTTTTTC CATAGCTCCCT
3451  GGGGGCTCAG CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA
3501  AACCCGACAG GACTATAAAG ATACCAAGGCG TTTCCCCCTG GAAGCTCCTT
3551  CTTGCGCTCT CCTGTTCGGA CCCGGCCTGCT TACCGGATAC CTGTCCGCCCT
3601  TTCTCCCTTC GGGAGAGTGT GGCCTTTTCT ATAGCTCACG CTGTAGGTAT
3651  CTCAGTGTCG TGTAGGCTGT TCGTCTCAAG GCTGGGCTTG TGCACGAAACC
3701  CCCCCGAGCG CCGACCCGCT GGCCTTTATC CGGTAACTAT CGTCTTGAGT
3751  CCAACCCGTT AAGACACGAC TTACCCGACG TCACAGCAGC TACCTGGTAAC
3801  AGGATTAGCA GAGCGAGGTA TGAGGCTGCT GCTAGCAGAT TCTGAAAGTG
3851  GTGGCCTAAC TACGCTACA CTGAGAAGGAC ATGATTTGGT ACCTGCGCTC
3901  TGCTGAAGCC AGTTACCTTC GGGAAAAAGG TGTGATGCTG TTGATCCGGC
3951  AAACAGGAAC CCGCTGGTAG AGGTTGTTTT TTTGTGGCAG AGCAGCAGAT
4001  TACGCGCAGA AAAAAGGAT CTCAAGAAGA TCCTTTGATC TTTTCTACGG
4051  GGTCTGACGC TCAGTGGAAC CAAAACGCTC GTAAAAGGGT ATGGGTCTATG
4101  AGATTACAA AAAGGATCTT CACCTAGATC TTTTTAATT AAAAATGAA
4151  TTTAAAATCA ATCTAAGTA TATATGAGTA AACTTGCTT GACAGTTACC
4201  AATGCTTAAT CAGTGAGGCA CCAACCATCT CGATCTCTG TTTCTGTTCA
4251  TCCATAGGTT CTCGACTGCC GTGCCTGTAG ATAACTACGA TACGGGAGGG
4301  GGGGCATCTT CTTDATCAGT GGCCTTTATG GAATGCGGAC CCACGCCGCC
4351  CGGCTCCAGA TTTACGCGA ATAAACCAGC CAGCCGAAAG GCCGAGCCGC
4401  AGAAGTGGTC CTGCAACTTT ATCCGGGCTCC ATCCAGTCTA TTAATGGTTG

88
4451  CC GG GAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG
4501  TTGC CCTT GC TACAGGCATC GTGGGTGCAC GCTCGTGCTT TGGTATGGCT
4551  TCA TT CAGCT CCGGTTCCCA ACGTCAAGG CGAGTTACAT GATCCCCCAT
4601  GTTG TGCA AA AA GC GG GTTA GCTC CTTCCGGG TCC TCGGATC GTGGT CAGAA
4651  GTA AGTTGGC CGCAGTGT TA TCACTCATGG TTATGGCAGC ACTGCATAAT
4701  TCTCTTACTG TCATGC CATC GTGAAGATGC TTTTCTGTGA CTGGTGAGTA
4751  CTCAACCAAG TCATTC TGG AATAGTGTAT GCGCGGACCG AGTTGCTCTT
4801  GCCC GCGTTC AACACGGGAT AATACCGGC CACATAGCAG AACTTTAAAA
4851  GTGCTCATCA TTGGAAAACG TTCTTCGGGG CAAAACCTCT CAAGGATCTT
4901  ACCGCTTGTTG AGATCC AGTTG CCGTCAACC CACTCGT GCA CCCAACCTGAT
4951  CTTCAGC CATC TTTTACTTTC ACCAGCGTTT CTGGTGAGC AAAAAACAGGA
5001  AGGCAAAAATG CCGCAAAAAA GGGAATAAGG GCGACACGGA AATGTTGAAT
5051  ACTCATACTC TTTCTTTTTT AATATTATTG AAGCATTTAT CAGGGTTATT
5101  GTCTCATGAG CGGATACATA TTGAATGTA TTTAGAAAAA TAAACAAATA
5151  GGGT TCCGC GCACATT TCC CCGAAAGTG CCACCTGACG TCTAAGAAAC
5201  CATTATTATC ATGACATTAA CCTATAAAAA TAGCCGTATC ACGAGGCCCT
5251  TTCGTC TTTC AAG


