INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9405510

Structure, expression and function of chicken proto-oncogene c-*ros*

Chen, Jianmin, Ph.D.

City University of New York, 1993

Copyright ©1993 by Chen, Jianmin. All rights reserved.

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Structure, Expression and Function of Chicken Proto-oncogene c-*ros*

by

Jianmin Chen

A dissertation submitted to the Graduate Faculty in Biomedical Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, the City University of New York

1993

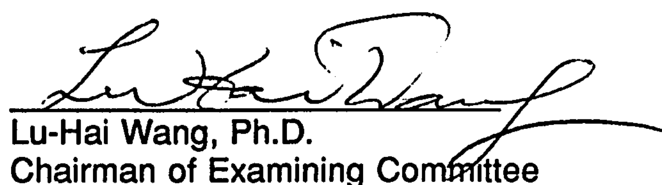
@ 1993

Jianmin Chen

All Rights Reserved

This manuscript has been read and accepted by the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 25, 1993
Date



Lu-Hai Wang, Ph.D.
Chairman of Examining Committee

July 28, 1993
Date

Terry A. Krulwich, Ph.D.
Exacutive Officer

Examining Committee

Irwin Gelman, Ph.D.

Hidesaburo Hanafusa, Ph.D.

Peter Palese, Ph.D.

Jay C. Unleless, Ph.D.

The City University of New York

Abstract

Structure, expression and function of the chicken proto-oncogene c-*ros*

by

Jianmin Chen

Advisor: Lu-Hai Wang, Ph.D., Professor

Proto-oncogene c-*ros* is the cellular counterpart of the transforming gene v-*ros* of avian sarcoma virus UR2 (Neckameyer et al., 1986; Chen et al., 1991). The transforming protein of UR2 is a *gag-ros* fusion polypeptide of 68-kDa with protein tyrosine kinase (PTK) activity (Feldman et al., 1982; Jong & Wang, 1987). Previous studies have shown that c-*ros* codes for a receptor-like PTK molecule (Neckameyer et al., 1986; Matsushime et al., 1986). However, the physiological function and potential ligand of the c-*ros* product remains unknown. In order to understand the function of the c-*ros* proto-oncogene, following aspects of c-*ros* were explored in my PH.D work:

1. Isolation and cloning of the c-*ros* cDNA. Using the v-*ros* DNA as a probe to screen chicken kidney cDNA libraries, I have isolated several overlapping cDNA clones and determined their sequences (Chen et al., 1991). Nucleotide sequence of the 8.1-kb c-*ros* cDNA shows that it codes for a transmembrane (TM) PTK molecule of 2311 amino acids (aa).

2. Functional characterization of the c-*ros* protein. Expression study showed that c-*ros* product is a 260 to 280-kDa glycosylated protein with very low kinase activity in the absense of ligand stimulation. The biochemical and biological properties of the c-*ros* protein and one of its activated variant were analyzed.

3. Analysis of the c-*ros* expression in vivo. Analysis of the c-*ros* RNA expression in various chicken tissues by RNase protection assay (RPA) and *in situ* hybridization showed

that the c-*ros* expression is under tight temporal and spatial regulation. The tissue and cell type specific expression of c-*ros* suggests that it may play important roles in their development and mature functions.

4. Clonning and charaterization of the c-*ros* promoter. A genomic DNA fragment corresponding to the potential c-*ros* promoter was isolated. Experiments of primer extension, RNase protection and functional demonstrations (CAT assay) were performed to verify that this region is indeed the c-*ros* promoter. One positive regulatory region was mapped by analysis of a series of deletion mutant.

Acknowledgements

First of all, I would like to give my sincere thanks to Dr. Lu-Hai Wang, not only for being a mentor, but also a friend. I am in debt to him for his stimulating discussions, helpful advises and continuous encouragement throughout my graduate study.

I would also like to express my special thanks to Dr. Unkeless, with whom I did my first rotation with. I will cherish all his intellegence, humor and friendship. Meanwhile, I am very grateful to Dr. Hanafusa, Dr. Palese, Dr. Gelman, Dr. Sickevitz and Dr. Bieker for serving on the committee, reading the manuscript and providing me with many helpful suggestions at various stages of my graduate study.

As a foreign student, I would like to give my special thanks to the graduate dean, Dr. Krulwich, for her parent-like guidance and care from the beginning of my study in the United States.

Countless thanks from deep in my heart goes to my wife, Xiaoqiao. Without her, I would not have been able to finish my graduate work so smoothly. I am deeply in debt to her understanding and companionship through so many weekends and midnights in the lab. I would also like to mention my 1-year old daughter, Rebecca, whose coming into the world gives me fulfilling happiness and enjoyment beyond the joy of science.

I am also very grateful to my colleagues, Delong, Susan, Song-Muh, Thambi, Joe, Betty, Diana, Lily, Yixing, Jing, Qinghua, Jie and Jennifer for their helpful discussion and sharing of reagents. I am in debt to Dr. Shudong Zong and Wei Chen for their help in *in situ* histochemistry, and Joseph Chan for some of the art work.

Lastly, I would like to dedicate this thesis to my parents.

Format of Thesis

This thesis was prepared in accordance with the guidelines of the City University of New York. Chapter III contains results published as Chen et al, Oncogene (1991) 6:257-264. Part of the results in chapter IV was submitted to Journal of Virology. Chapter V contains results submitted to Oncogene. Chapter III, IV, V and VI has an introduction and discussion specific for that section, with a general introduction and a general discussion at the beginning and end of the thesis. To avoid redundancy, "Materials and Methods" and "References" have been consolidated.

Table of Contents

Abstract....................IV

Acknowledgement....................VI

Format of Thesis VII

Table of Contents VIII

List of Tables....................IX

List of Figures....................X

Chapter I. Introduction....................1

Chapter II. Materials and Methods....................17

Chapter III. The proto-oncogene c-*ros* codes for a transmembrane tyrosine protein kinase sharing sequence and structural homology with *sevenless* protein of *Drosophila melanogaster*....................31

Chapter IV. Charaterization of the c-*ros* protein product....................50

Chapter V. Tissue and epithelial cell-specific expression of the chicken proto-oncogene c-*ros* in several organs suggests that it may play important roles in their development and mature function....................87

Chapter VI. Cloning and charaterization of the chicken proto-oncogene c-*ros* promoter....................111

Chapter VII. General discussion....................127

Bibliography....................132

List of tables

Table 5-1. Expression profile of chicken c-*ros* in different issues................................101

List of Figures

Fig. 3-1. Expression of c-*ros* in chicken tissues....42

Fig. 3-2. Isolation of c-*ros* cDNA clones....43

Fig. 3-3. Nucleotide and derived amino acid sequences of the c-*ros* cDNA....44

Fig. 3-4. Hydrophilicity plotting of the predicted c-*ros* protein....47

Fig. 3-5. Comparison of the predicted amino acid sequences of chicken c-*ros* cDNA with those of rat c-*ros* and *sevenless* cDNAs....48

Fig. 3-6. Comparison of the insertion sequences and autophosphorylation sites....49

Fig. 4-1. Constructs of c-*ros*, v-ros and VCros in mammalian expression vector and retroviral vectors....67

Fig. 4-2. Western blotting of pCMVros transfected COS cell clones selected by G418....69

Fig. 4-3. Expression of c-*ros* protein in CMVros6 cells....70

Fig. 4-4. Indirect immunostaining of COS-7 cells transiently expressing c-Ros, v-Ros and ppRos....71

Fig. 4-5. Membrane association, kinetics of biosynthesis and glycosylation of c-Ros and ppRos....73

Fig. 4-6. Subcellular fractionation of COS-7 cells transfected with v-ros, c-ros or ppros plasmid....74

Fig. 4-7. Surface labeling of the Ros proteins....75

Fig. 4-8. Covalent dimerization state of the c-Ros, ppRos and v-Ros....76

Fig. 4-9. Effect of the triple tyrosine residues on the kinase activity of c-Ros....77

Fig. 4-10. kinase activity of ppRos in vitro and *in vivo*....78

Fig. 4-11. Expression of c-Ros, ppRos and v-Ros in CEF....78

Fig. 4-12. Cell morphology and colony formation assay of CEF expressing v-Ros, c-Ros and ppRos........80

Fig. 4-13. Comparison of phosphotyrosine-containing proteins in UR2 and VCros transformed CEF and CEF transiently expressing ppRos........81

Fig. 4-14. The PI-3-Kinase activity associated with v-Ros and ppRos........82

Fig. 4-15. Phosphorylation of Shc was increased in UR2 and VCros transformed CEF but not in CEF expressing ppRos........83

Fig. 4-16. Phosphorylation states of PLCγ1 and a 88-kDa protein........84

Fig. 4-17. Construct of pCMVrosAP and pCDMrosFc........85

Fig. 4-18. Expression of RosAP and RosFc........86

Fig. 5-1 . Different c-*ros* probes used for RNA analysis........102

Fig. 5-2. RNase protection analysis of RNA samples from different chicken tissues........103

Fig. 5-3. RPA of RNA samples from chicken testis........104

Fig. 5-4. Northern blot Analysis of c-*ros* mRNA in different chicken tissues........105

Fig. 5-5. Kinetic study of c-*ros* expression in kidney and intestine........106

Fig. 5-6. Kinetic study of c-*ros* expression in lung, bursa and thymus........107

Fig. 5-7. Expression of c-*ros* in young and adult chickens........108

Fig. 5-8. c-*ros* mRNA *in situ* hybridization in chicken kidney, intestine and bursa......109

Fig. 6-1. Organization of the putative c-*ros* promoter region........122

Fig. 6-2. Nucleotide sequence of the putative c-*ros* promoter region........123

Fig. 6-3. Primer extension and RNase Protection assay for mapping the initiation site of the c-*ros* transcript........125

Fig. 6-4. The putative c-*ros* promoter is functionally active........126

Chapter I. Introduction

I. Oncogene and tumour suppressor gene

II. Protein tyrosine kinases (PTKs) and their signal transduction

III. Viral oncogene v-*ros* and proto-oncogene c-*ros*

I. Oncogene and tumor suppressor gene

Oncogene and proto-oncogene

Oncogenes described to date encode proteins that fall into four classes: growth factors, growth factor receptors, transducers of growth factor responses and transcription factors that mediate growth factor induced gene expression (Cantley et al., 1991). The normal cellular counterpart of oncogenes are proto-oncogenes, whose products are mainly involved in cellular signaling pathway (Hunter, 1991), controlling cell growth and differentiation. Genetic alteration of a proto-oncogene can covert it to an oncogene and cause uncontrolled cell growth or termination of differentiation.

Oncogenes can be dominant or recessive. Dominant mutation will cause gain of function (eg. *ras* gene) , whereas recessive mutation will lead to loss of function (eg. Rb gene). The latter are now widely recognized as tumor suppressor genes. The mechanisms of tumorigenesis include the following possibilities: point mutations usually activate proto-oncogenes; deletions are mainly involved in tumor suppressor gene inactivation; chromosomal translocation could activate oncogenes or inactivate tumor suppressor genes; and gene amplification can cause oncogenic activation (Bishop, 1991).

The first class of oncogenes is growth factor or growth factor-like molecules. *sis* (ie. PDGF B chain) is the first member of this family identified. It is widely accepted that oncogenic activation by growth factor is mainly through autocrine stimulation. The latest

members of this class include *int*-1, *int*-2, *hst* and FGF-5.

The second class of oncogenes encode mutant forms of growth factor receptors, many of which are receptor PTKs. Details will be discussed in the second part of this chapter.

The third class encode a large number of distinct molecules which are involved in different aspects of intracellular signalling. Among them, *src*, *fps* and *abl* represent nonreceptor PTKs; H-*ras*, K-*ras* and N-*ras* represent membrane associated G proteins; *mos* and *raf* represent cytoplasmic protein serine/threinine kinases; *Crk*, *Nck*, *Shc* and GRB2/*Sem5*/*Ash* represent cytoplsmic adaptors or regulators. Specific mutation, deletion or fusion with other protein may activate the oncogenic potential of these signalling molecules.

The final class is nuclear transcription factors. Among them, *fos*, *jun*, *myc*, *myb* and *rel* are well characterized. *jun*, *fos* and *myb* are activated because of the mutations that result in the loss of their negative regulatory elements. In contrast, for v-*rel*, the loss of the positive effector domain leads to the dominant negative prevention of expression of genes required for differentiation of cells. In many examples of human lymphomas, constitutive expression of c-myc leads to transformation.

Tumor suppressor gene and human cancer

Evidence for tumor suppressor genes comes from three sources: cell hybrids, familial cancer and loss of heterozygosity (Marshal, 1991). Cell hybrid experiments showed that normal cells could suppress the tumorigenic ability of certain tumor cells, suggesting that loss of certain normal gene functions could cause tumor (Harris, 1988). Recently, suppression of tumorigenicity was demonstrated by introducing a normal chromosome into some tumor cells (Sugawara et al., 1990). The identification of tumor

suppressor genes through studies of familial cancers such as retinoblastoma (Friend et al., 1986; Lee et al., 1987)) or loss of heterozygosity in colon cancers and colorectal carcinomas (Fearon and Vogelstein, 1990) supported the idea that inactivation of tumor suppressor genes is a major causal event in development of certain human cancers. Some tumor suppressor genes and their possible normal function will be discussed in the next section.

The Rb gene was recently shown to be responsible for familial and sporadic retinoblastoma (Friend et al., 1986; Lee et al., 1987). Inactivation of Rb gene can also be found in small cell carcinoma of the lung, bladder cancer and breast carcinoma (Marshall, 1991). It was shown that unphosphorylated Rb protein p105 prevented cell cycle progression through G1 and this activity was regulated by phosophorylation. There is new evidence suggesting that Rb could act as a major switch in cell cycle regulation (Wen-Hua Lee's unpublished results). Very recently, Rb was also shown to be present in the transcriptional complex of E2F (Bagchi et al., 1991; Chellappan et al., 1991). Homozygous Rb- mice died at day 15 of embryonic stage, resulting from abnormalities in nervous and hemotopoeitic systems (Lee et al., 1992; Jacks et al., 1992; Clarke et al., 1992). The appearance of pitutary gland tumors in heterologous Rb+/Rb- mice reflects different Rb function in different species. In fact, no retinoblatoma developed in these mice.

Loss of heterozygosity of p53 is frequently associated with astrocytomas, breast cancer, small cell lung cancer and colon cancer (Marshall, 1991). Inactivation of p53 could occur by point mutation or deletion, or dominant negative mutations that inhibit the normal p53 function (Eliyahu et al., 1988). There was evidence suggesting that p53 could function as a transcriptional activator (Fields and Jang, 1990; Raycroft et al., 1990). p53 was also

shown to bind some specific DNA sequences (Bargonetti et al., 1991; Kern et al., 1991; El-Deiry et al., 1992) and activate transcription in vitro (Farmer et al., 1992). Very recently, p53 was suggested to play a role at a certain cell cycle check point (Kastan et al., 1992). Homozygous p53-deficient mice are normal in their early life, but tend to develop a variety of neoplasms by 6 months of age (Donehower et al., 1992).

The functions of most of the other tumor suppressor genes identified so far remain unknown. NF1 for neurofibromatosis type 1 (Xu et al., 1990a) was recently shown to have GTPase activation activity (Xu et al., 1990b). WT1 for Wilm's tumor (Pridchard-Jones et al., 1990), a tissue-specific transcription factor containing four zinc fingers, was recently shown to be able to regulate expression of human IGFII (Drummond et al., 1992). DCC (deleted in colerectal carcinoma) (Fearon et al., 1988) encodes a protein sharing homology to neural cell adhesion molecules. DP2.5 was identified to be the gene responsible for familiar adenomatous polyposis coli (FAP) (Joslyn et al., 1991), but there was no clue to what its function might be from its primary sequence. NF2 for neurofibromatosis was identified to be a novel mesin-, ezrin-, radixin-like gene whose product was proposed to link cytoskeletal components with proteins in the cell membrane (Trofatter et al., 1993). The tumor suppressor gene for the von Hippel-Lindau disease was just cloned (Latif et al., 1993) and shows no homology to other proteins.

II. Receptor protein tyrosine kinases and their signal transduction

Protein tyrosine kinases (PTKs) represent the products of the largest family of the 60 or so oncogenes and proto-oncogenes known to date. They can be further divided into receptor and non-receptor PTKs (Hanks et al., 1988; Storm & Bose, 1989). These two PTKs transmit extracellular stimuli such as from growth factors intracellularly via a

cascade of reaction leading to changes in cellular metabolic activity and growth rate. Change in their normal functions due to over-expression or structural alterations often leads to uncontrolled cell growth, transformation and oncogenicity (Bishop, 1991; Cantley et al., 1991; Cross & Dexter, 1991; Hunter, 1991). Understanding pathways of signal transduction mediated by those PTKs is essential in elucidating the mechanisms of cell growth and oncogenic transformation.

Receptor Protein Tyrosine Kinases

A typical receptor PTK (RPTK) consists of an extracellular (EC) ligand binding domain, a transmembrane (TM) domain and cytoplasmic domain containing catalytic and substrate interacting regions. Binding of the ligand triggers structural alteration of these molecules such as oligomerization (Yarden & Ullrich, 1988; Ullrich & Schlessinger, 1990), resulting in autophosphorylation, activation of kinase activity, substrate association and finally passage of the signal to subsequent messengers most of which are yet to be identified. Oncogenic variants of these receptor PTKs include deletion of large portions of the EC domain (v-*erb*B/EGFR, v-*fms*/CSF-1 receptor, v-*ros*/c-*ros*, v-*kit*/c-*kit*, *trk*), single amino acid substitutions in the TM domain (*neu*), and mutations and truncations in the cytoplasmic domain (*erb*B/EGFR, v-*ros*, v-*fms*), all of which result in constitutive PTK activation in the absence of ligand stimulation. The receptor PTK oncogenes or proto-oncogenes for which their normal ligands are known include v-*erb*B/epidermal growth factor receptor (EGFR) (Downward et al., 1984; Ullrich et al.,1984), v-*fms*/monocyte colony stimulating factor-1 receptor (CSF1-R) (Sherr et al., 1985), v-*kit*/master cell growth factor receptor (Flanagan et al., 1990; Zsebo et al., 1990; Huang et al., 1990), *trk*/nerve growth factor receptor (NGFR) (Klein et al., 1991; Kaplan et al., 1991), *trk*B/Neurotrophin-3 (NT-3) and brain derived growth factor (BDNF) receptor (Klein et al., 1991; Glass et al., 1991;

Sqainto et al., 1991; Soppet et al., 1991; Hempstead et al., 1991), *trk*C/NF3 receptor (Lamballe et al., 1991), platelet-derived growth factor receptor (PDGFR) (Doolittle et al., 1983; Waterfield et al., 1983); *met*/Hepatocyte growth factor receptor (Naldini et al., 1991), insulin and insulin-like growth factor I receptor (IR and IGFR) (Wang et al., 1987; Kaleko et al., 1990), *erb*B2 or HER2/Neu diffentiation factor receptor or heregulin (Peles et al., 1992; Wen et al., 1992; Holmes et al.,1992), and Flk-1 (fetal liver kinase)/ VEGF receptor (Vascular endothelial growth factor receptor) (Millauer et al., 1993). However, ligands for other receptor-like PTKs including *ros*, *sea* and *ret* (Reddy et al., 1988; Cooper, 1990) remain unknown.

Many RPTKs function in growth control and the maintenance of normal homeostasis (Yarden & Ullrich, 1988). However, more and more RPTKs have been shown to be important for the development and differentiation. In *Drosophila*, DER(EGFR) has been proven to be important in early embryogenesis and eye devlopment; *sevenless* is critical for the differentiation of the photoreceptor R7 and *torso*, is important for establishment of terminal structures of the early embryo (Schejter & Shilo, 1989; Price et al., 1989; Basler et al., 1991; Sprenger et al., 1989). In C. elegans, *let*-23 (EGF receptor homolog) was shown to be the receptor for the inductive signal required for the vulval formation (Aroian et al., 1990). In the mouse, the c-*kit* (W locus) and steel (ligand for c-*kit*) are required for the proliferation and migration of various stem cells (Anderson et al., 1990; Martin et al., 1990). Hepatocyte growth factor/scatter factor and its receptor Met was shown to be able to control the differentiation of epithelial cells into branching tubules (Montesano et al., 1991). Trk, upon NGF stimulation, could mediate the formation of neurites from sympathetic neurons (Kaplan et al., 1991; Klein et al., 1991). Recent studies suggest that VEGF, ligand for the Flk-1, is a major regulator of vasculogenesis

and angiogenesis (Millauer et al., 1993).

Signal transduction for RPTKs

After years of intensive search of substrates for normal and oncogenic PTKs, a scheme of signal transduction for RPTKs has gradually come to light. The initial activation of kinase activity and autophosphorylation following ligand stimulation, appears to enable the RPTKs to associate with and/or phosphorylate their substrates.

One of the central components in the earlier events of signalling process is phosphoinositol-3-kinase (PI-3-Kinase)(Cantley et al., 1991; Berridge et al., 1984; Berridge & Irvine et al., 1989; Majerus et al., 1990). Other important components include phospholipase C-γ1 (PLCγ1), GTPase activating protein (GAP) and c-*raf* product p74 (Adari et al., 1988; William et al., 1989; Margolis et al., 1989; Meissenkelder et al., 1989; Morrison et al., 1988; Ellis et al., 1990; Kaplan et al., 1990; Kypta et al.,1990). Recently, c-*src* was placed downstream of siganl transduction of PDGFR (Kypta et al., 1990). The most recent finding suggested that PI-3-K and PLCγ1 are the downstream mediators of mitogenic signal of PDGF receptor (Valius and Kazlauskas, 1993).

From the study of the oncogene *crk*, coding for a PTK negative and SH2/SH3 sequence-containing protein (Koch et al., 1991; Mayer & Hanafusa, 1990; Anderson et al., 1990) as well as cloning and sequencing of other signal transducing molecules including Shc, Nck, Grb2/*Sem-5*/*Ash*/*Drk* (Koch et al., 1991), it becomes evident that SH2 sequences are involved in the interaction between the RPTKs and their substrates. Meanwhile, tyrosine phosphorylation of the receptor and/or the substrate molecules is important for their association (Koch et al., 1991). Not surprisingly, each receptor PTK has its own specific interacting components. For example, while activated PDGFR binds to all four components mentioned above, others such as IR, EGFR and CSF1-R bind only to

a subset of the four (Cantley et al., 1991). How those signalling proteins transmit the message is not completely clear.

The capability of the native and oncogenic PTKs to associate with PI-3-Kinase seems to correlate well with their normal and transforming functions, respectively (Cantley et al., 1991). PDGFR, EGFR, c-Kit and CSF1R were found to associate with PI-3-kinase. $P68^{gag\text{-}ros}$ and its closely related RPTKs, insulin receptor and IGFR, have been shown to be associated with PI-3-Kinase as well (Fukui et al 1989; Endemann et al., 1990; Ruderman et al., 1990; Liu et al., 1992). The cDNAs encoding the PI-3-Kinase subunits p85 and p110 were recently cloned (Escabedo et al., 1991; Otsu et al., 1991; Skolnik et al., 1991; Hiles et al., 1992). The regulatory subunit p85 has a structure of SH2-SH3-SH2, which can directly associate with activated receptors through the SH2-pTyr interaction. Some receptors such as IR associate with PI-3-Kinase through another adaptor molecule called IRS-1 which is a major substrate of activated IR (Sun et al., 1991). The p110 catalytic subunit shares great homology with Vps34p, a yeast protein involved in the sorting proteins to the vacuole, and recently, the Vps34p was shown to contain PI-3-Kinase activity, suggesting that PI-3-Kinase may play a role in intracellular protein sorting (Schu et al., 1993).

The finding that GAP promotes the GTPase activity of *ras* and inactivates *ras*-GTP complex provides a link between the PTK and G protein signalling pathways (Cantley et al., 1991). The GAP has two SH2 domains and one SH3 domain and the SH2 domains have been shown to be able to mediate the association of the GAP to the activated PDGFR and EGFR (Fantl et al., 1992). Growth factors like EGF and PDGF can induce the association of GAP with two phosphorylated cellular proteins, p62 and p190 (Ellis et al., 1990; Moran et al., 1990). p62 is a putative nucleic acid binding protein (Wong et al.,

1992) and p190 which shares homology with other GAP-like proteins including Bcr and n-chimaerin may link *ras*GAP to the nucleus (Broach et al., 1991; Settleman et al., 1992; Hall et al , 1992). Whether GAP is a upstream or downstream component in the *ras* pathway is not clear because recently there has been evidence suggesting that GAP could be downstream component of the *ras* signaling pathway, ie. the SH2 and SH3 domain of GAP are responsible for *ras*-dependent inhibition of muscarinic atrial potassium channel currents (Martin et al.,1992). The most recent reports placed the adaptor molecule Grb2 in the *ras* signalling pathway (Lowenstein et al., 1992). Grb2 binds to the activated receptor either directly through its own SH2 domain or through another SH2-containing adaptor molecule called Shc which was stronly phosphorylated upon the receptor activation. The Grb2 in turn will bind to the Sos protein which is the activator of Ras that promotes the GTP/GDP exchange. It was shown that the C-terminal proline-rich domain in the Sos protein binds to the SH3 domain of the Grb2 molecule (Egan et al., 1993; Rozaki-Adcock et al., 1993; Li et al., 1993; Gale et al., 1993). Most recently, the *vav* proto-oncogene containing SH2 and SH3 domains was shown to confer guanine nucleotide exchange activity (Gulbins et al., 1993). It was suggested that the SH3 domain may function to mediate protein-protein interactions, perhaps serving to couple the signalling of RPTKs to systems regulated by small GTP-binding proteins (Mayer and Baltimore, 1993).

Recently, some cytoplasmic protein tyrosine phosphatases (PTP) containing 2 consecutive SH2 sequences N-terminal to the phosphatase catalytic domain were shown to be recruited to some activated receptors, linking this type of PTPase to the signal transduction pathway of RPTKs. Among them, PTP1C (also known as SH-PTP1, SHP and HCP) (Shen et al., 1991;Mathews et al., 1992; Plutzky et al., 1992; Yi et al., 1992)

specifically expressed in the hematopoeitic and epithelial cells can bind to activated EGF receptor(Shen et al., 1991) and chimeric EGFR/Neu receptor (Vogel et al., 1993). PTP1D (Vogel et al., 1993), also known as Syp (Feng et al., 1993) was believed to be the mammalian homolog of the *Drosophila* gene *csw* (Perkins et al., 1992). The PTP1D was widely expressed throughout the mouse embryonic development and in adult tissues (Feng et al., 1993; Vogel et al., 1993). Whereas PTP1C led to partial and complete dephosphorylation of EGF-R, PDGFR, IR and IGFR (Vogel et al., 1993), PTP1D failed to dephosphorylate the above RPTKs. In contrast, PTP1D could associate with several activetd receptors including EGFR, PDGFR, Her2-Neu, c-Kit and was activated upon binding to the activated PDGFR (Vogel et al., 1993). It thus suggests that PTP1D has a positive rather than negative control in the PDGFR signalling (Vogel et al., 1993).

$p74^{raf}$ along with a number of ser/thr protein kinases such as microtubule-associated protein kinases (MAP2 kinase) and a series of newly characterized ser/thr kinases (*erk*s) (Boulton et al., 1991) serve to transmit signals from tyrosine phosphorylation into ser/thr phosphorylation. Raf-1 can be tyrosine-phosphorylated upon PDGF stimulation and become associated with the activated receptor (Morrison et al., 1989 & 1990). The fact that antisense RNA of c-raf and its dominant negative Raf-1 could inhibit serum-induced NIH3T3 cell proliferation and block *ras*-induced cell transformation strongly suggests that Raf-1 is a downstream signal transducer of Ras (Kolch et al., 1991). There is evidence supporting that, between c-raf and MAP Kinase, there could exist other components such as MAPKK (MAP Kinase Kinase). Very recently, the MAPKK was molecularly cloned by several groups (Gomez & Cohen, 1991; Crews & Erickson, 1992; Crews et al., 1992; Alessandrini et al., 1992). The latest evidence suggests that GTP-bound Ras, Raf and MAPKK could form a complex, thus connecting the *ras* pathway

to the ser/thr phosphorylation signalling pathway. Growth factors such as insulin, EGF, NGF and phobol esters activate MAP kinase by increasing its tyrosine and threonine phosphorylation. Activated MAP kinase then phosphorylates another serine/threonine kinase Rsk (ribosomal S6 protein kinase) (Sturgill et al., 1988; Sturgill & Wu, 1991; Gomez & Cohen, 1991; Chen et al., 1992). A fraction of the activated MAP kinase and Rsk enter the nucleus (Blenis, 1991; Chen et al., 1992) which could connect the *ras* signal transduction pathway to some nuclear proto-oncogene products, i.e. transcriptional factors. Among them *jun* and *fos* have been shown to be activated through PKC and PTK pathways (Cantley et al., 1991; Hunter, 1991).

Genetic evidence for the relationship between the substrates and their upstream components is largely lacking. However, the requirement of *raf* homologous gene *polehole* in *Drosophila* for the function of the fly EGFR homolog DER (Ambrosio et al., 1989; Rogge et al., 1991) supports the notion that *raf* is a downstream component of the EGFR signalling. The genetic study in fly eye devlopment also suggests that Grb2, *sos*, *ras* and c-*raf* are the downstream components in the *sevenless* signalling (Simon et al., 1991; Rogge et al., 1991; Bonfini et al., 1992; Olivier et al., 1993; Simon et al., 1993; Dickson et al., 1992). Similar signaling transduction pathway was identified in the system governing the vulval development of the Caenorhabditis elegans in which *let*-23, *let*-60 and *sem*-5 are the homolog of the vertebrate EGFR, *ras* and Grb-2, respectively (Aroian et al., 1990; Han & Sternberg, 1990; Beitel et al., 1990; Clark et al., 1992).

III. Viral oncogene v-*ros* and proto-oncogene c-*ros*

Avian sarcoma virus (ASV) UR2 is a replication-defective virus which was isolated together with its associated helper virus UR2AV from a spontaneous chicken tumor

(Balduzzi et al., 1981). The genome of UR2 contains 1.2-kb stretch of transformation-specific sequence, called *ros* (Wang et al., 1982). UR2 was presumably generated by recombination between UR2AV and c-*ros* at the expense of most of the replicative sequences in UR2AV. As a result, *ros* was fused to the 5' region of the UR2AV sequence which codes for portion of the viral structural protein p19 (Wang et al.,1982). The fused p19 and *ros* sequence in UR2 codes for a 68-kDa protein called $p68^{gag\cdot ros}$ resembling RPTKs (Neckameyer & Wang, 1985). The $p68^{gag\cdot ros}$ is a TM molecule with the $p19^{gag}$ portion protruding extracellularly (Jong & Wang, 1987). No post-translational processing or modification of P68, except phosphorylation, could be detected (Feldman et al., 1982; Garber et al., 1985). $p68^{gag\cdot ros}$ can be immunoprecipitated with antibodies against viral *gag* and is associated with a protein tyrosine kinase activity capable of phosphorylating itself and foreign substrates including rabbit IgG and α-casein at tyrosine residues (Feldman et al., 1982; Garber et al., 1985). The *in vitro* phosphorylated P68 migrates in SDS-PAGE as multiple species, suggesting that there are more than one phosphate-acceptor tyrosine residues (Jong & Wang, 1991). P68 kinase activity is different from those of other ASVs in its pH optimum, cation preference and phosphate donors (Feldmane et al., 1982). By subcellular fractionation, P68 was found to be mostly associated with the membrane fractions and became membrane associated very rapidly during its biosynthesis (Garber et al., 1985). Furthermore, the p19 portion was found to be essential for the transforming ability of $P68^{gag\cdot ros}$ but not important for its membrane association (Jong & Wang, 1990). The transmembrane domain of $p68^{gag\cdot ros}$ was also found to be important for its transforming activity most likely due to the role of the TM domain in directing initial membrane association of P68 (Jong & Wang, 1991). P68 was reported to be able to associate with PI-3-kinase (Fukui et al., 1989; Zong & Wang, unpublished results).

A recombinant DNA clone containing chicken cellular sequences homologous to v-*ros* was isolated from a genomic DNA library (Neckameyer et al., 1986). Later, a 2-kb 3' cDNA clone was isolated from a chicken kidney cDNA library (Podell & Sefton, 1987). Comparison of the nucleotide sequences of v-*ros* and c-*ros* shows 3 differences: (1) v-*ros* is truncated 7 aa before the TM domain and fused to *gag*. (2) v-*ros* contains a 9 nucleotide insertion within the the TM domain. (3) v-*ros* has an internal 36 nucleotide deletion between nucleotide 6918 and 6953 (Fig. 3-2) and a 3' 27 nucleotide terminal truncation and fusion to the *env* sequence. Based on the comparison of the sequences in the PTK domains, c-*ros* was found to share the greatest homology with insulin receptor (Ebina et al., 1985; Ullrich et al., 1985), insulin-like growth factor I receptor (Ullrich et al., 1986) and *Drosophila sevenless* protein (Basler & Hafen, 1988; Bowtell et al., 1988) among the known members of the PTK family. However, the presumed ligand for the c-*ros* product remains unknown.

Previous studies indicated that the expression of c-*ros* was highly regulated in chicken and was only detectable in kidney (Shibuya et al., 1982; Neckameyer et al., 1986) where a 8.3-kb mRNA was observed (Podell & Sefton, 1987). Most recently, a study of c-*ros* expression in rat, however, showed that it was expressed in lung, heart and testis in addition to kidney and tissue-specific alternative splice forms were observed (Matsushime & Shibuya, 1990). A survey of c-*ros* expression in various human tumor cell lines revealed a high incidence of elevated c-*ros* expression in glioblastoma cell lines (Birchmeier et al., 1987). In some of them, portions of the c-*ros* gene were found to be rearranged. Confirmation of tumorigenic potential of the c-*ros* gene was supported by the activation of c-*ros* via DNA rearrangement during NIH 3T3 cell transfection and nude mice passage of a mammary carcinoma cell line-derived DNAs (Fasano et al., 1984; Birchmeier

et al., 1986). Interestingly, the truncation point of that DNA rearrangement matches precisely with that of the spontaneous transduction of c-*ros* by avian retrovirus in UR2 (Neckameyer et al., 1986) which was suggested to have arisen by splicing between viral and c-*ros* sequence (Neckameyer et al., 1986). The human c-*ros* gene was mapped to the chromosome region 6q16-6q22 (Nagarajan et al., 1986). It thus maps at the vicinity of the c-myb gene which is located at the long arm of the chromosome 6. The possibility that c-*ros* may be involved in the chromosome 6q deletion and rearrangement in various malignancies was raised (Nagarajan et al., 1986).

cDNAs for proto-oncogene c-*ros* were recently identified and molecularly cloned from different species, including human, rat and chicken (Birchmeier et al., 1990; Matsushime & Shibuya, 1990; Chen et al., 1991). Partial cDNA clones for mouse c-*ros* were isolated (Sonnenbreg et al., 1991; Tessarolo et al., 1992). A cDNA of 8-kb was isolated and cloned from a glioblastoma cell line which might represent human c-*ros* cDNA. However, whether there were mutations in this human gene remains to be determined (Birchmeier et al., 1990). The c-*ros* cDNA sequences encodes a large RPTK sharing high homology with that of the *Drosophila* sevenless protein. The predicted c-*ros* protein contains one N-terminal signal peptide and one internal TM domain. The large EC domain (more than 1800 aa) contains multiple cystein residues and potential N-linked glycosylation sites. A cysteine-rich motif was identified within the first 100 aa of chicken c-*ros* cDNA (Chen et al, 1991). The spatial distribution of cystein residues in this motif was conserved in both rat and human c-*ros* cDNA, but not in the *sevenless* protein. Notably, eight fibronectin type-III repeats were conserved among *sevenless*, chicken, rat and human c-*ros* proteins (Norton et al., 1990; Sonnenberg et al., 1991; and this study). It has been known that the type III repeats of fibronectin participate in activities such as binding

to cell surface integrins and heparin, however, the function of those eight FN repeats in either sevenless or c-*ros* proteins is unknown. The kinase domain of c-Ros shares the greatest homology with the sevenless protein, and less to those of IR and IGFR. The four RPTKs also share another charateristic cluster of 3 tyrosine residues presumed to be the major autophosphorylation sites. It was thus suggested that c-Ros, Sevenless, IR and IGFR are closely related members of RPTK family (Chen et al., 1991). However, another characteristic 6 aa insertion in the c-*ros* kinase domain is shared only with Sevenless but not with IR and IGFR, suggesting that c-*ros* and *sevenless* have a closer genetic relationship.

The tissue-specific expression of c-*ros* in mouse was recently characterized by RNase protection and *in situ* mRNA hybridization (Sonnenberg et al., 1991; Tassarollo et al., 1992). It was shown that c-*ros* was expressed in kidney, intestine, lung and testis. The expression of c-*ros* was under a tight temporal and spatial regulation. The c-*ros* mRNA was mainly localized in the epithelial cells of the kidney collecting ducts and villi and crypts of the intestine. The onset and the pattern of c-*ros* expression in the kidney was coincident with a well known reciprocal epithelial/mesenchymal induction event during the kidney organogenesis. Based on the above phenomenon and the nature of the ligand for the Drosophila homolog Sevenless, which is a TM protein expressed on the neighboring cell surface of R7 precursor cells (Krämer et al., 1991), it was suggested that the ligand for c-Ros is very likely to be localized on the cell surface of the mesenchyme cells surrounding the epithelial cells of the branching collecting ducts (Sonnenberg et al., 1991; Tessarollo et al., 1992).

To study the biological function and search for the ligand of the c-*ros* product, I have isolated c-*ros* cDNA clones and determined their entire sequences. The c-*ros*

product shares a remarkable sequence and structural homology with the *Drosophila sevenless* protein. Functional study of c-*ros* protein demonstrates that the c-*ros* protein is a cell-surface protein with kinase activity. The putative c-*ros* promoter was isolated and shown to be functionally active. The tissue-specific expression pattern of c-*ros* was examined using RNase protection assay and *in situ* mRNA hybridization. Our data suggest that c-*ros* may play roles in the development and mature functions of kidney, intestine, bursa, lung, and possibly the thymus and testis.

Chapter II. Materials and Methods

Cell line and virus

COS-1 and COS-7 were maintained in DMEM with 10% fetal calf serum. Chicken embryo fibroblasts (CEF) were maintained in F10 medium supplemented with 5% calf serum and 1% chicken serum. The preparation of CEF and colony formation assay of virus-infected CEF were performed according to published procedures (Hanafusa, 1969). The retrovirus UR2, SrcXRos and VCros were described before (Jong & Wang, 1987; Jong & Wang, 1991; Zong et al., 1993).

Antiserum and purified antibodies

α-Ros antiserum 219 was described before (Jong & Wang, 1987). Pure α-Ros IgG was purified by affinity chromatography on a lacZ/Ros c-terminus fusion protein coupled affinity column. α-pTyr antibody Py20 is purchased from ICN and PT22-1 was made in collaboration with Dr. Tom Moran. α-Shc antiserum is a kind gift from Dr. Tony Pawson. Monoclonal α-PLCγ1 is purchased from UBI. Monoclonal Ab 5C2 was generated in collaboration with Dr. Moran. RC20 (Transduction laboratories) is a recombinant α-pTyr Py20 conjugated with alkaline phosphatase.

RNA preparation and dot blot hybridization

Total RNAs from various tissues were isolated by guanidinium thiocyanate extraction followed by centrifugation on a cesium chloride cushion according to the published methods (Glisin et al., 1974; Ullrich et al., 1977). Polyadenylated RNAs were isolated as described previously (Wang et al., 1974; Wang et al., 1981). The dot blot analysis was done following the procedure provided by the supplier of the nitrocellulose paper (Schleicher and Shuell). Hybridization was performed in a solution containing 0.75 M NaCl, 50 mM sodium phosphate (pH 7.4), 5 mM EDTA, 0.2% sodium dodecyl sulfate

(SDS), 50% formamide, 200 μg/ml each of salmon sperm DNA and yeast tRNA, 0.1% each of ficoll, polyvinylpyrrolidone and bovine serum albumin and ^{32}P-labelled v-*ros* riboprobes. The hybridization was performed at 60°C for 12 to 16 hours. The riboprobes were prepared from 5' v-*ros* DNAs subcloned in pGEM3 vectors using T7 RNA polymerase. The EcoRV cut 5' v-*ros* DNAs were chosen to exclude the region containing the conserved PTK domain to minimize cross hybridization with other PTK gene encoded messages. The specific activity of riboprobes ranged from $5X10^8$ to $1X10^9$ cpm per μg DNA. Filters were washed 3 times in a solution (1X SSPE) containing 0.15 M NaCl, 10 mM sodium phosphate (pH7.4), 1 mM EDTA and 0.1% SDS at 65°C for 20 minutes each time followed by washing in a solution of 0.1X SSPE and 0.1% SDS at 60°C for 50 minutes. After rinsing the filters in a buffer containing 0.5 M NaCl, 10 mM Tris.HCl (pH7.4), 1 mM EDTA, they were incubated in the same buffer containing 20 μg per ml of RNase A at 37°C for 30 minutes. Finally, the filters were rinsed in the same buffer without RNase and dried and processed for autoradiography.

Construction of cDNA libraries

For the oligo(dT)-primed library, 15 μg of polyadenylated kidney RNA was mixed with 5 μg of oligo$(dT)_{12\text{-}18}$ and cDNA synthesis was carried out as described (Watson & Jackson, 1985; Dorai & Wang, 1990). The cDNAs were ligated to phage vector λgt10 at the EcoRI sites. A library of $3.5X10^7$ recombinant phages was obtained. For the *ros* oligonucleotide-primed cDNA library, a 15-mer deoxyoligonucleotide representing the 5' v-*ros* sequence was synthesized and used to prime cDNA synthesis from 6 μg of kidney polyadenylated RNA. After first strand cDNA synthesis, the second strand DNA synthesis and subsequent linker ligation was performed using a commercial cDNA synthesis kit (Pharmacia) according to the conditions suggested by the manufacturer. In this method

a linker containing staggered and dephosphorylated EcoRI and an internal NotI site was used. It thus avoided polymerization of linker and omitted the steps of methylation and EcoRI digestion following linker ligation. The cDNAs were then phosphorylated and ligated to dephosophorylated EcoRI arms of phage λZAPII DNAs (Stratagene). A library of 1.6X10^6 recombinant phages was obtained.

Isolation of cDNA clones

A EcoRV DNA fragment derived from the 5' v-*ros* (Neckameyer & Wang, 1985) was used as a probe to isolate a phage clone 84-1 containing the 3' 3-kb c-*ros* cDNA insert from the oligo(dT)-primed library. A probe derived from the 5' region of this initial cDNA clone was used to isolate further upstream cDNA clones from the oligonucleotide-primed library. With such a "cDNA walking" strategy, several overlapping upstream clones were isolated. All the cDNA inserts except the initial 3-kb clone were subcloned into the plasmid vector pBluescript SK(-) by being rescued into phagemid with helper phage R408 or VcsM13 according to conditions provided by the supplier (Stratagene).

Nucleotide sequencing

Two strategies of sequencing were undertaken. The cDNAs were digested with appropriate restriction enzymes and subcloned into M13 phage vectors for sequencing. Alternatively, individual full-length cDNA inserts were unidirectionally deleted from either end using exoIII (Henikoff, 1984) and S1 nuclease (Stratagene, BRL and USB) and then religated after blunting the ends to generate a series of overlapping deletion clones. Clones spanning the entire insert and with less than 200 bp overlapping between adjacent clones were isolated for sequencing. Final sequence was constructed after sequencing both directions of each cDNA insert using Sanger's dideoxy method (Sanger et al., 1977). The RACE clones were sequenced using primer in the vector pBluescript and primer

JMC4. For the 2.5-kb HindIII genomic DNA fragment containing the potential c-*ros* promoter, the restriction sites as shown in Fig. 6-1 were utilized to subclone those fragments into pBluescript SK+ or SK- and the sequencing was done using the primers in the vector.

Construction of full length c-*ros* cDNA from overlapping kidney cDNAs

Four previously described overlapping cDNA clones 5b, 19b, 10a and 84-1 (in 5' to 3' order) (Chen et al., 1991) were used to construct the full length c-*ros* cDNA in plasmid pBluescript SK(+) or SK(-) (Stratagene) (Fig. 4-1A). This was done by using the unique BstBI site in clone 5b, 19b and 10a, the unique SacI site in 3' end of 19b and 84-1 to form the full length cDNA. Most of the 5' and 3' non-coding sequence of the resulting full length cDNA was then deleted by using PCR method with a pair of synthesized oligodeoxynucleotides each containing a NotI site. The resultant plasmid was called pSKros.

Construction of expression plasmids for c-*ros* and mutant ppros

For transient expression of c-*ros* in the COS-7 cell, pECE vector (Ellis et al., 1986) containing SV40 early promoter, SV40 replication origin and polyadenylation signal was used. The 7-kb full length c-*ros* was freed by NotI digestion and inserted into a modified pECE vector containing a NotI site in the polylinker region, resulting in pECEros (Fig. 4-1B). The deletion variant ppros was initially engineered in pSKros by deleting 4983 bops of the EC sequence flanked by PstI and PvuII sites to produce pSKppros. The ppros sequence was then excised for pSKppros by NotI digestion and inserted into the expression vector pRc/CMV (Invitrogen) under the control of the human cytomegalovirus (CMV) early promoter and enhancer. The resulted plasmid was named as pCMVppros (Fig. 4-1B). Similarly, the full length c-*ros* cDNA was inserted into the pRC/CMV to

generate pCMVros (Fig. 4-1B).

Construction of retroviral expression plasmids

For expression of c-*ros* proteins in CEF, the full length c-*ros* and ppros were introduced individually into the pUIGFRΔATG in which the *gag* initiation codon was mutated (Liu et al., 1992, 1993). The resulted plasmids are called pUfcros and pUppcros respectively (Fig. 4-1C). These plasmids encode the full length or 5' internally deleted c-*ros* proteins using its native initiation codon. The c-*ros* sequence is flanked by the viral long terminal repeats (LTRs) in these plasmids and can be transfected directly into CEF for their expression.

Establishment of permanent c-*ros* expressing cell line

The c-*ros* expression plasmid pCMVros was transfected into COS-1 cells (Gluzman et al., 1981) as described above. After 2 weeks of selection in the presence of 400 µg/ml G418, the single colonies were picked up and amplified. Western Blotting was then performed to screen for the *ros* positive clones. Clone CMVros6 was thus selected and used for further protein analysis.

Extraction of protein from tissues and assay for the c-*ros* protein

One gram of tissue (kidney or intestine) was suspended in 5 ml of PBS/1mM PMSF and homogenized with a polytron at high speed for 1 min. The pellet was resuspended in 5 ml of RIPA lysis buffer, sonicated and extensively vortexed. After centrifugation for 10 min at 7000 rpm (JA17 rotor), the supernatant was incubated with 3 µg purified rabbit α-Ros IgG in cold room overnight followed by addition of 50 µl of protein A-sepharose beads and incubation for more than one hour. The immunoprecipitate was washed extensively and subjected to kinase assay.

Fluorescence immunostaining of *ros* proteins transiently expressed in COS-7 cells

48 h after transfection, cells in 3.5-cm dishes were washed three times with PBS, fixed with 4% paraformaldehyde/PBS at room temperature for 10 min. The fixed cells were washed three times with PBS and permeablized with cold methanol for 10 min before being washed with PBS. The cells were blocked with 5% goat serum/PBS for half an hour and sequentially incubated with purified α-Ros antibody, biotinylated donkey α-rabbit antibody and avidin-FITC. DAPI was used to specifically stain the nucleus.

Biotinylation of cell surface proteins

Transient transfected COS-7 cells or UR2 and SrcXRos (JONG & Wang, 1991) transformed CEF cells in 6 cm dishes were washed with cold PBS three times and put on ice. 1 ml of PBS containing 0.4 mM Sulfo-NHS-biotin (Pierce) was added and incubated for 2 hours on ice with occasional mixing. The unreacted Sulfo-NHS-biotin was quenched with the addition of regular DMEM medium. After extensive wash with PBS, the cells were lysed in RIPA buffer and α-Ros immunoprecipitates were subjected to SDS-PAGE and transferred to nitrocellulose (NC) paper. The blot was incubated with avidin-alkaline phosphatase (Boehrenger Mannheim, 1:5000 dilution) and developed by method provided by the manufacturer (Promega).

PI-3-kinase assay

The PI-3-kinase assay was done essentially as described previously (Liu et al., 1992). In brief, UR2 transformed CEF or ppros transfected CEF were lysed in Nonidet P-40 (NP40) buffer (20 mM Tris-HCl, pH7.5, 5 mM EDTA, 150 mM NaCl, 1% NP40, 1 mM Na_3VO_4, 1 mM phenylmethanesulfonyl fluoride (PMSF), 100 mM NaF, 50 mM sodium pyrophosphate). After clearing of the cell lysates by centrifugation at 12,000g for 10 min, the supernatant was incubated with α-Ros antiserum (1:1,000 dilution) for 1 hour at 4°C. 15 μl of protein A-agarose beads (Repligen) was added and the mixture was incubated

for 1 hour at 4°C. The immunoprecipitates were washed as described previously (Fukui et al., 1989), and the washed beads were resuspended in 25 μl of TGN buffer (20 mM Tris-HCl, pH7.5, 100 mM NaCl, 0.5 mM EGTA). 10 μg of the substrate PI (20 μg/μl in dimethyl sulfoxide, Avanti Polar Lipids, Inc.) was then added to the resuspended immunoprecipitates and mixed to make micelles of PI. The mixture was incubated at room temperature for 10 min. Pre-mixed [γ-^{32}P]-ATP (10 μCi/assay, NEN) and $MgCl_2$ (final concentration 20 mM) were then added, and the mixture was incubated at room temperature for 10 min. PIP was extracted and analyzed on a thin layer chromatography silica gel 60 plate (Merck) exactly as described previously (Fukui et al., 1989).

Protein analysis

Fractionation and sucrose gradient sedimentation was done according to the published methods (Jong & Wang, 1990).

For metabolic labeling, cells were transferred at 50% confluence one day earlier or used directly 48 hours after transfection. Cells were starved for 2 hours in Met-free MEM followed by [^{35}S]-Met labeling at 100 μCi/ml for 4 hours. In the case of the pulse-chase experiment, cells were labeled with 200 μCi/ml [^{35}S]-Met for 20 min before complete medium was added for chasing. For tunicamycin treatment, cells were pretreated with the drug at a concentration of 10 μg/ml for 2 hours and continued throughout the labeling period.

For immunoprecipitation (Garber et al., 1985), labeled cells were extracted with RIPA (50 mM Tris-Cl pH-7.5, 150 mM NaCl, 5 mM EDTA, 1% Triton X-100, 0.1% SDS, 1% sodium deoxycholate, 1% Trasylol and 1 mM phenylmethyl sulfonyl fluoride). The α-Ros antiserum (Jong & Wang, 1987) was then added to cleared supernatant of the cell extract for another 45 min, followed by protein A-sepharose (Pharmacia) binding for

another 45 minutes. The immune complex was washed with high salt RIPA (300 mM NaCl) three times and low salt RIPA (10 mM NaCl) once before it was boiled in the sample buffer for SDS-PAGE (Laemmli et al., 1970).

For kinase assays (Feldman et al., 1982), unlabeled cells were extracted with RIPA containing 0.05% SDS and the cell extract was immunoprecipitated as describe above. The immune complex was then washed three times with high salt RIPA, and twice with kinase buffer containing 50 mM Tris-HCl (pH8.0) and 10 mM MnCl2. The kinase reaction was carried out in the presence of 5 μCi [γ-P^{32}] ATP at room temperature for 15 minutes. After kinase reaction the immune complex was further washed with RIPA as described above before SDS-PAGE.

Western blot analysis was done according to the method described by Hamaguchi et al. (1988) with the following modifications. The blots were blocked at room temperature for 2 minutes with 5% nonfat dry milk in 1XTBST (10 mM Tris-HCl pH8.0, 150 mM NaCl, 0.1% Triton X-100, 0.02% Sodium azide). Binding with different antibodies was performed in 1XTBST with 5% nonfat milk overnight in cold room or on hour at room temperature. Then the blots were rinsed with blocking solution. If necessary, the filter was incubated with secondary antibodies (usually rabbit α-mouse IgG) for one hour at room temperature. Finally, the blot was incubated with 1 μCi of [^{125}I]protein A (ICN) in 1XTBST containing 5% milk at 37^{0}C for 1 hour. The blot was washed thoroughly, dried and autoradiographed with an intensifying screen at -70^{0}C.

Construction of pCMVrosAP and pCDMrosFc and production of RosAP and RosFc

Details are described in the Results and Discussion of chapter 4.

Preparation of various c-*ros* riboprobes

The isolation of the genomic clone corresponding to the 5' c-*ros* will be described

elsewhere (Chen, J., Tong, J., and Wang, L.-H., unpublished results). A 1.3-kb genomic HindIII-MscI fragment was subcloned into the HindIII and SalI (blunted with klenow) sites of pBluescript SK+ (SK+) (Invitrogen), which was called pSKrospro. Probe A was synthesized with T7 RNA polymerase using the ClaI digested pSKrospro as template. Probe A contains 809 nt from genomic sequence and some from the vector. Probe B was synthesized with T7 RNA polymerase using the SacI linearized cDNA clone 5b in SK+ which contains 231 nt from 5b (Chen et al., 1991). The 0.48-kb XbaI-HindIII fragment in cDNA clone 10a was cloned into SK+, resulting in plasmid pSKXH. Probe C was synthesized from the SpeI linearized pSKXH containing 235 nt of c-*ros*. Probe E has sense and anti-sense orientations. The sense probe was synthesized with the T3 RNA polymerase using HindIII linearized pSKXH as template, and the antisense probe, using T7 RNA polymerase with XbaI digested pSKXH as the template. Both E probes contain about 480 nt from c-*ros*. Probe D was synthesized with T3 RNA polymerase and contains 189 nt from 3' c-*ros* cDNA including 29 nt of non-coding sequence. Probe N was synthesized with T7 RNA polymerase using the PvuII linearized clone 2a in SK+ (Fig. 3-2), which contains 503 nt of clone 2a.

Northern blot analysis

For Northern analysis, 6 μg of poly(A)+ RNA was denatured and fractionated by electrophoresis in a 1% agarose/2.2 M formaldehyde gel (Sambrook et al., 1989). Subsequently the gel was denatured with 0.1 N NaOH, neutralized with 20X SSC and transferred onto a nitrocellulose membrane by capillary action and then baked at 80°C for 2 hrs. Following prehybridization, the blot was hybridized with a ^{32}P-labeled antisense riboprobe corresponding to the 500 nt of the 3' end of the c-*ros* cDNA clone 2a (Fig. 3-2). Hybridization was carried out overnight at 72°C in the presence of 50% formamide, 5X

SSC, 0.1% SDS, 5X Denhardt's solution and 200 μg/ml denatured salmon sperm DNA. The membrane was then washed in 0.1X SSC at increasing temperatures up to 80°C and exposed at -70°C for 3 days using an intensifying screen.

RNase Protection Assay

RNase Protection Assay (RPA) was performed with the RPA kit (Ambion) according to the manufacturer's instruction. In short, 30 μg of total RNAs were mixed with $1X10^5$ cpm ^{32}P-labeled antisense riboprobe, precipitated and resuspended in 20 μl hybridization buffer containing 40 mM PIPES (pH6.4), 1 mM EDTA, 0.4 M NaCl and 80% formamide. After overnight hybridization at 50°C, RNase A/T1 digestion was performed at 37°C for 30 minutes followed by protease K digestion and phenol/chloroform extraction. The protected RNA fragments were precipitated and resuspended in 10 μl loading buffer and separated by electrophoresis in a 6% polyacrylamide/urea gel at 15 Watts for 2 hrs. After fixation and drying, the gel was exposed at -70°C for desired time under an intensifying screen.

mRNA *in situ* hybridization

In situ hybridization was done according to the method developed by Tiedge (1990) with the following modification. The fresh chicken tissue was dissected, embedded in OCT medium, quickly frozen in dry ice/acetone bath, sectioned at 5 microns and stored at -70°C before subjected to hybridization. The frozen sections were fixed with 4% paraformaldehyde in PBS at room temperature for 20 minutes before UV-illuminated at 30 W from 30 cm for another 20 minutes. The sections were pre-hybridized in the buffer containing 1.2 M NaCl, 20 mM Tris/HCl (pH7.5), 0.04% ficoll, 0.04% polyvivylpyrrolidone, 0.2% BSA, 20 mM EDTA (pH7.5), 1 mg/ml salmon sperm DNA (Sheared), 1 mg/ml total yeast RNA and 0.1 mg/ml yeast tRNA at 50°C for 1 hr, and the [^{35}S]-UTP labeled

riboprobe was added at a concentration of 2500-5000 cpm per ml in the hybridization buffer containing 1.2 M NaCl, 20 mM Tris/HCl (pH7.5), 0.04% ficoll, 0.04% polyvivylpyrrolidone, 0.2% BSA, 20 mM EDTA(pH7.5), 0.2 mg/ml salmon sperm DNA (Sheared), 0.1 mg/ml total yeast RNA and 0.1 mg/ml yeast tRNA. After overnight hybridization, the sections were washed sequentially in 2X SSC at 50°C for 1 hr, RNase A digestion in the buffer containing 0.5 M NaCl and 10 mM Tris/HCl (pH 8.0) at 37°C for 1 hr, 2X SSC at 50°C for 1 hr, 0.1X SSC, 14 mM 2-mercaptoethanol and 0.05% Sodium pyrophosphate at 50-65°C for 3 hrs, and finally cooled down at room temperature overnight. The dried sections were first exposed to X-ray film at room temperature to determine the exposure time for the NBT2 emulsion before being dipped in NBT2 emulsion, exposed in dark at 4°C for 10-20 days and then developed. Nuclei were stained by crystal violet. The slide pictures were taken under a Leitz microscope.

Primer extension and RPA for mapping the transcription initiation site

Primer extension was performed according to the method described in Sambrook et al. (1989) with following modifications. $1X10^5$ cpm ^{32}P-labeled oligodeoxynucleotide JMC54 (5'-CCATACAAACTCCTTCAAGTCTAGGTGACA-3') complementary to the cDNA sequence from nucleotide +182 to +151 (Fig. 6-2) was annealed with 50 µg total RNA in 30 µl hybridization buffer containing 40 mM PIPES (pH 6.4), 1 mM EDTA (pH 8.0), 0.4 M NaCl and 80% formamide at 30°C for 12-15 hours. After precipitation, the reverse transcription was performed using Mo-MuLV reverse transcriptase (BRL) at 37°C for 1 hr. After RNase digestion and subsequent phenol/chloroform extraction, the product was precipitated and resuspended in 10 µl loading buffer, denatured at 100°C for 5 min, and separated by electrophoresis in a 6% polyacrylamide/urea gel. The fixed and dried gel was then exposed at -70°C for 7 days under an intensifying screen.

RPA was performed as described earlier. In order to make the probe A protected product have the same 5' end as the primer extension product by JMC 54, the 2.5-kb HindIII fragment in pBluescript SK(+) was cut with MscI (fig. 6-1) and SalI (this site is in the vector) and self ligated to yield plasmid pSKrospro. The plasmid was linearized with ClaI (Fig. 6-1) and the antisense riboprobe was made by using T7 RNA polymerase.

Cloning of the 5' cDNA

The 5' RACE method (Frohman et al., 1988) was used to isolate clones corresponding to the 5' untranslated region of the c-*ros* mRNA. In short, 2 µg chicken kidney mRNA was reverse-transcribed with RT primer JMC1 (5'-AGCTGCAGCTGAGAGGTTAT-3'). The first strand cDNA was then tailed with dATP using terminal transferase (BRL). Reaction mixture was diluted to 500 µl with TE buffer and 10 µl was amplified for 30 cycles using 200 ng gene-specific primer JMC2 (5'-CGGAATTCATATGCATCCTCTGCACGA C-3') upstream of the RT primer, 50 ng dT_{14}-adapter primer(5'-GATCTAGAGTCGACATCGATTTTTTTTTTTTTTT-3'), and 200 ng adapter primer (5'-GATCTAGAGTCGACATCGAT-3') (Dorai & Wang, 1991). The PCR product was purified, blunt-ended and ligated into plasmid pBluescript SK(+) (Stratagene). After transformation, about 500 colonies were screened using ^{32}P-labeled oligodeoxynucleotide JMC4 (5'-GGGTCGACAATCCATTCCTTGAACAA-3') and 4 clones turned out to be positive. Plasmid DNA was then prepared and sequenced.

Isolation of Genomic Clones

A chicken genomic library (Vennstron & Bishop, 1980) was screened using the most 5' 0.7-kb cDNA as probe to isolate the c-*ros* promoter region. Out of $2X10^5$ plaques, 3 positive clones were picked. The 15-kb clone 6a which is positive for a 5' cDNA primer JMC4 was amplified for further study. A JMC4 positive 2.5-kb HindIII fragment was then

subcloned and sequenced (Fig. 6-1).

Cat Expression Vectors and CAT Assay

The putative *ros* promoter region HindIII-XbaI fragment (1.3-kb)(Fig. 6-1 and 6-2), of which a 3' XbaI site was added by PCR, was subcloned into the pCat-enhancer (Promega) vector between HindIII and XbaI sites. The resulting plasmid was called pCatHind. The vector pCat-enhancer (Promega) contains a SV40 enhancer at the 3' of the Cat gene which stimulates a foreign promoter inserted before the Cat gene. Since the basic vector has a high basal CAT activity when transfected into COS-7 cell, a 0.13-kb HpaI-BamHI fragment from pSV2Cat (Gorman et al., 1982) containing the SV40 polyadenylation signal (SVpA) was then inserted between HindIII and BamHI to reduce the background caused by some non-specific transcription initiated upstream of the Cat gene. The new plasmid was called pCatpA (Fig. 6-4A). A series of deletion mutants were generated by deleting 5' distal region between BamHI and ClaI, PacI or XbaI sites, and named pCatpACla, pCatpAPac and pCatXba, respectively (Fig. 6-4B). Some other deletion mutants were generated by PCR method (Fig. 6-4B). 20 μg each of the plasmids together with 2 μg of pLacZ was transfected into COS-7 cells using calcium phosphate method. 48 hrs after transfection, cell extracts were subjected to CAT assay, and duplicate aliguotes of the cell extracts were used for assay of the galactosidase activity. 48 hours later, the cells were collected and lysed by repeated freezing and thawing for three times. After pelleting the cell debries at 4°C, the supernatant protein concentration was measured and 10 μg protein was used for each CAT assay in a reaction mixture containing 2 μl of [^{14}C]-Chloramphenicol (0.1μCi, ICN), 200 ng Acetyl-CoA (Boehringer manheim) and 0.25 M Tris/HCl (Gorman et al., 1982). The CAT activity was obtained by determining the percentage of Cholorophenicol that was acetylated. The β-galactosidase assay was

performed as previously described (Chen et al., 1991). 10 μg of total protein extract was incubated with 0.77 mM 4-methyl-umbelliferyl-β-D-galactoside in a final volume of 230 μl of buffer Z (60 mM Na_2HPO4, 40 mM NaH_2PO4, 10 mM KCl, 1 mM MgSO4, 50 mM 2-mercaptoethanol, pH7.5) at 37^{0}C for 25 min in a Titer-tek, Fluoroskan II. The absolut fluorescence units were used to normalize each CAT activity.

Chapter III. The proto-oncogene c-*ros* codes for a transmembrane protein tyrosine kinase sharing sequence and structural homology with *sevenless* protein of *Drosophila melanogaster*

Jianmin Chen, Diana Heller, Betty Poon, Lily Kang and Lu-Hai Wang

The results of this chapter were published in Oncogene (1991) 6, 257-264.

Permission was obtained from publisher to use the published data in this chapter.

Abstract

Our earlier study predicted that proto-oncogene c-*ros* codes for a receptor-like protein tyrosine kinase. To further understand its protein structure and physiological function, we have analyzed its expression in various tissues of chicken and have isolated and sequenced cDNA clones containing the entire coding region of the gene. Confirming our earlier study, we found that kidney is the organ that expresses the highest level of c-*ros* mRNA. In addition, we found a lower level of expression in gonad, thymus, bursa and brain. A distinctive 8.3-kb c-*ros* mRNA is present in kidney. No detectable amount of c-*ros* mRNA was found in the rest of tissues examined. Nucleotide sequence of the c-*ros* cDNA predicts that it codes for a TM PTK molecule of 2311 aa. The EC domain consists of 1873 amino acids which share 20 to 43% homology with that of the *Drosophila sevenless* protein and PTK domains of the two genes have 58 to 74% homology. The EC domain containing 37 potential N-linked glycosylation sites is preceded by a 5' hydrophobic sequence resembling a typical signal peptide. An internal hydrophobic domain of 26 amino acids, the presumed TM domain, is followed by a spacer sequence of 58 amino acids, a PTK domain of 270 amino acids and a carboxyl tail of 84 amino acids. Overall, our result indicates that c-*ros* codes for a glycosylated TM PTK molecule which shares a remarkable sequence and structural homology with that of *Drosophila*

sevenless protein.

Introduction

Proto-oncogene c-*ros* is the cellular counterpart of the transforming gene v-*ros* of avian sarcoma virus UR2 (Wang et al., 1982; Neckameyer & Wang, 1985; Neckameyer et al., 1986). Previous sequence analysis of the v-*ros* DNA and partial genomic clones of the c-*ros* suggested that c-*ros* codes for a receptor-like PTK molecule (Neckameyer et al., 1986; Matsushime et al., 1986). Based on the comparison of the sequences in the PTK domains, c-*ros* was found to share the greatest homology with insulin receptor (Ebina et al., 1985; Ullrich et al., 1985), Insulin-like growth factor I receptor (Ullrich et al., 1986) and *Drosophila sevenless* (Basler & Hafen, 1988; Bowtell et al., 1988) among the known members of the PTK family.

Previous studies indicated that the expression of c-*ros* was highly regulated in chicken and was only detectable in kidney (Shibuya et al., 1982; Neckameyer et al., 1886) where a 8.3-kb mRNA was observed (Podell and Sefton, 1987). Most recently, a study of c-*ros* expression in rat, however, showed that it was expressed in lung, heart, and testis in addition to kidney and alternative splicing was observed among those tissues (Matsushime and Shibuya, 1990).

However, the physiological function and the potential ligand of the c-*ros* product remains unknown. To explore those questions, we have further studied the c-*ros* expression in chicken using riboprobe hybridization and have isolated c-*ros* cDNA clones and determined its entire coding sequence. Our results indicate that c-*ros* product shares remarkable sequence and structural homology with that of the *Drosophila sevenless* protein.

Results

Tissue-specific expression of c-*ros*

Previous analyses by liquid hybridization (Shibuya et al., 1982) and Northern analyses (Neckameyer et al., 1986) identified kidney as the sole organ in chicken that expresses a significant level of c-*ros* mRNA. Our previous kinetic study indicated that the level of c-*ros* mRNA in kidney peaked at 10 to 14 days post hatching and maintained a slightly lower, but constant level thereafter (Neckameyer et al., 1986; and our unpublished data). To further investigate the tissue-specific expression of c-*ros*, in the current study, we used slot blot hybridization with riboprobes to reinvestigate the question. Our result confirmed that kidney is the tissue having the highest level of c-*ros* mRNA (Fig. 3-1A). In addition, much lower but significant levels of RNAs were also observed in gonad, thymus, and bursa. The kidney mRNAs were analyzed by Northern blotting as well. A 8.3-kb mRNA species was reproducibly detected in different mRNA samples (Fig. 3-1B & C). Our result confirmed that reported previously by Podell and Sefton (1987). Substantial degradation of mRNA was observed in current and previous studies mentioned above. This is consistent with the presence of multiple mRNA instability motifs in the 3' c-*ros* mRNA (see below).

Isolation and restriction mapping of c-*ros* cDNA clones

Clone 84-1 representing the 3' 3-kb c-*ros* mRNA was isolated from the oligo(dT)-primed library. All the other 6 overlapping clones were isolated from the oligonucleotide-primed library. Their relationship was established first by restriction mapping (Fig. 3-2) and was further confirmed by sequencing. Except for clone 84-1, all the rest were rescued into phagemid vector pBluescript SK(-) for quantitative preparation and nucleotide sequencing. The total length of these overlapping cDNA clones is 8.12-kb.

Nucleotide sequence and predicted product of the c-*ros* cDNA

The nucleotide sequence (Fig. 3-3) covering the entire cDNA clones was derived by the sequencing strategy described in the materials and methods. All regions were sequenced in both directions. We predict from the sequence that c-*ros* codes for a polypeptide of 2311 amino acids. The sequence represents the only uninterrupted open reading frame spanning the bulk of the cDNA with the initiation codon conforming to the Kozak rule (Kozak, 1984). Two long stretches of hydrophobic region are present (Fig. 3-3 & 3-4). One is located at the N-terminus and the other is located internally. Each of them has an average hydrophobicity of greater than 1.5 (Fig. 3-4) and is presumed to function as signal peptide and transmembrane domain respectively. Excluding the presumed signal peptide, the unmodified c-*ros* protein is expected to have a molecular weight of 252,000 daltons. The predicted extracellular domain of 1873 aa contains 37 potential N-linked glycosylation sites composed of Asn-X-Ser/Thr (marked by open stars in Fig. 3-3) (Kornfeld & Kornfeld, 1985) and 34 cysteine residues (marked by closed stars in Fig. 3-3). Ten of the cysteine residues are within the N-terminal 100 amino acids. However, this cysteine rich region shares no homology with those of insulin receptor (Ebina et al., 1985; Ullrich et al., 1985), *sevenless* protein (Basler & Hafen, 1988; Bowtell, et al., 1988) or EGF receptor (Ullrich et al., 1984). Interestingly, a cysteine-rich domain is also present in the N-terminus of predicted products of rat and human c-*ros* cDNAs respectively (Matsushime & Shibuya, 1990; Birchmeier et al., 1990). No potential cleavage sites similar to those of insulin receptor (Ebina et al., 1985; Ullrich et al., 1985), IGFI receptor (Ullrich et al., 1986) or *sevenless* protein (Basler & Hafen, 1988; Bowtell et al., 1988) are present in the predicted extracellular domain of the c-*ros* product.

The intracellular domain consists of 412 aa including the PTK domain of about 270

aa and a carboxyl tail of 86 aa. The conserved PTK domain is about 60 aa downstream from the transmembrane domain.

Our cDNA clone contains a 3' non-coding sequence of 1013 nucleotides. Within this region there are three AATAAA motifs for polyadenylation and the third one used in polyadenylation in our cDNA clone is different from that of a previously reported cDNA clone (Podell & Sefton, 1987). These results indicate that there exists 3' heterogeneity among the c-*ros* mRNAs due to choice of polyadenylation sites. In addition, there are seven $A(U)_nA$ mRNA instability motifs that may be involved in post-transcriptional regulation of gene expression (Shaw and Kamen, 1986) within this region. This is consistent with the observed lability of chicken c-*ros* mRNA (Fig. 3-1; Neckameyer et al., 1986; Podell & Sefton, 1987). Interestingly, such motifs are not present in the reported 3' non-coding sequence of the rat c-*ros* cDNA (Matsushime & Shibuya, 1990). A Comparison of our cDNA sequence with the sequence of a 3' 2-kb cDNA clone reported previously (Podell & Sefton, 1987) indicates that there exist 8 differences in the 3' non-coding tail. We have individually confirmed the sequences at positions of difference in our cDNA clone. The reason for those differences are not clear. In addition, the penultimate, instead of the same, polyadenylation site shown in the previous cDNA clone (Podell & Sefton, 1987) is used for polyadenylation in our cDNA clone.

Comparison of the c-*ros* protein with other PTKs

Comparison of the amino acid sequence of the c-*ros* with those of known PTKs revealed that *sevenless* has the highest homology (data not shown). Fig. 3-5 shows the aa sequence homology among chicken c-*ros*, rat c-*ros* and *sevenless* proteins. As expected, the PTK domains of the three proteins share high degree of homology (75% identity and 88% homology if conserved aa are included). In addition, the *ros* proteins

also share considerable homology with the *sevenless* protein in their extracellular domains (20% identity and 43% homology if conserved aa are included). By contrast, very little homology is present in their carboxyl tails and the rat c-*ros* carboxyl tail is 50 aa longer than that of the chicken c-*ros*.

Comparison of the sequences in catalytic domain of c-*ros* product with those of various PTKs reveals certain features: 1) the cytoplasmic domain of c-*ros* product does not contain the SH2 and SH3 regions shared by PTKs of *src* family and a number of non PTK molecules (Sadowski & Pawson, 1986); 2) as pointed out previously, the PTK domain of v-*ros* protein contains a unique 6 aa insertion in comparison with a series of cytoplasmic and receptor PTKs (Neckameyer & Wang, 1985). This feature of 5 to 6 aa insertion in the PTK domain of *ros* proteins is shared with the *sevenless* protein, but not with IR and IGFR proteins (Fig. 3-6). On the other hand, the *ros* proteins lack the extended insertion sequence, called KI, interrupting the kinase domain, a characteristic shared by c-*fms*, PDGF receptor and c-*kit* proteins (Hampe, et al., 1984; Woolford et al., 1988; Doollitle et al., 1983; Besmer, 1986); 3) the presumed major autophosphorylation sites of *ros* proteins contain three instead of one tyrosine residues (Fig. 3-6). This characteristic is shared with *sevenless*, IR and IGFR proteins.

Discussion

Our predicted chicken c-*ros* product shares a remarkable sequence and structural homology throughout the entire molecule with the *Drosophila sevenless* protein. In addition, the sequence and several structural features in the cytoplasmic domain, particularly the PTK region, are also very similar to those of insulin receptor and IGFI receptor.

Both the predicted *ros* product and the *sevenless* protein have extremely large extracellular domains with numerous potential glycosylation sites. However, the spatial distribution of cysteine residues is different between the two proteins except that both have a cysteine-rich region in their N-termini. Interestingly the punctuation of several exons in the PTK domains, where information on genomic clones from *sevenless* and *ros* are available (Neckameyer et al. 1986; Matsushime et al., 1986; Matsushime and Shibuya, 1990), is identical between them.

In the cytoplasmic domain certain structural features are shared among *ros*, *sevenless*, IR and IGFR proteins. They include the length of the spacer between the TM and PTK domains, the lack of SH2, SH3 domains and large kinase insertions, as well as the presence of multiple tyrosine residues at the presumed major autophosphorylation sites. By contrast, the characteristic five to six aa insertion in the PTK domain of *ros* is only shared with *sevenless*, but not with IR or IGFR. These observations suggest that *ros*, *sevenless*, IR and IGFR are closely related members of the receptor PTK gene family, and particularly, *ros* and *sevenless* appear to share an even closer homology. It remains to be seen whether *sevenless* represents the *Drosophila* counterpart of the chicken and mammalian c-*ros* genes.

A typical signal peptide sequence is present in the N-terminal region of *ros* and

sevenless protein respectively. The signal peptide in *sevenless* is about 50 aa downstream from the N-terminus (Basler & Hafen, 1988; Bowtell et al., 1988) and instead of being cleaved, the sequence was proposed to insert into the membrane of R7 or its neighboring photoreceptor cells of *Drosophila* (Basler & Hafen, 1988; Simon et al., 1989). The signal peptide of *ros* proteins by contrast is located at the N-terminus and is likely to be cleaved at the early stage of protein synthesis and processing. Moreover, the *sevenless* protein was shown to be initially synthesized as a 280-kDa glycoprotein, followed by cleavage into a 220-kDa extracellular and a 58-kDa intracellular subunits, and the 58-kDa subunit was further processed into 48- or 49-kDa protein with simultaneous degradation of the rest of the protein (Simon et al., 1989). No sequences corresponding to the cleavage signal sequences in *sevenless*, IR or IGFR (Simon et al., 1989; Ebina et al., 1985; Ullrich et al., 1985; Ullrich et al., 1986) are present in the *ros* product. Initial characterization of the human c-*ros* protein in glioblastoma cell did not reveal any processing of the primary 280-kDa product (Sharma et al., 1989). However, further analysis of the *ros* protein in normal cells is needed to clarify the differences.

The function of the *sevenless* gene has been well established genetically and shown to be required for the development of the photoreceptor cell R7 of *Drosophila* ommatidia (Rubin, 1989). Despite the fact that *sevenless* is expressed in many *Drosophila* tissues including photoreceptor cells other than R7, its expression in R7 cells is crucial for their differentiation into photoreceptor cells. Recently, it was shown that expression of another gene, called *boss*, in the R8 cell was required for its neighboring cell to become R7 (Reinke & Zipursky, 1988). The *boss* gene product was thus suggested to be the ligand of *sevenless* protein. The potential ligand of the c-*ros* protein is unknown and whether there exists a similar signal transducing pathway for c-*ros*

remains to be seen. However, repeated efforts to detect expression of c-*ros* in chicken eye tissues have so far yielded negative results (Neckameyer et al., 1986; and this study). Instead, c-*ros* has been shown to be expressed in kidney, gonad, thymus and bursa of chicken (Fig. 3-1) and lung, heart, kidney and testis of rat (Matsushime & Shibuya, 1990). This discrepancy may be due to functional divergence of the gene in *Drosophila* versus higher animals. Alternatively, c-*ros* may be closely related to, but not the counterpart of, the gene *sevenless*.

Distinct cDNAs were isolated from various tissues of rat, suggesting that alternative splicing may yield different forms of c-*ros* proteins(Matsushime & Shibuya, 1990). Our cDNA sequence corresponds to that isolated from rat heart which contain a 21 aa insertion in comparison with that from lung and kidney (Matsushime & Shibuya, 1990). The 21 aa insertion in our kidney c-*ros* cDNA (flanked by the closed triangles between position 431 and 451 in Fig. 3-3) is present in all three different clones that we have isolated (2a, 1c & 5b). In addition to alternative splicing, multiple polyadenylation sites appear to be used for 3' end processing as mentioned above. An intriguing difference between chicken and rat c-*ros* mRNAs is the apparent absence of mRNA instability motifs and multiple AAUAAA polyadenylation signals in the latter c-*ros* mRNA. The presence of instability motifs may account for the observed high tendency of degradation of the chicken c-*ros* mRNA. But it may not fully explain the low abundance of c-*ros* mRNA in various chicken tissues. The significance for those differences between chicken and rat c-*ros* mRNAs is unclear.

Acknowledgement

We thank Song-Muh Jong and Thambi Dorai for their helpful suggestions, and David Clayton for the gift of a chicken actin DNA plasmid. This work was supported by NIH grant CA29339. DH was supported by a fellowship from American Leukemia Society.

Fig. 3-1 Expression of c-*ros* in chicken tissues. (A) 10 μg of polyadenylated RNA from 10-14 day old chickens were analyzed by slot blotting and hybridization with v-*ros* riboprobe (A and B) or a chicken actin DNA probe (C). A and B were treated with RNase subsequent to hybridization. 1 μg of polyadenylated RNA from uninfected or UR2 transformed CEF and 1 μg of yeast tRNA were included as controls. Lane A and B are the same strip with different times of exposure. Lane A is a 14 h exposure, lane B is a 20 h exposure. The origins of mRNAs are: B, brain; R, retina; G, gonad; K, kidney; L, liver; M, muscle; BM, bone marrow; Bu, bursa; S, spleen; T, thymus. (B & C) Two separate Northern analysis of kidney polyadenylated RNAs. 8 μgs of mRNA was used and the filters were probed with a ^{32}P-labeled v-*ros* DNA.

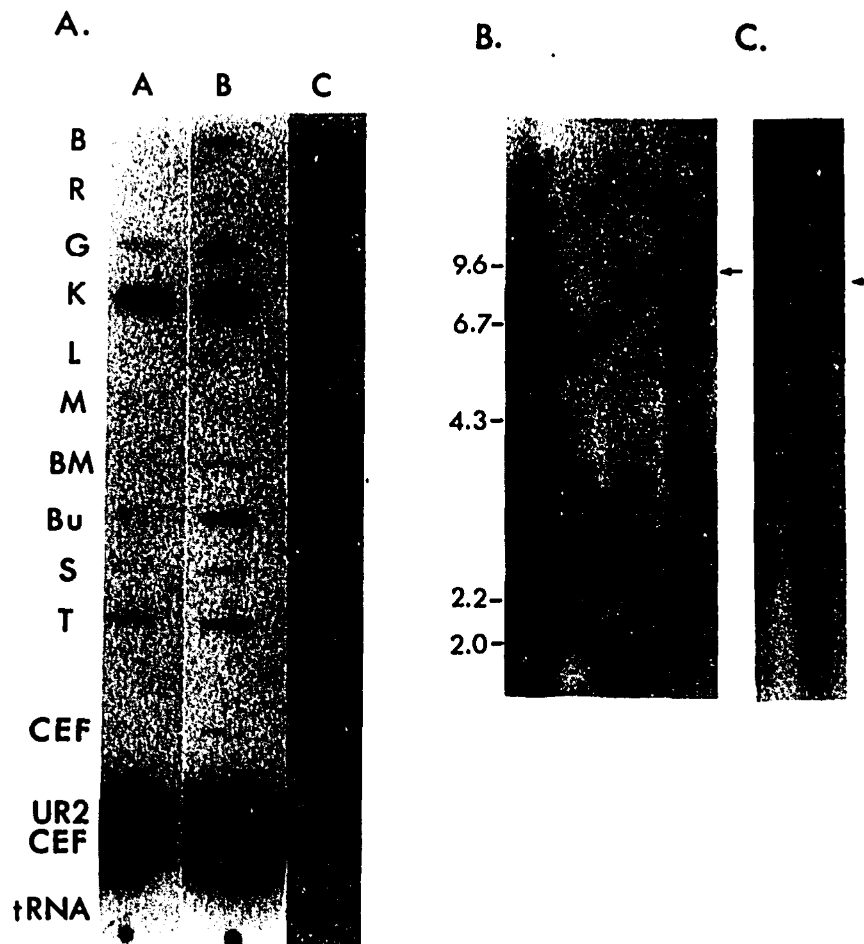


Fig. 3-2 Isolation of c-*ros* cDNA clones. The open box depicts the derived structure of the c-*ros* product. The two blackened vertical bars represent the two hydrophobic sequences. EC stands for extracellular domain. TPK stands for protein tyrosine kinase domain. Several overlapping clones were isolated from two 14-day chicken kidney cDNA libraries (see materials and methods). Abbreviations for restriction enzyme sites are: P, PstI; E, EcoRI; S, SacI; H, HindIII; X, XbaI.

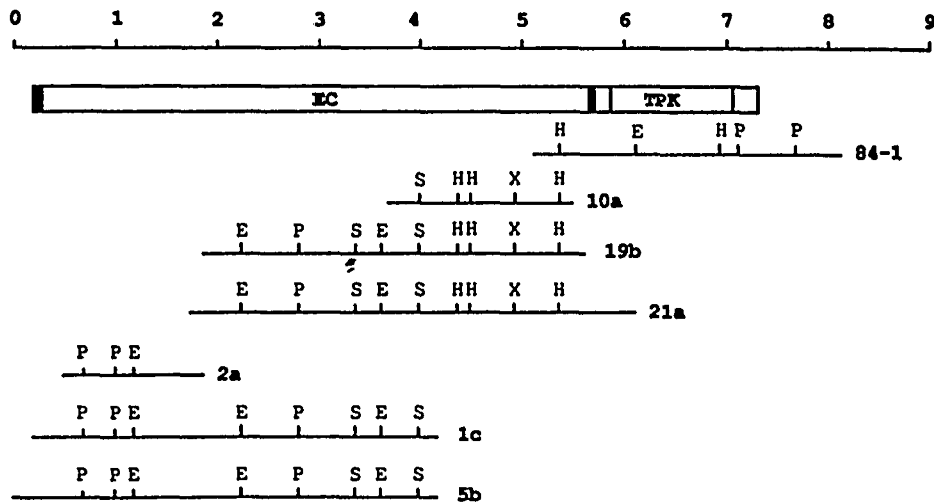


Fig. 3-3 Nucleotide and derived amino acid sequences of the c-*ros* cDNA. The sequence was constructed from sequencing data of the cDNA clones shown in Fig. 3-2. The closed and open stars denote cysteine resides and potential N-linked glycosylation sites, respectively, in the extracellular (EC) domain. Two stretches of hydrophobic sequences are highlighted by closed diamonds above them. The two closed triangles in the EC domain define a 21 aa sequence corresponding to an identical insertion in rat heart c-*ros* cDNA, but absent in rat lung cDNA. The closed circles in the intracellular domain denote the sequence motifs for ATP binding. The asterisks point to the presumed major autophosphorylation tyrosine residues. The 6 aa insertion is highlighted by an overline. The mRNA instability motifs $A(U)_nA$ as well as the polyadenylation addition signal AATAAA in the 3' non-coding region are underlined.

```
                                                                                  AGGTTGTTCAAGGAATGGATTGTAAACAAAGAATTAAA   38
AATATTTAATAAATACTAAAGTAGAAATTTCGTAGTCGTGTCTCTGTCACCTAGTCTTGAAGGAGTTTGTATGGCCAAGCCAAGGAAGGAATCAGACTGAAAGATGAAATTTCTCAACTT  158

MetArgAsnAlaCysLeuLeuLeuAsnArgLeuGlyAlaPheTyrPheIleTrpIleSerAlaAlaTyrCysSerPheSerLysAsnCysGlnAspLeuCysThrSerAsnLeuGluGly   40
ATGAGGAACGCTTGCCTGCTGCTGAACAGACTTGGTGCTTTTTACTTCATTTGGATTTCTGCTGCATATTGTTCTTTCTCAAAAAATTGTCAAGACCTATGTACAAGTAACCTGGAAGGA  278

GluLeuGlyIleAlaAsnLeuCysAsnValSerAspIleAsnValAlaCysThrGlnGlyCysGlnPheTrpAsnAlaThrGluGlnValAsnCysProLeuLysCysAsnLysThrTyr   80
GAACTTGGAATTGCTAACCTTTGCAATGTTAGTGACATCAATGTGGCATGTACACAGGCATGTCAGTTTTGGAATGCAACAGAGCAAGTCAATTGCCCACTGAAATGTAACAAGACATAC  398

ThrArgGluCysGluThrValSerCysLysPheGlyCysSerArgAlaGluAspAlaTyrGlyValGluAlaGlnAsnCysLeuAsnLysProGlyAlaProPheAlaSerSerIleGly  120
ACCAGAGAATGTGAGACAGTTTCCTGCAAGTTTGGCTGTAGTCGTGCAGAGGATGCATATGGTGTTGAAGCACAGAACTGTTTGAACAAGCCTGGAGCACCATTTGCATCCTCCATTGGA  518

SerHisAsnIleThrLeuGlyTrpLysProAlaAsnIleSerGluValLysTyrIleIleGlnTrpLysPheHisGlnLeuProGlyAspTrpArgTyrThrGluValValSerGluThr  160
AGCCACAATATCACACTGGGTTGGAAGCCAGCTAATATCTCTGAGGTTAAATATATTATCCAGTGGAAGTTTCATCAGCTCCCTGGAGACTGGAGATACACAGAGGTCGTATCTGAAACT  638

SerTyrThrValLysAspLeuGlnAlaPheThrGluTyrGluPheArgValValTrpIleIleThrSerGlnLeuGlnLeuHisSerProProSerProSerTyrArgThrHisAlaSer  200
TCATATACAGTCAAGGACCTTCAAGCCTTCACAGAGTATGAGTTTCGAGTAGTTTGGATCATAACCTCTCAGCTGCAGCTCCATTCTCCACCAAGTCCCAGTTACCGGACCCATGCTTCT  758

GlyValProThrThrAlaProIleIleLysAspIleGlnSerSerSerProAsnThrValGluValSerTrpPheProProLeuPheProAsnGlyLeuIleValGlyTyrAsnLeuVal  240
GGAGTTCCTACTACAGCTCCCATCATTAAAGATATTCAGAGTTCAAGTCCAAACACTGTTGAAGTTAGTTGGTTTCCACCACTTTTTCCCAATGGATTGATTGTTGGATACAACTTGGTC  878

LeuThrSerGluAsnHisGluLeuLeuArgAlaSerArgGlyHisSerPheGlnPheTyrSerThrPheProAsnSerThrTyrArgPheSerIleValAlaValAsnGluAlaGlyAla  280
CTGACCAGTGAGAATCATGAACTGTTGAGAGCATCGAGAGGGCACAGCTTTCAGTTCTACTCTACCTTCCCAAACAGCACTTACAGGTTCAGTATTGTAGCTGTTAATGAGGCTGGAGCT  998

GlyProProAlaGluAlaAsnIleThrThrProGluSerLysValLysGluLysAlaLysTrpLeuPheLeuSerArgAsnGlnSerLeuArgLysArgTyrMetGluHisPheLeuGlu  320
GGACCCCCTGCAGAAGCCAACATCACAACACCGGAATCCAAAGTTAAGGAAAAAGCAAAATGGCTCTTTCTATCCAGAAACCAGTCTTTAAGGAAAAGATACATGGAACATTTCCTTGAA 1118

AlaAlaHisCysLeuGlnAsnGlyIleIleHisHisAsnIleThrGlyIleSerValAsnValTyrGlnGlnValValTyrPheSerGluGlyAsnSerIleTrpValLysGlyValVal  360
GCAGCCCACTGTCTGCAAAATGGTATAATACACCACAACATTACAGGAATTTCTGTGAATGTTTATCAGCAAGTTGTCTATTTTTCTGAGGGGAATTCCATCTGGGTGAAAGGAGTTGTG 1238

AspMetSerAspValSerAspLeuThrLeuPheTyrThrGlyTrpGlyAsnIleThrSerIleSerValAspTrpLeuTyrGlnArgMetTyrPheValMetAsnGluLysIleHisVal  400
GATATGTCTGATGTATCTGACTTAACACTCTTTTATACTGGCTGGGGGAACATTACATCAATCTCAGTAGATTGGCTTTACCAGAGGATGTACTTTGTCATGAATGAAAAGATACATGTT 1358

CysGlnLeuGluAsnCysThrAlaAlaGluAspIleThrProProTyrGluThrSerProArgLysIleValAlaAspProTyrAsnGlyTyrIlePheCysLeuLeuGluAspGlyIle  440
TGTCAGTTAGAGAACTGCACAGCAGCTGAAGACATCACTCCTCCTTATGAGACATCTCCTAGGAAGATTGTAGCTGATCCCTATAATGGGTATATTTTCTGTCTTCTGGAGGATGGTATA 1478

TyrArgAlaAsnLeuProLeuPheProAspThrAlaSerAlaAlaSerLeuValValLysSerHisThrLeuArgAspPheMetIleAsnPheGlnSerLysArgLeuIlePhePheAsn  480
TACAGAGCAAACCTTCCACTCTTTCCTGACACTGCCTCAGCAGCTTCACTAGTTGTGAAATCCCACACTTTGAGAGACTTCATGATCAACTTCCAGTCTAAGAGACTCATTTTCTTCAAC 1598

LysThrGluGlnAlaPheValSerGlyPheLeuAspGlySerGluPheHisThrLeuArgAlaHisValProLeuAspAspMetGluSerPheValTyrGluAspAsnIlePheThrVal  520
AAAACAGAACAGGCCTTTGTGTCGGGTTTTCTTGATGGTTCAGAATTTCACACTCTGCGAGCACACGTGCCACTTGATGACATGGAGAGCTTTGTTTATGAGGATAACATATTTACTGTC 1718

ThrAspGlyArgAlaValPheHisGluGluIleSerGlnValGlySerSerSerPheAsnGluTyrValValAspCysSerLeuGluTyrProGluTyrPheGlyPheGlyAsnLeuLeu  560
ACAGATGGCCGGGCAGTTTTCCATGAGGAGATCTCACAAGTGGGAAGTTCTAGCTTCAACGAGTATGTTGTAGATTGCAGTTTGGAATATCCAGAGTATTTTGGCTTTGGCAATTTGCTT 1838

PheTyrAlaAlaSerThrGlnProTyrProLeuProThrLeuProArgLeuValThrValLeuPheGlySerAspGlnAlaValIleSerTrpSerProProGluTyrThrIleGlyThr  600
TTCTATGCAGCATCAACTCAGCCTTATCCTTTACCAACTCTTCCTCGGCTTGTGACAGTACTCTTTGGATCGGATCAAGCTGTGATCAGCTGGAGTCCACCTGAATACACAATTGGAACC 1958

SerArgSerAlaTrpGlnAsnTrpThrTyrAspValLysValSerSerGlnSerThrPheGluGluGluTrpValValSerAsnIleThrAspThrArgPheAlaValLysAsnLeuVal  640
AGTCGGTCTGCTTGGCAGAACTGGACCTATGATGTGAAAGTGTCATCTCAAAGTACTTTTGAGGAGGAATGGGTTGTCTCAAACATTACTGACACCAGGTTTGCCGTGAAGAATCTGGTC 2078

SerPheThrGluTyrGluMetSerValArgAlaValSerProAlaGlyGluGlyProTrpSerGluProPheArgGlyMetThrPheGluGluAlaGluGluGluProTyrIleLeuAla  680
AGCTTCACAGAGTATGAGATGTCAGTCAGAGCCGTGTCTCCTGCTGGTGAAGGACCATGGAGCGAGCCATTTAGAGGCATGACATTTGAAGAAGCTGAAGAAGAACCTTACATATTGGCA 2198

ValGlyAlaGluGlyLeuTrpLysGlnArgLeuAspSerTyrGlyProGlyGluPheLeuTyrProHisIleArgAsnIleSerAspLeuAspTrpTyrAsnAspThrLeuTyrTrpSer  720
GTAGGGGCGGAAGGGCTCTGGAAGCAGCGTCTGGATAGCTACGGGCCAGGGGAATTCCTTTATCCTCATATTAGAAATATTTCAGACCTCGACTGGTATAATGACACTCTGTATTGGAGT 2318

AsnSerMetGlyLysValGlnThrTrpSerMetAsnLysLysGluGlyThrThrGluAsnSerTyrValProAspIleLysAsnAlaArgMetLeuAlaPheAspTrpLeuGlyGlnCys  760
AATTCCATGGGGAAAGTTCAGACCTGGTCAATGAATAAGAAAGAGGGTACCACTGAGAACAGTTATGTACCTGACATCAAAAATGCAAGAATGTTGGCCTTTGACTGGCTGGGACAATGC 2438

LeuTyrTrpAlaGlyLysAlaAsnThrIleTyrArgLysSerLeuLeuGlyAspHisMetAspValValAlaHisValValTyrValValLysAspLeuAlaValAspSerValAsnGly  800
TTGTATTGGGCTGGCAAAGCAAATACGATCTACAGAAAATCACTGCTAGGAGACCATATGGATGTTGTGGCCCATGTTGTATATGTAGTGAAGGATCTAGCGGTGGATTCAGTGAATGGC 2558

TyrLeuTyrTrpAlaThrThrTyrThrValGluSerAlaArgLeuAsnGlyGluGluTyrLeuIleLeuGlnGluHisLeuGlnPheSerGlyLysGlnValValGlyLeuAlaLeuAsp  840
TATTTGTATTGGGCCACAACGTATACAGTGGAGAGTGCAAGACTGAATGGAGAAGAGTACCTTATATTACAGGAGCATCTACAGTTTTCTGGAAAACAGGTGGTTGGTTTAGCTTTGGAT 2678

LeuThrSerGlyPheLeuTyrTrpLeuValGlnAspGlyLeuCysLeuAsnLeuTyrArgIleSerIleCysLysGluSerCysGlyAsnIleMetValThrGluIleSerAlaTrpSer  880
CTCACCTCTGGTTTTCTTTACTGGCTTGTGCAAGATGGTCTATGCCTAAACCTATACAGAATTTCTATTTGCAAAGAAAGTTGTGGTAATATCATGGTCACAGAAATCTCAGCATGGAGT 2798

ValSerGluValSerGlnAsnAlaLeuGlnTyrTyrSerGlyArgLeuPheTrpIleAsnArgLeuLysPheIleThrThrGlnGluLeuAsnGlnSerIleSerIleProPheSerGlu  920
GTTTCGGAAGTATCTCAGAATGCACTGCAGTATTACAGTGGTCGTCTGTTCTGGATTAATAGGCTTAAATTCATTACAACTCAGGAACTCAACCAGAGCATTAGCATTCCCTTTTCAGAA 2918

ProAlaGluPheAlaAlaPheThrLeuValHisThrSerLeuLysProLeuProGlyAsnPheSerPheThrProLysValIleProSerSerValProGluSerSerPheLysIleLys  960
CCAGCTGAGTTTGCAGCCTTTACCCTTGTTCACACATCACTTAAACCTTTGCCAGGAAATTTCTCTTTTACACCAAAAGTGATTCCAAGTTCGGTGCCAGAATCTTCTTTCAAGATTAAA 3038

SerSerGlyAsnSerPheHisIleIleTrpAsnAlaSerThrAspValGlnTrpGlyThrValPheTyrCysValGlySerAsnAlaLeuGlnMetArgThrLeuGluSerGluArgCys 1000
GGAAATTCTTCAAGTTTTCACATCATTTGGAATGCCTCCACAGATGTGCAATGGGGAACTGTCTTTTACTGCGTGGGATCAAATGCTCTACAAATGAGGACTTTGGAGAGTGAACGGTGC 3158

LeuHisProHisAspLeuThrValProSerTyrLysValAspTrpLeuGluProPheThrLeuPheAspPheSerValThrProTyrThrTyrTrpGlyLysAlaProThrThrSerVal 1040
CTCCATCCGCATGACCTCACAGTTCCATCCTACAAAGTTGATTGGCTGGAACCATTCACACTCTTTGATTTTAGTGTCACTCCGTACACATACTGGGGCAAGGCACCAACAACATCTGTG 3278

TyrLeuArgAlaProGluGlyValProSerAlaProAlaAsnProArgIleTyrValLeuHisSerAsnThrHisGluGlyGluLeuLysValLeuValGluLeuArgTrpAspLysPro 1080
TATCTTCGAGCCCCTGAAGGAGTTCCTTCAGCCCCTGCAAACCCTAGAATATATGTGTTGCATAGTAACACACATGAAGGTGAAGAGAAAGTCCTTGTGGAGCTCAGATGGGACAAACCT 3398

GluArgAspAsnGlyValLeuThrGlnPheArgValTyrTyrGlnLeuLeuTyrGluSerGlyAlaAlaAspThrLeuMetGluTrpAsnValSerAspValLysProThrAlaLeuLeu 1120
GAAAGGGACAATGGAGTATTAACACAATTCAGGGTTTACTATCAACTGTTGTATGAGAGTGGTGCTGCTGACACTCTCATGGAGTGGAATGTGAGCGATGTCAAACCCACTGCATTGCTC 3518

PheSerIleArgAspGluHisProArgLeuThrValArgPheGlnValGlnAlaPheThrSerValGlyProGlyProMetSerAspValAlaGlnArgAsnSerSerAspIlePhePro 1160
TTTTCAATCAGAGATGAGCATCCTAGGCTCACAGTAAGGTTTCAGGTGCAAGCTTTTACATCTGTGGGTCCAGGACCTATGTCTGATGTAGCACAGAGGAATTCATCAGATATTTTCCCT 3638

ValProThrLeuIleThrPheSerSerAsnLysLeuPheLeuThrAspIleAspSerAsnHisThrIleTrpGluValLeuThrAsnArgAsnIleLysAspIleCysTyrThrAlaAsp 1200
GTCCCAACTCTTATAACTTTTTCTTCCAACAAGTTGTTTCTTACTGACATAGACAGTAATCATACCATATGGGAAGTTTTGACAAATAGAAATATAAAGGACATCTGTTATACTGCTGAT 3758
```

```
AspAspLysValTyrTyrIleLeuGluAspSerLeuPheLeuLeuAsnValGlnSerThrSerGluSerGlnLeuPheGluAspValPheLeuArgAsnValThrAlaIleThrValAsp 1240
GATGACAAAGTGTATTATATCCTAGAAGACTCTCTTTTTCTTCTGAATGTACAGAGTACTTCAGAGTCTCAGTTGTTCGAAGATGTATTTCTTCGCAATGTCACAGCCATCACAGTAGAT 3878

TrpIleAlaArgHisLeuPheValAlaMetLysThrSerTrpAsnGluThrGlnValPhePheIleAspLeuGluLeuLysThrLysSerLeuLysAlaLeuAsnIleGlnLeuGlyLys 1280
TGGATTGCAAGACATCTCTTTGTTGCCATGAAAACGTCATGGAATGAAACTCAAGTATTTTTTATTGATCTGGAGCTCAAGACAAAGTCCTTAAAAGCATTGAACATACAGTTGGGTAAA 3998

ArgAsnSerThrIleSerSerLeuLeuSerTyrProPheLeuSerArgLeuTyrTrpIleGluGluLeuAspTyrGlySerArgMetPheTyrTyrAspIleLeuAsnAsnThrMetTyr 1320
CGTAATTCAACTATATCATCTCTACTGTCATATCCCTTTTTAAGCCGCTTGTATTGGATTGAAGAGCTTGATTATGGCAGCCGCATGTTTTATTATGATATACTGAATAATACCATGTAC 4118

HisIleLeuGlyTyrGluSerValGluGluLysMetArgAsnTyrCysAsnCysAsnValAlaGluAlaGluLeuGlyArgProIleSerIleAspValThrAspIleLysLysProGln 1360
CACATTTTGGGATACGAATCAGTAGAAGAAAAGATGAGGAACTACTGCAACTGTAACGTGGCAGAAGCAGAGCTAGGAAGACCTATAAGTATAGATGTAACTGACATCAAGAAACCCCAA 4238

LeuLeuPheIleArgGlyArgAspGluIleTrpAlaSerAspValAspGlyCysHisCysTrpArgIleThrLysIleProSerPheGlnGlyThrLysIleGlySerLeuThrValAsp 1400
CTGCTCTTTATCAGAGGAAGAGATGAAATATGGGCCTCAGATGTGGATGGTTGCCACTGCTGGAGGATTACCAAGATACCAAGTTTTCAAGGTACAAAAATTGGAAGCTTGACTGTAGAT 4358

LysGlnPheIleTyrTrpThrIleGluLysLysGluTyrThrGluIleCysLeuAlaAspLysGluSerThrArgHisSerLeuGlnArgLysAlaAsnHisGluLeuLysIleLeuAla 1440
AAACAATTTATATATTGGACCATTGAAAAGAAGGAATACACAGAAATTTGTCTAGCAGATAAAGAAAGCACAAGACACTCTCTTCAGAGGAAAGCAAATCATGAGCTGAAAATCTTAGCC 4478

TyrSerSerAlaMetGlnSerTyrProAspLysLysCysLeuThrProLeuLeuAspThrGluLysProThrIleLeuAspThrThrAsnThrSerPheThrLeuSerLeuProSerVal 1480
TACAGCTCAGCAATGCAATCATATCCAGATAAGAAGTGCCTGACTCCATTACTAGACACTGAGAAACCTACTATCCTCGATACAACAAACACTAGTTTTACGTTAAGCTTACCATCTGTC 4598

ThrThrGlnGlnLeuCysProSerIleSerGlnProThrProThrTyrLeuValPhePheArgGluIleThrSerAsnHisGluAsnSerThrTyrHisPheSerThrLeuLeuGlnLys 1520
ACAACACAGCAGCTATGCCCCAGCATATCCCAGCCTACACCAACATATCTGGTATTTTTCAGAGAAATAACTAGCAACCATGAGAACTCGACCTACCATTTCTCTACTCTACTGCAAAAA 4718

ThrLeuGluIleGlnGluProIleAlaValIleAsnAsnLeuLysProPheSerThrTyrAlaIleGlnValAlaValLysAsnTyrTyrSerAsnGlnAsnGlnLeuAlaValGlyArg 1560
ACTTTGGAAATTCAGGAACCTATTGCTGTCATAAACAACCTAAAACCATTTTCAACGTACGCCATACAAGTTGCTGTGAAAAACTACTATTCAAATCAAAACCAGCTGGCTGTAGGAACG 4838

GluAlaIleSerThrThrLeuTyrGlyValProGluGlyValAspSerIleLysThrValValLeuSerAspThrThrIleAsnIleSerTrpSerGluProLeuGluProAsnGlyPro 1600
GAGGCAATCAGCACAACTCTATATGGAGTACCAGAGGGAGTTGATTCCATTAAGACAGTGGTTTTGTCCGACACTACTATCAACATATCTTGGAGTGAACCTCTGGAGCCAAATGGACCT 4958

LeuGluSerIleArgTyrGlnIleSerValAsnLeuLeuSerLeuPheProGluAlaProLeuArgLysSerGluPheProAsnGlyThrLeuSerTrpSerValSerAspLeuGlnSer 1640
CTAGAATCCATCCGCTATCAAATCTCTGTCAATTTATTGTCTTTGTTCCCAGAAGCTCCCTTGCGAAAAAGTGAATTTCCAAATGGGACGCTTTCATGGTCTGTCAGTGATCTCCAGTCT 5078

GlyThrAsnAsnLeuPheLysValLeuAlaPheHisProAsnGluAsnTrpPheSerGluSerValProValIleAlaLysThrPheGluThrProLeuSerProSerAsnIleIlePro 1680
GGAACAAACAATTTGTTCAAGGTTCTTGCTTTTCATCCCAATGAGAACTGGTTTTCTGAAAGTGTCCCTGTTATTGCAAAAACTTTTGAGACTCCACTTTCACCTTCCAATATAATCCCC 5198

ArgAsnThrSerPheGlnLeuGluTrpArgAlaProLeuHisIleAsnGlyThrSerPheTrpPheGluLeuSerLysTrpGlnThrArgSerAspTrpPheSerProAlaSerThrThr 1720
AGAAATACTAGTTTCCAGCTAGAATGGAGAGCTCCTCTGCACATCAATGGAACCAGCTTCTGGTTTGAGCTGAGTAAGTGGCAAACAAGAAGTGACTGGTTTTCACCTGCAAGCACTACA 5318

CysThrValGlyProValTyrThrCysAsnLeuThrGlyThrLeuProSerAlaAsnTyrLeuValArgAlaThrValValTyrValThrGlyMetLysSerThrSerSerProThrSer 1760
TGCACTGTAGGCCCTGTTTACACATGCAATCTCACTGGAACTCTTCCTTCTGCCAATTATCTTGTAAGAGCAACAGTAGTATATGTTACAGGAATGAAAAGCACTTCCTCTCCAACAAGC 5438

PheLysThrThrAlaGlyValProSerLysProGlyThrProLysArgAlaGluAspSerLysAsnSerValGlnTrpGluLysAlaGluAspAsnGlySerAsnLeuThrTyrTyrIle 1800
TTTAAAACTACAGCTGGTGTTCCCAGTAAGCCAGGAACACCCAAAAGAGCAGAAGACTCTAAAAACTCTGTACAGTGGGAAAAGGCTGAAGACAATGGAAGCAACCTCACCTACTACATC 5558

LeuGluSerArgLysGlnSerGlyAsnThrAsnLysValLysSerLeuTrpValValValTyrAsnGlySerCysAspAsnIleCysThrTrpLysAlaGluAsnLeuGluGlyThrPhe 1840
TTAGAAAGCAGAAAGCAGTCTGGCAACACCAACAAAGTGAAGTCCTTATGGGTGGTGGTTTACAATGGTTCTTGTGACAACATTTGCACATGGAAGGCCGAGAACTTAGAAGGAACTTTT 5678

GlnPheArgAlaAlaAlaAlaAsnMetLeuGlyLeuGlyGluTyrSerAspThrSerLysAspIleValLeuAlaLysAspThrValThrSerProAspIleThrAlaIleValAlaVal 1880
CAGTTCAGAGCAGCAGCTGCAAATATGCTGGGGCTTGGTGAATACAGTGACACAAGCAAAGATATTGTGTTGGCTAAAGATACTGTGACCTCTCCAGATATCACTGCTATTGTTGCTGTG 5798

IleGlyAlaValValLeuGlyLeuThrIleIleIleLeuPheGlyPheValTrpHisGlnArgTrpLysSerArgLysProAlaSerThrGlyGlnIleValLeuValLysGluAspLys 1920
ATTGGAGCAGTTGTACTGGGCTTGACTATAATCATACTGTTTGGTTTTGTATGGCACCAAAGATGGAAATCCAGAAAACCAGCCTCAACTGGGCAGATTGTGCTTGTCAAGGAAGATAAA 5918

GluLeuAlaGlnLeuArgGlyMetAlaGluThrValGlyLeuAlaAsnAlaCysTyrAlaValSerThrLeuProSerGlnAlaGluIleGluSerLeuProAlaPheProArgAspLys 1960
GAATTAGCTCAACTTAGGGGAATGGCTGAGACAGTGGGATTAGCCAATGCTTGTTATGCTGTCAGCACTCTTCCTTCTCAAGCAGAGATTGAGTCATTGCCAGCTTTTCCTCGGGACAAA 6038

LeuAsnLeuHisLysLeuLeuGlySerGlyAlaPheGlyGluValTyrGluGlyThrAlaLeuAspIleLeuAlaAspGlySerGlyGluSerArgValAlaValLysThrLeuLysArg 2000
CTGAACTTACACAAGTTGTTAGGAAGTGGAGCATTTGGAGAGGTGTATGAAGGGACTGCATTAGATATCCTGGCAGATGGAAGTGGAGAATCCAGAGTAGCAGTCAAGACTTTGAAGAGA 6158

GlyAlaThrAspGlnGluLysSerGluPheLeuLysGluAlaHisLeuMetSerLysPheAspHisProHisIleLeuLysLeuLeuGlyValCysLeuLeuAsnGluProGlnTyrLeu 2040
GGTGCAACAGACCAAGAGAAGAGTGAATTCTTGAAGGAGGCACACTTAATGAGTAAATTTGATCATCCCCACATTCTGAAGCTACTTGGAGTGTGTCTGTTAAATGAACCTCAGTACCTT 6278

IleLeuGluLeuMetGluGlyGlyAspLeuLeuSerTyrLeuArgGlyAlaArgLysGlnLysPheGlnSerProLeuLeuThrLeuThrAspLeuLeuAspIleCysLeuAspIleCys 2080
ATACTGGAGCTGATGGAAGGAGGAGATCTGCTTAGCTATTTACGAGGAGCCAGAAAGCAAAAGTTCCAGAGTCCCTTACTGACATTGACTGATCTCTTGGATATATGCTTGGATATTTGC 6398

LysGlyCysValTyrLeuGluLysMetArgPheIleHisArgAspLeuAlaAlaArgAsnCysLeuValSerGluLysGlnTyrGlySerCysSerArgValValLysIleGlyAspPhe 2120
AAAGGTTGTGTCTATTTAGAGAAAATGCGTTTCATACACAGGGACCTGGCTGCTCGCAACTGCCTTGTGTCTGAGAAGCAATATGGGAGCTGCTCCCGAGTGGTAAAGATTGGTGATTTT 6518

GlyLeuAlaArgAspIleTyrLysAsnAspTyrTyrArgLysArgGlyGluGlyLeuIleAsnValArgTrpMetAlaProGluSerLeuIleAspGlyValPheThrAsnHisSerAsp 2160
GGACTTGCCAGAGATATCTATAAAAATGATTACTACAGGAAAAGAGGAGAAGGCCTAATAAATGTCAGATGGATGGCTCCTGAAAGCCTCATTGATGGCGTCTTTACAAATCACTCTGAT 6638

ValTrpAlaPheGlyValLeuValTrpGluThrLeuThrLeuGlyGlnGlnProTyrProGlyLeuSerAsnIleGluValLeuHisHisValArgSerGlyGlyArgLeuGluSerPro 2200
GTTTGGGCTTTTGGAGTCTTAGTGTGGGAAACATTAACTTTGGGTCAACAGCCATATCCGGGTCTCTCCAATATAGAAGTTTTACACCATGTACGATCAGGAGGAAGGCTGGAATCTCCG 6758

AsnAsnCysProAspAspIleArgAspLeuMetThrArgCysTrpAlaGlnAspProHisAsnArgProThrPhePheTyrIleGlnHisLysLeuGlnGluIleArgHisSerProLeu 2240
AATAACTGTCCTGATGACATACGTGATTTAATGACAAGATGCTGGGCCCAAGATCCTCACAACAGACCTACTTTCTTTTATATTCAGCACAAACTGCAAGAGATAAGGCACTCTCCACTG 6878

CysPheSerTyrPheLeuGlyAspLysGluSerValAlaGlyPheIleAsnGlnAlaPheGluAspIleAspValProProAlaAspSerAspSerIleLeuSerThrThrLeuMetGlu 2280
TGCTTCAGCTACTTCCTTGGAGACAAAGAGTCAGTGGCTGGGTTCATCAACCAAGCTTTTGAGGATATTGATGTTCCACCTGCGGATTCAGACAGCATTCTTTCAACCACTCTAATGGAA 6998

AlaArgAspGlnGluGlyLeuAsnTyrLeuValValValLysGluSerAsnGlnAspGlnGlySerIleSerSerAlaGluLeuThrSerVal 2311
GCAAGGGATCAAGAAGGCCTGAATTATTTAGTAGTAGTCAAAGAAAGCAACCAGGATCAAGGTAGCATCAGCTCTGCAGAACTTACTTCAGTCTAATTCCATCTGTTTGAGGACAGAAGA 7118

GAGAACTCTAACAAATAAGACGTAAGACAGAGAAAATATTATGCACTTTATTATACATTTAAAGATTTCTTAAAAGAAACAACAACATTCTAAAACACCAGAACTTCCAGAGAATGTCAT 7238
AAGGTACTGAAACTTGTGTACAATAAGACAACAGTCATATATTTATGATTTTCTAAGTGCTAAGGTTCATGTTTTTTACAAGAAATAGTTTAGGGAAATAAAACTGGCCATCACACAGAA 7358
AATGATTTTACTGTGATCCCATTAACTTAATGCAGAGACGTAAACTTTGGACACAAGCAGTCCTTCCAAAAGAGAAGTTCTGTCATCCTGTCTCATAGAAGGGAACAACCTGATCATGAA 7478
AAGAAAAACAGTTCAAATATTGCTCCGTTGAGCACACATGAGATTATACAAATTGAAAAGCTAACAAATGCTAGTTGTATTTTTCTGATCTCCAAATCTGTAAGCAATTGAAGAAAGGTG 7598
GCTGGGTTCTTCACTGCAGGGATGGTCTATTTTATCATTTCATTTAAACACACACTTGGAATCATCTCAAATTTTAAGTTGCACAGTGATTGATCTGTGAATCTACTATCCACATTCCTT 7718
TTCCCTTTTAACATTGATCACCTTTCATTTCATAGTTGATTGGATAAAGGTAAATGTTTTTCCTTGATGTCCTTGGAGATCAAATCAAAACATATAATAAACAATATATAAATCATTCC 7838
TGAGTAAAAGGAAGAAATAAGCTCTTTGGAATGCTTCATCATGAATAGCAATTTGCATATTTTGCAGAGCTATTCACTGACGTCTGTCAGGAAGCTAGAAACGTATCCTGCTTCTATGCA 7958
AATTATTGATAGTGGCAAATACTGGAGAGTAGTAAAAATGCGTTGTTTAATGCTAAAATGTAGATATTTTTAAGGAACATTGCCACTGAGATCAAGCTCCAAAAGTTTCTTAAATGGAGA 8078
TTCAATAAAGATAAACAAGATGTTTGAAAAAAAAAAAAAAAA 8121
```

Fig. 3-4 Hydrophilicity plotting of the predicted c-*ros* protein. (A) Hydrophilicity profile of chicken c-*ros* gene product. IBI Macvector Hydrophilicity program was used to generate the figures in A, B and C. Window sizes: Hydrophilicity, 7; Amphiphilicity, 11; Hydrophilicity scale, Kyte-Doolittle. (B) Blow-up of the 5' sequence from residue 3 to 27. (C) Blow-up of the internal hydrophobic sequence from residue 1874 to 1897.

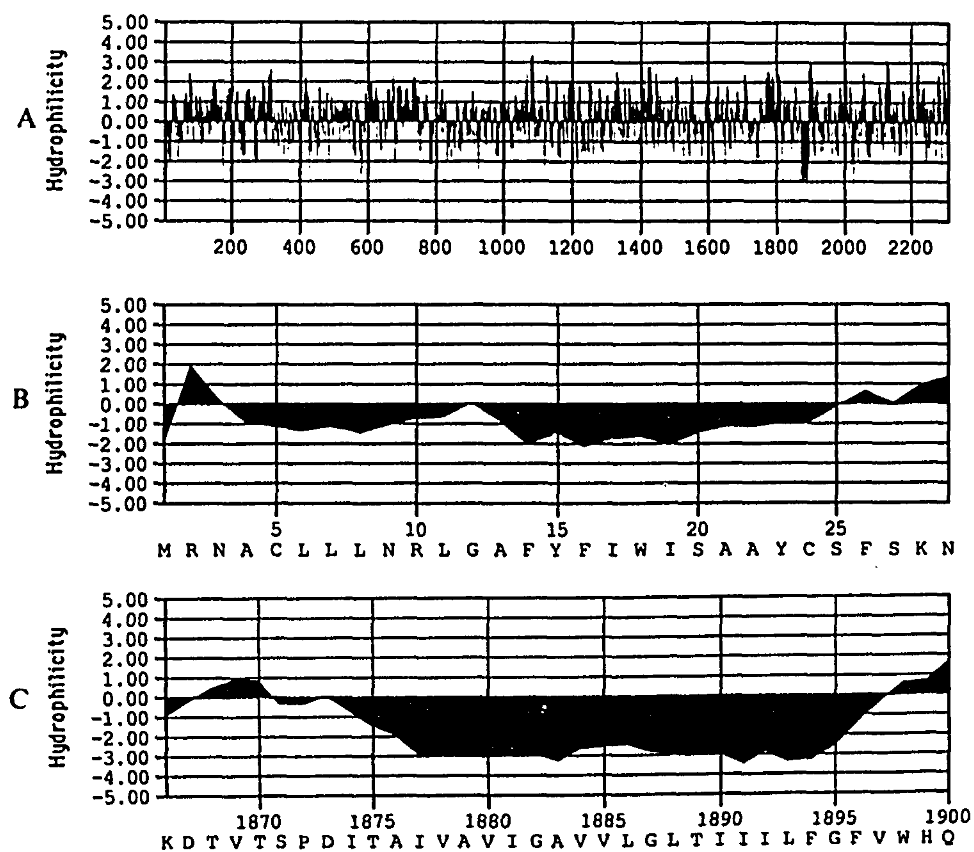


Fig. 3-5 Comparison of the predicted amino acid sequences of chicken c-*ros* cDNA with those of c-*ros* and *sevenless* cDNAs. The smaller percentage numbers represent the amino acid identity and the larger numbers are derived if conserved aa are included in the calculation. The rat c-*ros* sequence was from Matsushime & Shibuya (1990) and that of *sevenless* from Basler & Hafen (1988) and Bowtell et al.(1988).

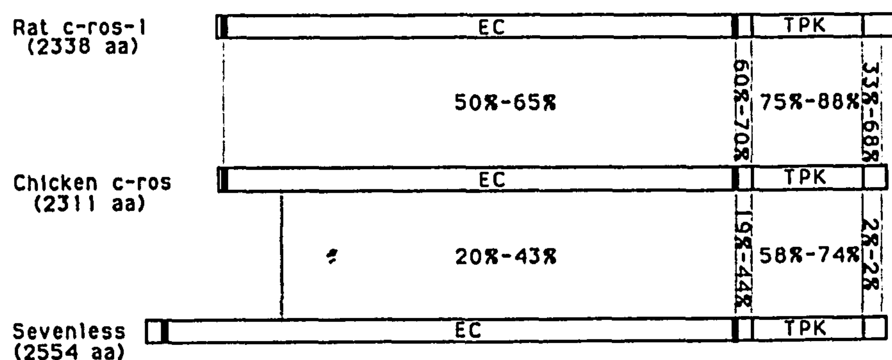


Fig. 3-6 Comparison of the insertion sequences and autophosphorylation sites. The sequence of various PTK's are from respective references cited in the text except that the *src* sequence is from Tekeya et al. (1983). Numbers indicate their positions in the respective amino acid sequences.

A. Insertion Sequence

```
src       NILV------GENL   398
ros^ck    NCLVSEKQYGSCSR  2113
ros^rat   NCLVSVKDY-TSPR  2067
ros^hu    NCLVSVKDY-TSPR  2096
sev^-     NCLVTESTG-STDR  2365
```

B. Autophosphorylation Sites

```
                    *
src       ARLIE--DNEYRK   418
              *   * *
ros^ck    ARDIYKNDY-YRK  2134
              *   * *
ros^rat   AREIYKHDY-YRK  2088
              *   * *
ros^hu    ARDIYKNDY-YRK  2117
              *   * *
sev^-     ARDIYKSDY-YRK  2388
              *   * *
IR        TRDIYETDY-YRK  1165
              *   * *
IGFR      TRDIYETDY-YRK  1138
```

Chapter IV. Characterization of chicken c-*ros* protein product

Abstract

Proto-oncogene c-*ros* is the cellular counterpart of the transforming gene v-*ros* of the retrovirus UR2. v-*ros* encodes a *gag-ros* fusion protein P68 with very high kinase activity. The structure of P68 mimics that of a RPTK. However, little was known about the biochemical properties of the protein product of c-*ros*. Here, we constructed the full length chicken c-*ros* cDNA in different expression vectors and expressed the c-*ros* protein in COS-7 and CEF cells. The results from COS cell expression showed that the c-*ros* was expressed as a 260-280-kDa glycosylated protein with low tyrosine kinase activity. The protein was not covalently dimerized. No cleavage was observed as seen in other closely related RPTKs such as IR, IGFR and *Drosophila sevenless* protein. A mutant protein ppRos was generated by deleting most of the extracellular portion of the c-*ros* protein but retaining the signal peptide. Biosynthesis of the ppRos protein mimics that of the parental full length c-*ros* protein. The major difference is that ppRos was more stable and constitutively active in kinase activity. When the c-*ros* and ppros were inserted into a retroviral vector and expressed in CEF, similar biochemical properties were observed, except that the expression level of neither c-Ros nor ppRos was nearly as high as that of the v-Ros. Both c-*ros* and ppros fail to transform CEF. The signal transduction of ppRos and v-Ros was compared with respect to phosphorylation of some specific cellular substrates and activation of the PI-3-Kinase. The kinase active ppRos was found to be unable to phosphorylate a 88-kDa protein, Shc and PLCγ and fail to associate with PI-3-Kinase. These differences correlate with the inability of ppros to transform CEF.

Introduction

The UR2 encoded P68$^{gag\text{-}ros}$ is a transmembrane protein with the p19 portion protruding extracellularly (Jong & Wang, 1987). Recent studies demonstrated that both the *gag* sequence and the TM domain in P68$^{gag\text{-}ros}$ are important for its transforming ability (Jong & Wang, 1991, 1992). The v-*ros* protein was shown to be able to associate with PI-3-Kinase activity (Fukui et al., 1989; Zong & Wang, unpublished result) and phosphorylate certain cellular proteins involved in the signalling process including Shc and IRS1 (Zong & Wang, unpublished observation).

cDNAs for the proto-oncogene c-*ros* of chicken and rat were recently cloned and sequenced (Mitsushime & Shibuya, 1990; Chen et al., 1991). A 8-kb cDNA for human c-*ros* was cloned from a human glioblastoma cell line (Birchmeier et al., 1990). All the three c-Ros proteins share a high homology with the *Drosophila sevenless* protein, suggesting that c-*ros* might be the vertebrate homolog of the *Drosophila sevenless* gene.

Little was known about the biochemical property and structure of the vertebrate c-*ros* protein. In this study, c-*ros* protein was expressed in CEF under a retroviral promoter, and in COS-7 cells under the control of human CMV promoter and SV40 promoter. The biochemical properties of the c-*ros* protein were studied mainly in the COS-7 transient expression system. A c-*ros* mutant ppRos with most of the EC domain deleted was generated, resulting in an activated kinase activity comparable to that of the UR2 P68$^{gag\text{-}ros}$. Neither the full length c-*ros* nor the ppros when expressed in a retroviral vector is able to transform CEF. The tyrosine phosphorylated substrate pattern of ppRos was compared with that of the v-Ros. Several differences have been noticed, including phosphorylation of a 88-kDa protein, Shc and PLCγ1 and association with PI-3-Kinase. These differences are correlated with the differential transforming ability of ppRos and v-Ros.

Since c-*ros* is specifically expressed in the epithelial cells of various tissues (see next chapter) where epithelial/mesenchymal inductive events occur, it is reasonable to speculate that the ligand for the vertebrate c-*ros* could be a cell surface protein of mesenchymal cells. Based on this hypothesis, a possible expression cloning system was established in this study.

Results and Discussion

Construction of full length cDNA for c-*ros* and the mutant ppros

A full length c-*ros* cDNA of 7-kb from overlapping cDNA clones was constructed by the strategy described in Material and Methods and shown in Fig. 4-1. Since a mutant *sevenless* protein with part of the EC domain deleted conferred the ability to rescue the *sevenless* phenotype (Basler et al., 1991), a mutant *ros* protein ppros was created by deleting most of the extracellular domain of c-*ros*. Since the c-*ros* protein is very large and its expression level is very low (see below), the activated mutant ppRos mimicking a ligand activated c-Ros should facilitate our study of the c-*ros* protein.

Expression of c-*ros* and ppros in Cos cell

I first tried to establish permanent c-*ros* expressing cell lines in mammalian cells. The full length c-*ros* and pp*ros* cDNA was subcloned into the pRc/CMV vector containing the human CMV early promoter and enhancer resulting in pCMVros or pCMVppros. After those plasmids were transfected into COS-1 cell, G418 resistant clones were selected. Fig. 4-2 shows a Western blot analysis of the total cell extracts of some of those G418-resistant clones. Some clones expressed 260-280-kDa protein doublet which was not present in the control COS-1 cell. Some clones expressed a 130-kDa band, which is most likely an rearranged or alternatively spliced *ros* product. The clone 6 named as CMVros6 was used for further protein analysis. [^{35}S]-Met labeling of CMVros6 cells followed by immunoprecipitation of the extract with α-Ros serum revealed the similar doublet (Fig. 4-3A) and the c-Ros doublet was shown to have kinase activity in an in vitro kinase assay (Fig. 4-3B). Similar *ros* proteins were observed in the chicken kidney and intestine extracts (Fig. 4-3C). The expression level per cell was examined by indirect immunostaining using a purified rabbit α-Ros antibody. Surprisingly, only about 1% of cells in CMVros6

expressed the *ros* protein but the expression level in those positive cells was high (data not shown). The experiment was repeated with other vectors containing different promoters including SV40, Mu-MoLV LTR, MMTV LTR or metallathione promoter. Similar results were obtained: no pure and high *c-ros* expressing cell line could be established in COS-7, NIH3T3, Rat1, Rat2, and several glioblastoma and renal tumor cell lines. No cell lines expressing the activated mutant ppRos could be established either. The reason for the failure to establish a c-Ros over-expressor will be discussed together with the c-Ros expression in CEF.

Biochemical properties of c-Ros and ppRos

Since it was difficult to establish a c-Ros over-expressor, transient expression system was used to study the biochemical properties of the c-Ros and ppRos. The COS-7 cell was used for all the transient expression studies because it is a SV40 large T transformed cell line which can promote high level expression from SV40 promoter and can amplify the foreign plasmid with an SV40 origin to multiple copies (Gluzman 1981). *c-ros* was cloned into pECE vector (Ellis et al., 1986) containing the SV40 early promoter and SV40 origin. The ppros was cloned into vector pRc/CMV containing the human CMV promoter which is one of the strongest promoters identified so far. The *c-ros* was also cloned into pRc/CMV, resulting in pCMVros. pCMVros (12-kb) gave much lower expression than pECEros (10-kb), probably because the larger plasmid had lower transfection efficiency. By indirect fluorescence immunostaining, fewer cells were shown to be stained positively in pCMVros than in pECEros transfected cells (data not shown). The pECEUR2 expressing $P68^{gag\text{-}ros}$ was always used as a control because its biochemical properties had been characterized (Garber et al, 1986; Jong & Wang, 1987; Jong & Wang, 1990). The subcellular localization of c-Ros and ppRos transiently expressed in

COS-7 cells was examined by indirect immunofluorescence using affinity purified antibody that recognizes the cytoplasmic domain of v-Ros. Overall a similar pattern of distribution of c-Ros, ppRos and v-Ros was observed (Fig. 4-4). The majority of the immunoreactive protein was localized in the intracellular structures possibly representing membrane vesicles which is consistent with the high level expression in the transient expression system. The same antibody fails to stain the unpermeablized cells as predicted (data not shown), since the cytoplasmic PTK domain should be in the cytoplasm. However, v-Ros protein displayed a more uniform pattern implying a cell membrane staining. The c-Ros and ppRos, particularly the former, were localized in the perinuclear area and displayed more clearly the endoplasmic reticulum (ER) network staining, implying that the majority of the protein was associated with the ER vesicles and was less abundant on the cell surface.

The intracellular localization of these proteins was further studied by subcellular fractionation. Individual *ros* plasmid-transfected COS-7 cell homogenates were fractionated into S100 and P100 fractions (Fig. 4-5A, B and C). Confirming previously published results (Jong & Wang, 1991), most of the P68$^{gag\text{-}ros}$ was detected in P100 fractions with high kinase activity. Similarly, both c-Ros and ppRos were also mainly associated with the p100 fraction. By comparing the kinase activity and protein amounts, ppRos has a specific kinase activity about half of that of P68$^{gag\text{-}ros}$ and at least twice of that of c-Ros. The specific activity of c-Ros could be overestimated because c-Ros was transferred to nitrocellulous paper very inefficiently during western blot. The lower kinase activity of c-Ros is predictable in the absence of its cognate ligand. Activation of ppRos is apparently due to the deletion of most of the EC domain which may exert negative constraint on the c-Ros PTK activity.

The P100 membrane fractions were further fractionated in a discontinuous sucrose gradient and analyzed for the kinase activity and protein amount (Fig. 4-6). Similar to v-Ros, the distribution of ppRos peaked at both heavy and light membrane fractions. In contrast, the c-Ros kinase activity was mainly associated with the heavy membrane fraction (fraction 2). The reason for the appearance of relatively higher amount of c-Ros protein in fraction 5 was unclear (Fig. 4-6C). That less abundance of the c-Ros protein or activity in the light membrane was probably due to its instability and failure to be properly processed and transported to the cell as also suggested by the immunostaining described above. Actually, more than one-ninth of v-Ros and ppRos are associated with the light membrane fraction. Confirming the light membrane association, both v-Ros and ppRos could be labeled with a cell impermeable biotinylation reagent Sulfo-NHS-biotin (Pierce). As a control, the same reagent failed to label the SrcXRos chimera protein locating intracellularly (Fig 4-7). Under similar conditions, the c-Ros could not be labeled. However, the interpretation is complicated by the very low level of c-Ros expression. Taking together the fact that both c-Ros and ppRos are associated with membrane fractions and some with light membranes, and ppRos can be labeled on cell surface by Sulfo-NHS-biotin in live cells, it is reasonable to conclude that both c-Ros and ppRos are expressed on the cell surface or the vesicle membranes. Similar results were shown previously for the *Drosophila sevenless* protein (Simon et al., 1989). The half life of the c-*ros* protein was about one hour. The kinetic study of its biosynthesis by [^{35}S]-Met pulse-chase labeling showed that the c-Ros was first synthesized as a lower 260-kDa species and further glycosylation and/or phosphorylation yielded its upper 280-kDa species which was likely the mature form. Both of the 260- and 280-kDa c-Ros doublet were apparently glycosylated as evidenced by their disappearance in the presence of the

glycosylation inhibitor tunicamycin (Fig. 4-5C, lower panel). Without N-linked glycosylation, c-Ros appeared as a 250-kDa polypeptide consistent with what was predicted from its primary sequence (Chen et al., 1991). Since there are remaining potential N-linked glycosylation sites in the EC domain of ppRos, it was expected to be glycosylated also (Fig. 4-5C, the top panel). Indeed tunicamycin converted the apparent molecular weight of ppRos from 85-kDa to 70-kDa predicted from its primary sequence. The half life of ppRos was much longer than that of c-Ros protein. In another experiment (data not shown), the half life of ppRos was estimated to be about 8 hours.

Oligomeric state of c-Ros, ppRos and v-Ros

Since the c-Ros-related IR and IGFR are both covalently linked heterotetramer with a structure of $\alpha2\beta2$, it is worthwhile to examine the covalent oligomerization state of the *ros* proteins. The *ros* proteins were analyzed in parallel in non-reducing and reducing gel (Fig. 4-8A). Kinase assay was used for its high sensitivity, however similar results were obtained by Western blot analysis (data not shown). Initial result showed that c-Ros was not covalently dimerized, but the $P68^{gag\text{-}ros}$ and ppRos were dimerized and possibly oligomerized Ffig. 4-8A). T6 was used as a positive control because it codes for a Gag-IR fusion protein which is most likely a dimeric protein. Since in both CEF and COS cell transient system, $P68^{gag\text{-}ros}$ and ppRos were over-expressed, artificial inter-molecule disulfide bonds could be formed during or after the lysis of cells. To avoid the formation of new disulfide bonds in vitro, an alkylating reagent iodoacetamide (IAA) was included at a concentration as high as 1 mM in the RIPA lysis buffer. Similar experiment was done as in Fig. 4-8A except NM1, coding for a Gag-IGFR fusion was included in this case as a positive control (Liu et al., 1992). ppRos and $P68^{gag\text{-}ros}$ from COS cell and CEF respectively were free of covalently linked dimers. Confirming the previous result, NM1

was shown to be covalently dimerized (Liu et al., 1992). It is concluded that unlike IR and IGFR, *ros* proteins appear not to form covalent dimers. Whether non-covalently linked oligomers exist for c-Ros and ppRos needs to be determined.

Kinase activity of c-Ros, ppRos and v-Ros

One of the common features among Ros, IR and IGFR and *sevenless* protein is the presence of a cluster of 3 tyrosine residues in the presumed major autophophorylation site. To investigate the importance of these tyrosine residues on the kinase activity of c-Ros, a mutant RosPF3 was made by mutating all the three tyrosine residues into phenylalanine. The kinase activity of RosPF3 was compared with c-Ros in terms of autophosphorylation and phosphorylation of the exogenous substrate enolase (Fig. 4-9). The autophosphorylation activity of RosPF3 decreased more than ten fold (Fig. 4-9A), so was its ability to phosphorylate the exogenous substrate enolase (Fig. 4-9B). This result indicates that the three tyrosine residues are important for the kinase activity of c-Ros and very likely are its major autophophorylation sites.

Since the ppRos has an elevated kinase activity similar to that of v-*ros* oncogene but fails to transform CEF (see below), there exists the possibility that the activated ppRos is capable of autophosphorylation but not phosphorylating exogenous substrates. However, ppRos and v-Ros can phosphorylate exogenous substrate, lysozyme, to the same extent (Fig. 4-10A). The in vitro phosphorylation of c-Ros and ppRos are mainly on phosphotyrosine (pTyr) residues because treating the gel with 1 N KOH at 55^{0}C does not decrease the ^{32}P amount associated with them (data not sown). Furthermore, the ppRos is phosphorylated on tyrosine residues to the same extent as v-Ros in transiently transfected COS-7 cells (Fig. 4 -10B), suggesting that ppRos can autophosphorylate *in vivo*.

Expression of c-Ros and ppRos in CEF and their transforming ability

The cDNAs of c-*ros* and ppros were inserted into a retroviral vector under the control of the retrovirus UR2 LTR and used for expression in CEF. The retroviral vector containing c-*ros* or ppros were transfected together with UR2AV DNA into CEF and after two rounds of selection in soft agar, the cells were assayed for expression of c-Ros and ppRos and colony formation. In CEF, c-Ros was expressed as a broad 260-280-kDa protein band and ppRos as a 85-kDa polypeptide (Fig. 4-11). Since the calculated molecular weights for c-Ros and ppRos are 250-kDa and 70-kDa respectively, the c-Ros and ppRos proteins expressed in CEF must have undergone some post-translational modifications such as glycosylation and/or phosphorylation. Consistent with the results obtained from COS cells, the kinase of ppRos is much more active than that of c-Ros. The amount of ppRos protein expressed in CEF was much less than that of v-Ros after two transfers post transfection (data not shown). Therefore, if normalized for the amount of protein, the kinase activity of ppRos was comparable to that of v-Ros observed in COS cells. CEF cells expressing c-Ros and ppRos showed no morphological change when compared with the parental CEF. In contrast, UR2 rapidly transformed CEF into elongated fusiform morphology (Fig. 4-12A). Unlike UR2 transformed CEF, CEF expressing ppRos and c-Ros failed to form colonies (Fig. 4-12B). In another experiment, pUppcros and UR2 was transfected into CEF and their protein expression level was analyzed 48 hours later and found to be similar (data not shown), indicating there is no transcriptional or translational difference among UR2, c-*ros* and ppros retroviral constructs. Virus stocks collected from the c-*ros* and ppros transfected CEF medium could be used to reinfect CEF for c-Ros and ppRos expression. Therefore, the reasons for failure to amplify the c-Ros and ppRos expressing CEF include the following possibilities: 1) as mentioned above, c-

Ros and ppRos are unable to transform CEF so that the transformed cells have no growth advantage and can not be selected. 2) c-Ros and ppRos proteins are toxic to CEF. 3) the c-*ros* proteins are differentiation related rather than proliferation related so that the cells expressing c-*ros* proteins will stop mitosis. Failure of c-Ros or ppRos to transform CEF could be due to their inability to interact with certain downstream signalling molecules which are important for transformation. However, comparison of tyrosine phosphorylation substrate patterns of c-Ros or ppRos expressing CEF with that of UR2 was hampered by the low level expression of c-Ros proteins.

Signal transduction of c-*ros*, ppros and v-*ros* proteins

Comparison of tyrosine-phosphorylated substrates by ppRos with those of transforming proteins v-Ros and VCRos (Zong et al., 1993) was made possible by using transient over-expression system in CEF since ppRos could not be expressed at a similar level as v-Ros or VCRos in UR2 retroviral vector. This was achieved by transfecting the CEF with pCMVppros. Comparison of substrate phosphorylation 48 hours after transfection was done by pretreating the cells with 50 μM Vanadate overnight. Equal amount of cell extracts were immunoprecipitated with α-pTyr Py20 and PT22-1 and the immunoprecipitates were subjected to western blot analysis with an α-pTyr antibody RC20 (Fig. 4-13). The untransfected CEF had very few tyrosine-phosphorylated proteins. UR2 and VCRos transformed CEF had many tyrosine-phosphorylated proteins and some of them are different between the two virus transformed CEF populations. The VCRos is similar to v-Ros except that the cytoplasmic portion of v-Ros was replaced by corresponding sequences from c-Ros (Fig. 4-1; and Zong et al., 1993). The use of VCRos could distinguish that the signalling difference between ppRos and v-Ros was not due to their different cytoplasmic domain. The ppRos expressing CEF displayed a distinctive

pTyr-containing protein pattern from those of v-Ros and VCRos. Notably, ppRos failed to recognize protein species at the range of 120-140-kDa, a 88-kDa band and a 36-kDa band, instead, it phosphorylated a 100-kDa protein more strongly than v-Ros and VCRos. Failure of ppRos to interact with substrates important for mitogenic signalling or its stimulation of certain growth suppressing molecules, or combination of both, could account for the inability of ppRos to transform CEF.

Tremendous progress has been made in identifying the molecules involved in signal transduction pathways after activation of RPTK by ligand binding. PI-3-kinase, PLCγ1, and Shc coupled to *ras* signaling pathway have been shown to be involved in the signal transduction of a number of activated receptors (Cantley et al., 1991). Recently, PLCγ1 and PI-3-kinase were shown to be important for PDGF receptor mediated mitogenic signal (Valius & Kazlauskas, 1993). To study the signal transduction of *ros* proteins and molecular basis for the differential transforming capability between ppRos and v-Ros, several signaling components, including the Ros-associated PI-3-kinase activity, phosphorylation of PLCγ1 and Shc, were analyzed.

Strong PI-3-kinase activity was found to be associated with α-Ros immunoprecipitate of UR2 transformed CEF extract, but very little PI-3-kinase activity was associated with ppRos (Fig. 4-14) despite that more ppRos was brought down by the anti-serum as reflected in the kinase assay (Fig. 4-14B).

Shc is a recently identified member of the SH2-containing proteins which couples the receptors to Grab2, another SH2-containing protein which in turn will associate with yet another downstream signalling component Sos, leading to activation of ras (Egan et al., 1993; Rozaki-Adcock et al., 1993; Li et al., 1993; Gale et al., 1993). Again, in v-Ros and VCRos transformed CEF, the phosphorylation of the three species of Shc was

increased as compared with normal CEF, whereas no increase of Shc phosphorylation was seen in ppRos transfected CEF (Fig. 4-15A). The amount of Shc protein was similar for all cells (Fig. 4-15A). This finding suggests that the *ras* pathway might not be activated in ppRos signalling but that Shc does play a role in v-Ros and VCRos signalling. The phosphorylation of Shc was studied in a UR2 ts mutant 251 in a temperature shift experiment (Fig. 4-15B). The 66-kDa Shc is most sensitive to the temperature shift. Shc was dephosphorylated within 1 hour when temperature was shifted from 35 to 41°C, suggesting that Shc is important for the v-Ros signalling. However, the phosphorylation of Shc was not affected much when temperature was shifted from 41 to 35°C, suggesting that Shc might not be a direct substrate of v-Ros. Again the differential Shc phosphorylation was not due to a change in its level of expression. The nature of the ts mutant was confirmed by showing that ts251 v-Ros is strongly phosphorylated at 35°C, but not at 41°C (Fig. 4-15C).

It was previously shown that phosphorylation of PLCγ1 is important for its activation (Kim et al., 1991). The phosphorylation of PLCγ1 was compared among CEF cells expressing various *ros* protein (Fig. 4-16A). For the first time, PLCγ1 was shown to be phosphorylated in v-*ros* and VCros transformed CEF, suggesting that PLCγ1 could be involved in the downstream signalling pathway of v-Ros. The cells expressing ppRos showed much less phosphorylation of PLCγ1 despite no variation in its protein level. To confirm the phosphorylation of PLCγ1 is v-Ros-specific, PLCγ1 phosphorylation by ts251 was analyzed (Fig. 4-16B). The α-pTyr blot shows that the phosphorylation of PLCγ1 decreased when temperature was shifted from 35 to 41°C and increased when temperature was down-shifted without change of PLCγ1 protein level throughout the course (Fig. 4-16B). This suggests that PLCγ1 could be a downstream substrate of v-Ros.

Recently, a panel of monoclonal antibodies against tyrosine phosphorylated proteins in UR2 transformed CEF were generated in our laboratory (Chan, J, Moran, T., Jong, S.-M., and Wang, L.-H., unpublished result). Among them, the 5C2 MAb recognizes multiple protein bands around 85- to 88-kDa which are phosphorylated in UR2 transformed CEF. Interestingly this protein(s) referred to as 88-kDa protein can not be phosphorylated in a mutant of v-Ros in which the 3 amino acid SLT insertion in the v-Ros TM sequence was deleted (Zong & Wang, unpublished observation). Phosphorylation of the 88-kDa protein was compared in v-Ros, VCRos and ppRos expressing CEF (Fig. 4-16C). The 88-kDa protein was phosphorylated in UR2 and VCros transformed CEF cells, but, little phosphorylation was detected in CEF expressing ppRos. Whether this protein is important for the transforming activity of v-Ros remains to be determined.

The failure of c-Ros to transform the CEF could be explained by its low kinase activity in the absence of its ligand. However, the inability of ppRos to transform CEF could not be explained by its weak kinase activity since its kinase activity was comparable to that encoded by the viral oncogene v-*ros*, neither could it be explained by its stability because ppRos has a half life as long as 8 hours. Nor could it be accounted for by its subcellular localization since ppRos was synthesized and transported to the cell membrane when expressed in COS cells. Since a v-*ros* and c-*ros* chimeric molecule VCros with the intracellular domain replaced by that of c-*ros* could transform CEF as well as UR2 (Zong et al., 1993), it seems that the *gag* sequence or the TM domain may play an important role in *ros* mediated cell transformation. The difference between c-*ros* and v-*ros* TM domain is the presence of a three amino acid (SLT) insertion in the C-terminal half of the latter. It seems that this SLT insertion may affect the interaction and phosphorylation of the v-Ros substrates (Zong & Wang, unpublished result). It would be

interesting to see what the effect will be by introducing this three amino acids into the TM domain of ppRos. Since previous studies by our laboratory suggested that the immediate EC sequence upstream of the TM domain of gag-IR, gag-IGFR and gag-Ros (Poon et al., 1990; Liu et al., 1991; Zong et al., unpublished results) exerts negative effect on the transforming potential of those *gag* fusion proteins, one may suspect that the inability of ppRos to transform CEF is due to the presence of 27 additional amino acids immediate upstream of the TM domain in comparison with v-Ros. However, deletion of those 27 amino acids does not activate the transforming potential of ppRos despite the fact that ppRos was again as active as that of v-Ros (data not shown). Since it was previously shown that *gag* sequence in v-*ros* is important for transforming activity and fusion of *gag* to truncated IR and IGFR could activate their transforming potential, lack of *gag* sequence in ppRos could account for its lack of transforming ability. The above results suggest that the EC domain could exert a very strong influence on the cytoplasmic domain including its kinase activity and the interaction with the downstream signalling molecules. Apparently ppRos interacts with its downstream signalling molecules differently from v-Ros and VCRos due to their different EC sequences since the cytoplasmic domain of ppRos and VCRos are identical and all of their kinase activities are comparable. The fact that the constitutively activated ppRos signals differently from v-Ros also implies there exists crucial difference between the signal transduction pathways of c-Ros and v-Ros. The mutant represented by ppRos, i.e. kinase positive and transforming negative, is instrumental in helping to identify the important components mediating the *ros*-induced cell transformation.

Exploration of cloning systems for the identification and isolation of the cDNAs encoding the ligand for the c-Ros receptor

The ligand for the *Drosophila sevenless* protein has recently been identified to be the product of the *bride of sevenless* (*boss*) which is located on the cell surface of the neighboring cell of photoreceptor cell R7 where *sevenless* is expressed (Krämer et al., 1991). Our and others' studies of c-*ros* expression in chicken and mouse suggest that c-*ros* may play a role in the epithelial/mesenchymal induction during kidney and intestine organogenesis (Sonnenberg et al., 1991; Tessarollo et al., 1992; Chen et al., 1992). Based on the previous result that the induction of kidney ureter to branch and differentiate requires the contact between the ureteric bud and the surrounding mesenchyme (Gilbert, 1987), it is reasonable to speculate that the ligand for vertebrate c-*Ros* is very likely to be located on the cell surface of the mesenchyme. We and others have failed so far to isolate a *boss* homologous gene in vertebrate probably due to the long distance and functional divergence between *Drosophila* and chicken and mammals. I have tried to establish an expression screening and cloning system based on the hypothesis that c-Ros ligand is expressed on the cell surface. The chicken kidney and intestine cDNA expression library was made in the plasmid vector pcDNAI (Invitrogen) containing CMV early promoter and enhancer and SV40 origin which allows the maximum replication and expression of transfected plasmids in COS-7 cells. (Aruffo & Seed, 1987). I have made kidney and intestine cDNA libraries in pcDNAI with a complexity as high as 1 to $5X10^7$. Such high complexity should be large enough to represent all the mRNAs, including rarely transcribed genes.

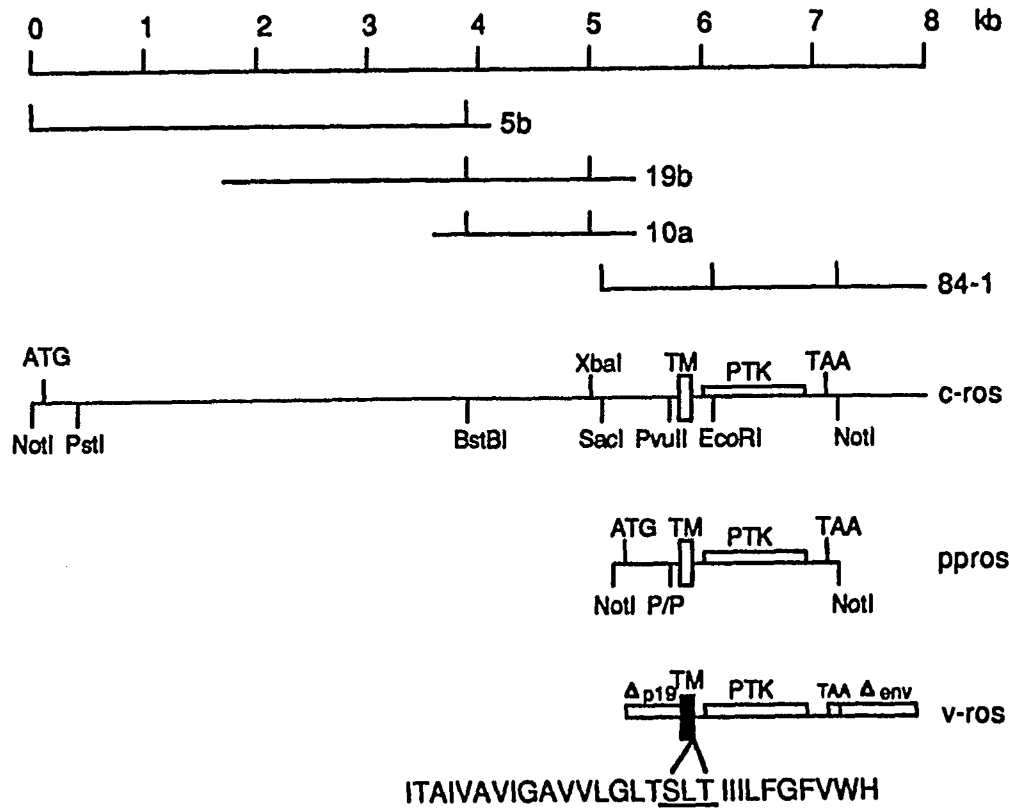
Since we do not have good antisera against EC domain of chicken c-Ros and the expression level of c-Ros protein is very low, much effort was made to produce a good c-Ros extracellular (EC) domain as a probe for screening and cloning. The whole c-Ros EC domain was fused to a human placenta alkaline phosphatase catalytic domain

(Flanagan & Leder, 1990) and the resulting plasmid pCMVrosAP (Fig. 4-17A) was transfected into COS-7 cell and G418 resistant cell clones expressing the RosAP protein were isolated. Since the pCMVrosAP contains the signal peptide from the c-Ros but not the TM and cytoplasmic domain, the fusion protein RosAP should be secreted into the medium. As shown in Fig. 4-18A, a large proportion of the RosAP could be detected in the supernatant of RosAP36 but not in the control COS-7 cells. The highest expressor clone RosAP36 was amplified which can be used for the production and purification of large quantity of RosAP. The RosAP in the concentrated supernatant can be used to stain the positive cells expressing the Ros ligand or to screen the expression library.

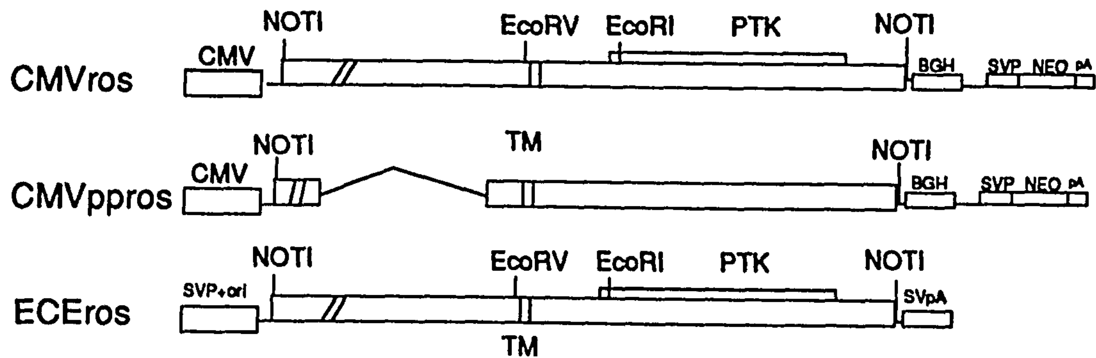
Another fusion protein was made utilizing human IgG_2Fc. The c-Ros EC domain was inserted in frame before the human genomic DNA encoding the human IgG Fc portion in the vector pCDM8 (Kindly provided by Dr. Seed) and the resulting plasmid was called pCDMrosFc (Fig. 4-17B). The pCDMrosFc was transfected into COS-7 and the RosFc protein could be detected in the medium (Fig. 4-18B). Less RosFc was secreted to medium than RosAP, probably due to the fact that the expression of RosFc was assayed in a transient expression system while the RosAP is secreted from an established cell line. The use of RosFc could ease the purification of the RosFc protein simply by passing through protein A affinity column or α-human Fc antibody affinity column. Furthermore, considering the high background when using alkaline phosphatase and color reaction (RosAP), RosFc may turn out to be a better probe because the [^{125}I]-protein A could be used to bind to RosFc. The combination of expression cloning, radiolabeling and emulsion developing (Wang et al., 1991) should be a useful approach to identify the c-Ros ligand.

Fig. 4-1. Constructs of c-*ros*, v-*ros* and VCros in mammalian expression vector and retroviral vectors. (A) The structure of c-*ros*, ppros and v-*ros* is shown in scale with various structural domains and restriction sites. Different cDNA clones were used to construct full length c-*ros* cDNA as described in Materials and Methods. The respective terminal codon TAA is indicated with v-*ros* terminating within the viral *env* region. The TM domains of v-*ros* and VCros containing 3 aa-insertion are marked as solid boxes and the aa sequence around the SLT insertion is also shown. (B) c-*ros* and ppros in various mammalian expression vectors. pCMVros and pCMVrospp contain human CMV early promoter and enhancer which are marked as CMV. The poly A signal for pCMV vectors has been derived from the bovine growth hormone gene and is marked as BGH. Both pCMV plasmids have Neo^r gene under the control of SV40 early promoter. pECEros uses SV40 early promoter and contains a SV40 replication origin to allow maximal amplification of the plasmid in COS cells. (C) The various c-*ros* and its recombinant retroviral plasmids were constructed as described in Materials and Methods or have been described previously (Zong et al., 1993). Small open boxes denote viral *gag* and *env* sequences. Large speckled boxes indicate the v-*ros* sequence. Large open boxes represent the c-*ros* sequence. The TM domains of c-*ros* and v-*ros* are represented by the open and solid boxes respectively. The ectodomains of Ufcros and Uppcros are interrupted to reflect their actual lengths. The deletion in Uppcros is shown by the two bent lines.

A



B



C

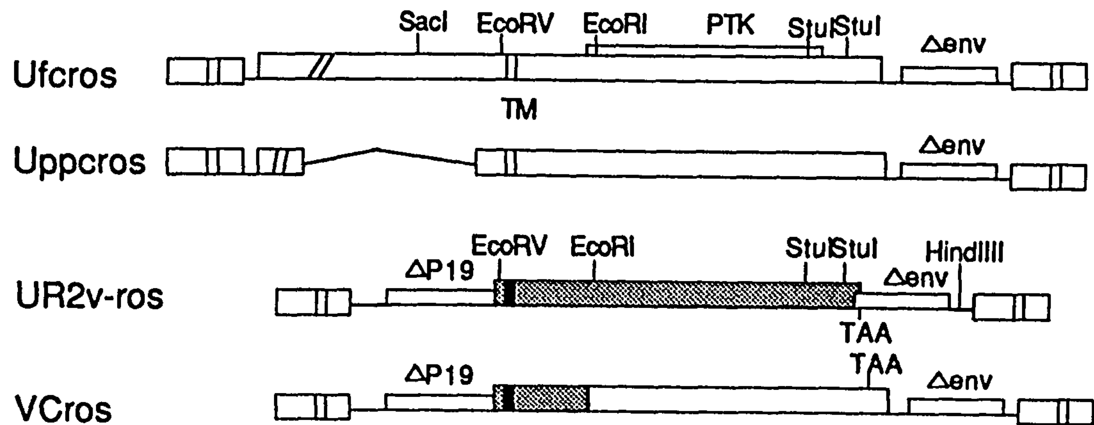


Fig. 4-2 Western blotting of pCMVros transfected COS cell clones selected by G418. For each clone, cells from a confluent 6-cm dish were lysed with cell lysis buffer and one quarter of the extract was loaded onto a 6% SDS polyacrylamide gel for Western analysis. The blot was probed with α-Ros serum followed by [^{125}I]-protein A. COS is untransfected COS cells served as the negative control. Numbers on the top of each lane represent different G418 resistant clones. Clone 6 was named as CMVros6. Arrows on the right side indicate the position of the c-*ros* protein doublet. Numbers on the left are marker molecular weights in kilodaltons.

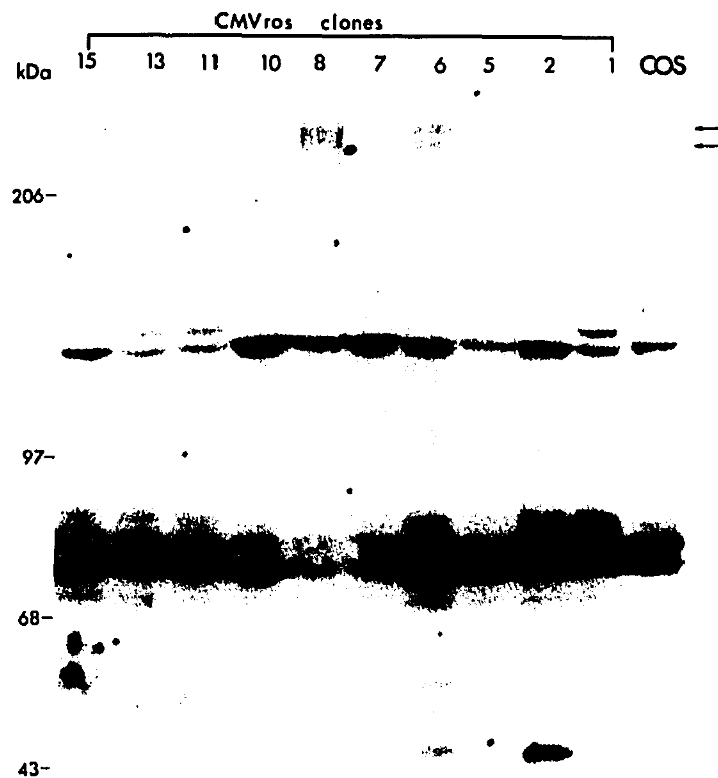


Fig. 4-3 Expression of c-*ros* protein in CMVros6 cells. (A) Effect of tunicamycin on c-*ros* expression. One 6-cm dish of each culture was labeled with [^{35}S]methionine (100 μCi) for 2 h and immunoprecipitated with α-Ros serum before analysis by SDS-PAGE (6%). Tunicamycin was used at a concentration of 10 μg/ml and the cells were pretreated for 2 h and continued throughout the labeling period. Untransfected COS cell ere served as negative control. (C) Kinase activity of the c-*ros* protein. One 6-cm dish of COS cell (lane 1) or CMVros6 cell (lane 2) was extracted with RIPA buffer containing 0.05% SDS, immunoprecipitated with α-Ros serum and then assayed for kinase activity before being separated on SDS-PAGE (5%). Arrows indicate the c-*ros* doublet. (C) The native c-*ros* protein in chicken kidney and intestine. Tissues were extracted as described in Materials and Methods and α-Ros immunoprecipitates were similarly subjected to kinase assay. K, 9 day chick kidney; I, 9 day chick intestine.

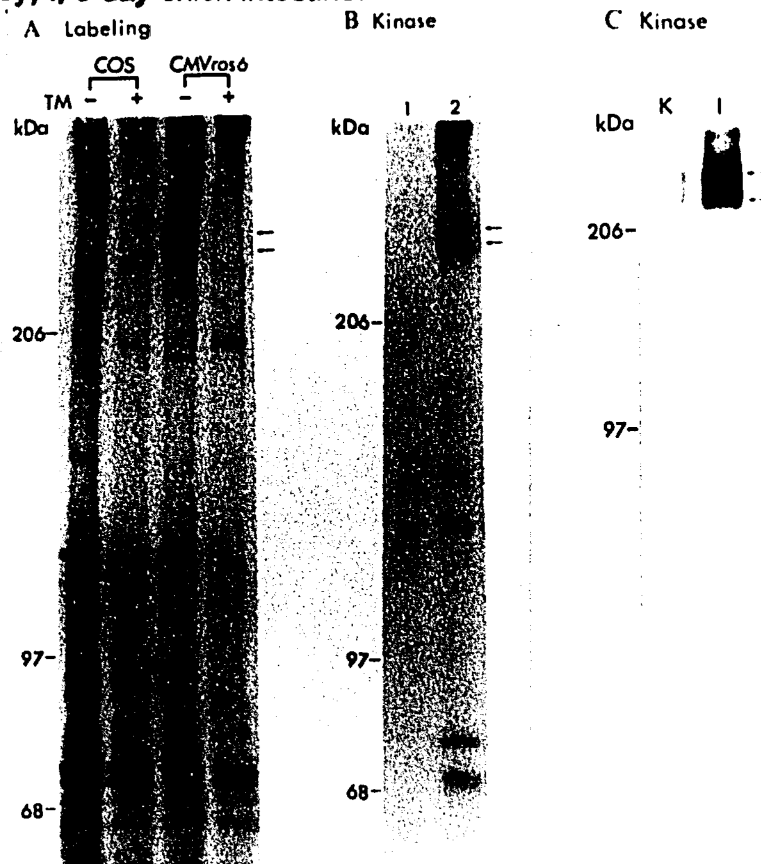
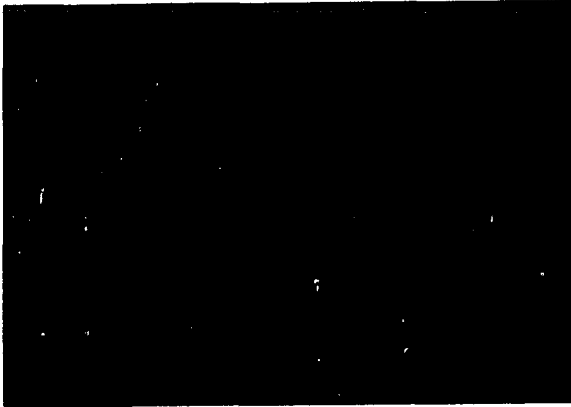


Fig. 4-4 Indirect immunostaining of COS-7 cells transiently expressing c-Ros, v-Ros and ppRos. COS-7 cells were transfected with pECEros, pECEUR2 or pCMVppros. 48 h later, cells were fixed and permeablized and sequentially incubated with purified α-Ros antibody, biotinylated donkey α-rabbit antibody and avidin-FITC. DAPI was used to specifically stain the nucleus.

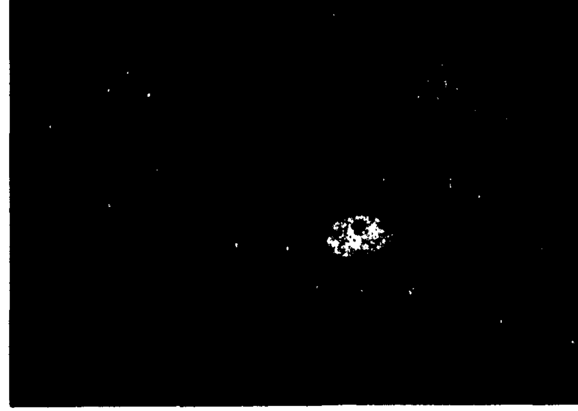
A. UR2



B. DAPI

C. PP

D. DAPI



E. ROS

F. DAPI

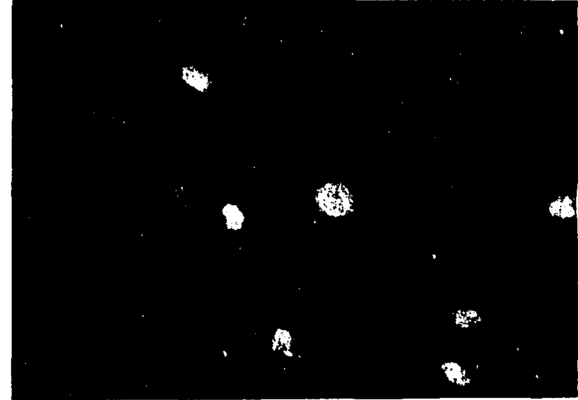


Fig. 4-5 Membrane association, kinetics of biosynthesis and glycosylation of c-Ros and ppRos. Fractionation of the pUR2ECE, pCMVppros and pECEros transfected cell homogenates were fractionated into S100(S) and P100(P) fractions and the α-Ros immunoprecipitates were analyzed by kinase assay (A) or Wester blot (B). Each lane represents 1/3 of the cell extract from a 10-cm dish. (C) Stability of ppRos and c-Ros. One 6-cm dish each of the transfected COS-7 cells was pulse-labeled with 200 µCi of $[^{35}S]$-Met for 20 min and then chased for the time indicated with or without tunicamycin treatment as described in Fig. 4-3. Top panel: ppRos; lower panel: c-Ros. UR2 represents $P68^{gag\text{-}ros}$, ROS represents c-Ros protein and PP stands for ppRos protein. Exposure time for A: 1 h for UR2 and PP lanes and 6 h for ROS lanes; B: 24 h for UR2 and PP lanes and 72 h for ROS lanes; C: top panel, 24-hour exposure, lower panel, 72-hour exposure.

Kinase Assay

Western Blot

pulse-chase

Fig.4-6 Subcellular fractionation of COS-7 cells transfected with v-*ros*, c-*ros* or ppros plasmid. The P100 fractions from 3 10-cm dishes were subjected to discontinuous sucrose gradient sedimentation for 16 h at 100,000 g. Each fraction was diluted with RIPA buffer and immunoprecipitated with α-Ros. Half of the immunoprecipitate was subjected to western blot and the other half to kinase assay. A. v-Ros; B. ppRos and C. c-Ros. Only the kinase assay was shown for c-Ros because the protein quantity was too little to be revealed by Western blot analysis.

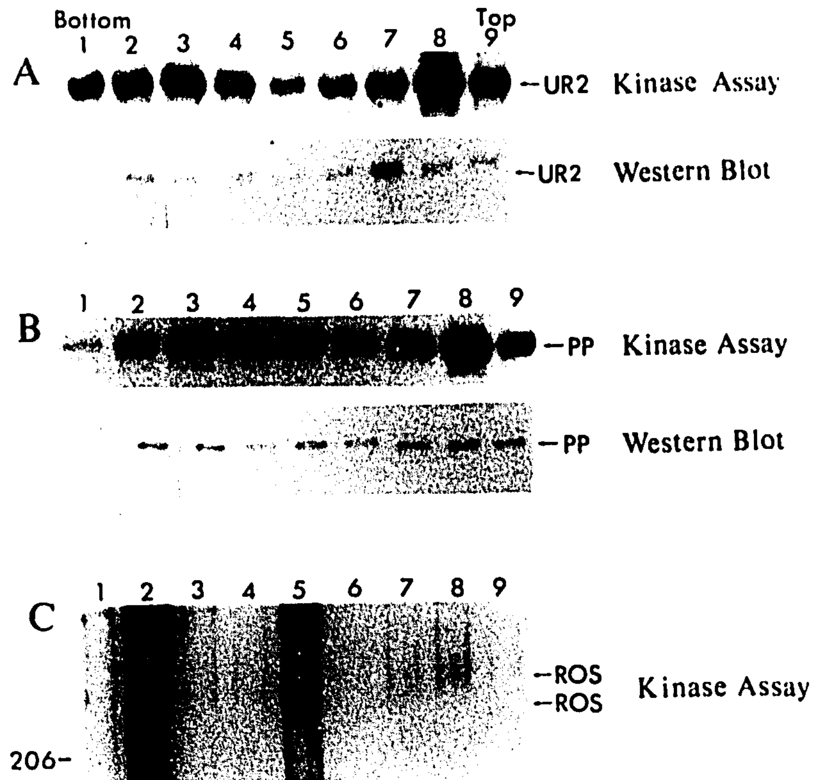


Fig. 4-7 Surface labeling of the *ros* proteins. Transient transfected COS-7 cells (for ppRos) or retrovirus transformed CEF (for v-Ros and SrcXRos) was labeled with water-soluble and cell-impermeable Sulfo-NHS-biotin and analyzed as described in Materials and Methods. The expressed SrcXRos protein is 75-kDa (Jong & Wang, 1992).

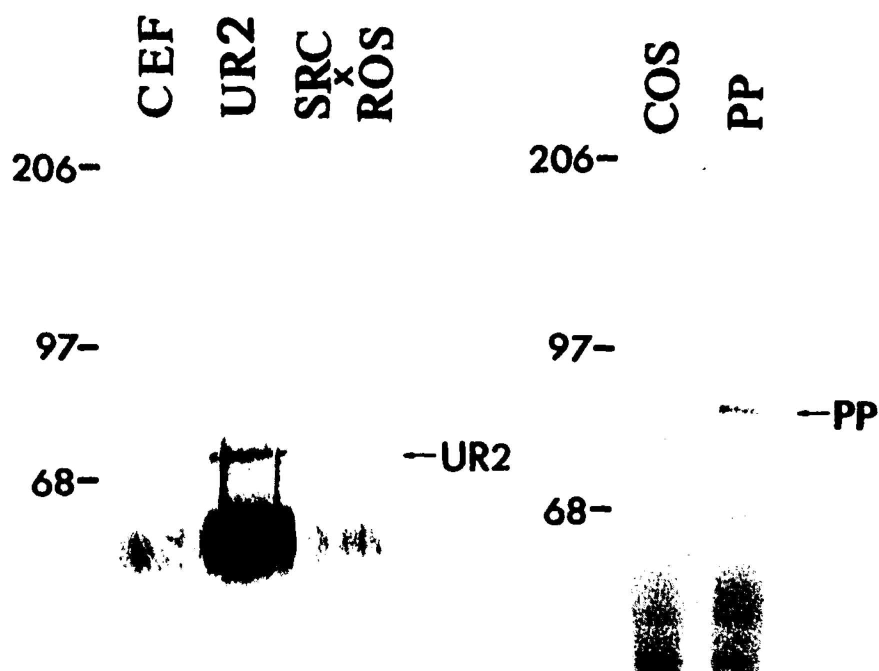


Fig. 4-8 Covalent dimerization state of the c-Ros, ppRos and v-Ros. (A) Cell lysis in the absence of IAA (iodoacetamide). c-*ros*, ppros transfected COS-7 cells and UR2 and T6 (encodes gag-IR) transformed CEF cells were extracted in RIPA lysis buffer and α-Ros or α-IR immunoprecipitates were subjected to kinase assay. Half of the immunoprecipitates were separated in a reducing SDS-PAGE and the other half in a non-reducing gel. B. Similar experiment was done as in A, except that 1mM IAA was added to the lysis buffer to prevent the formation of disulfide bond during or after lysis of cells. UR2-C representing COS cells transient transfected with v-*ros* was also also included in this experiment. NM1 (gag-IGFR) which was shown to be covalently-linked dimer (Liu et al, 1992) was used as a positive control.

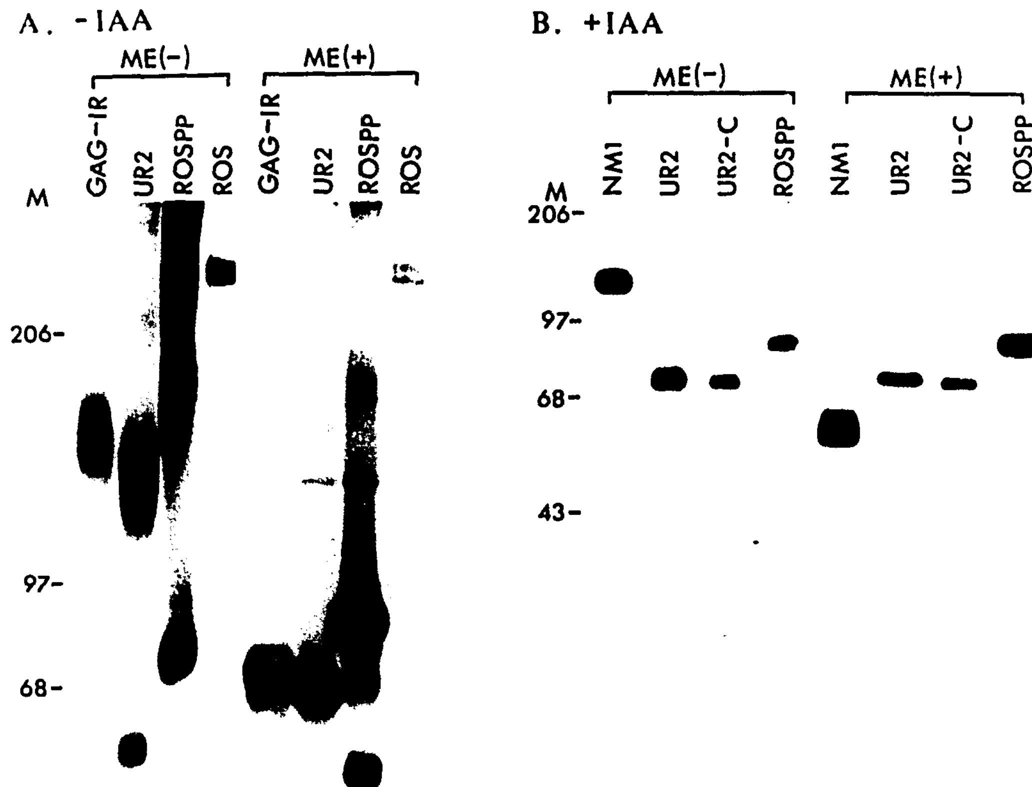


Fig. 4-9 Effect of the triple tyrosine residues on the kinase activity of c-Ros. The characteristic triple tyrosine residues of Ros kinase domain was mutated by PCR into triple phenylalanine. The mutated pECEros is named as pECEPF3. The pECEros and pECEPF3 were transfected into COS-7 cells. 48 h later, one dish of cells were lysed in RIPA and α-Ros immunoprecipitates were subjected to kinase assay in the presence of enolase, another dish was labeled with [^{35}S]-Met for 4 h, lysed and immunoprecipitated with α-Ros antibody for quantitation of protein amount. (A) Autophosphorylation of c-Ros and RosPF3. (C) Phosphorylation of enolase by c-Ros and RosPF3. A and B are from the same reaction. (C) [^{35}S]-Met labeling of c-Ros and RosPF3 proteins.

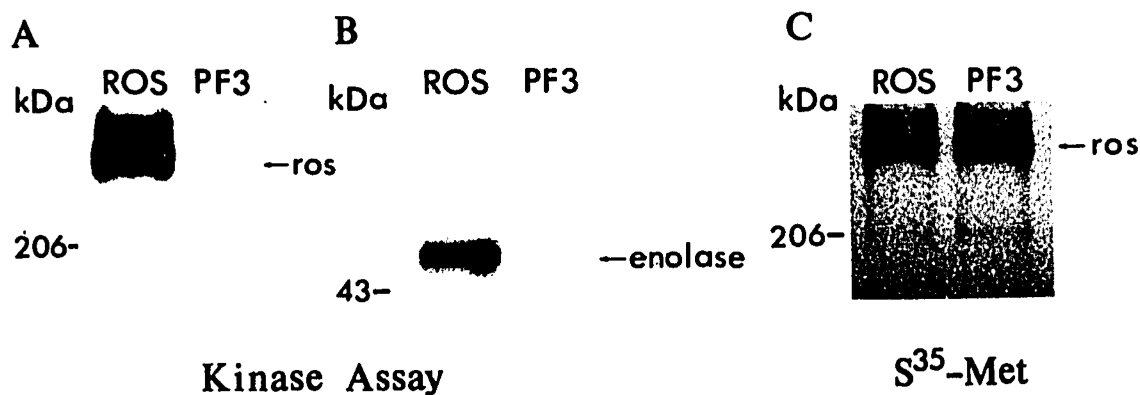


Fig. 4-10 kinase activity of ppRos in vitro and *in vivo*. pCMVppros and pECEUR2 were transfected into COS-7 cells and 48 h later, cell lysates were prepared and subjected to immunoprecipitation for *in vitro* kinase assay in the presence of an added exogenous substrate lysozyme (A). In parallel dishes, cells were treated with 250 μM Sodium Vanadate for 4 h, lysed and the α-Ros immunoprecipitate were separated in SDS-PAGE, transferred to nitrocellulose paper. The blot was hybridized with the α-pTyr antibody RC20, developed with alkaline phosphatase color reaction (B).

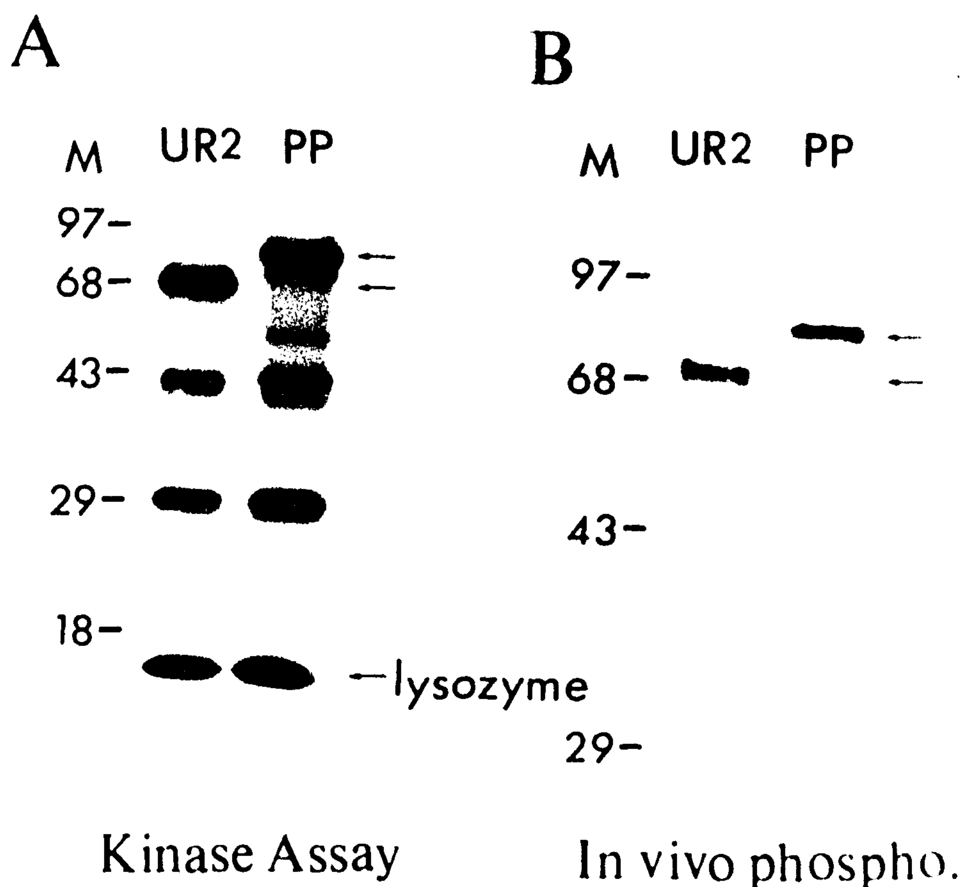


Fig. 4-11 Expression of c-Ros, ppRos and v-Ros in CEF. The pUfcros, pUppcros and pUR2 (2LTR) were individually transfected together with the help virus plasmid pUR2AV into CEF cells. After two rounds of soft agar overlay, cells from a 10-cm dish each were lysed and the extracts were precipitated with α-Ros and subjected to *in vitro* kinase assay. (A) Kinase assay of c-Ros and ppRos with an exposure time of 6 h. (C) Kinase assay of v-Ros with an exposure time of 10 min.

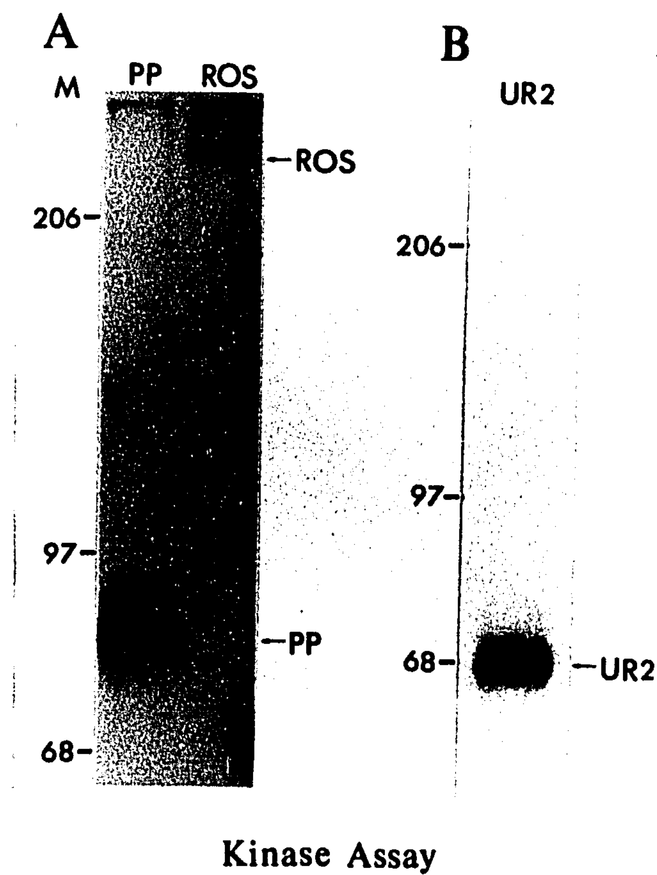


Fig. 4-12 Cell morphology and colony formation assay of CEF expressing v-Ros, c-Ros and ppRos. Fresh CEF cells were transfected with pUfcros, pUppcros or pUR2 (2LTR) together with the help virus pUR2AV. After two rounds of soft agar overlay, cell morphology was examined under the microscope (A), and colony formation assay was set up with $1X10^5$ cell per 6-cm dish and pictures were taken 12 days later (B).

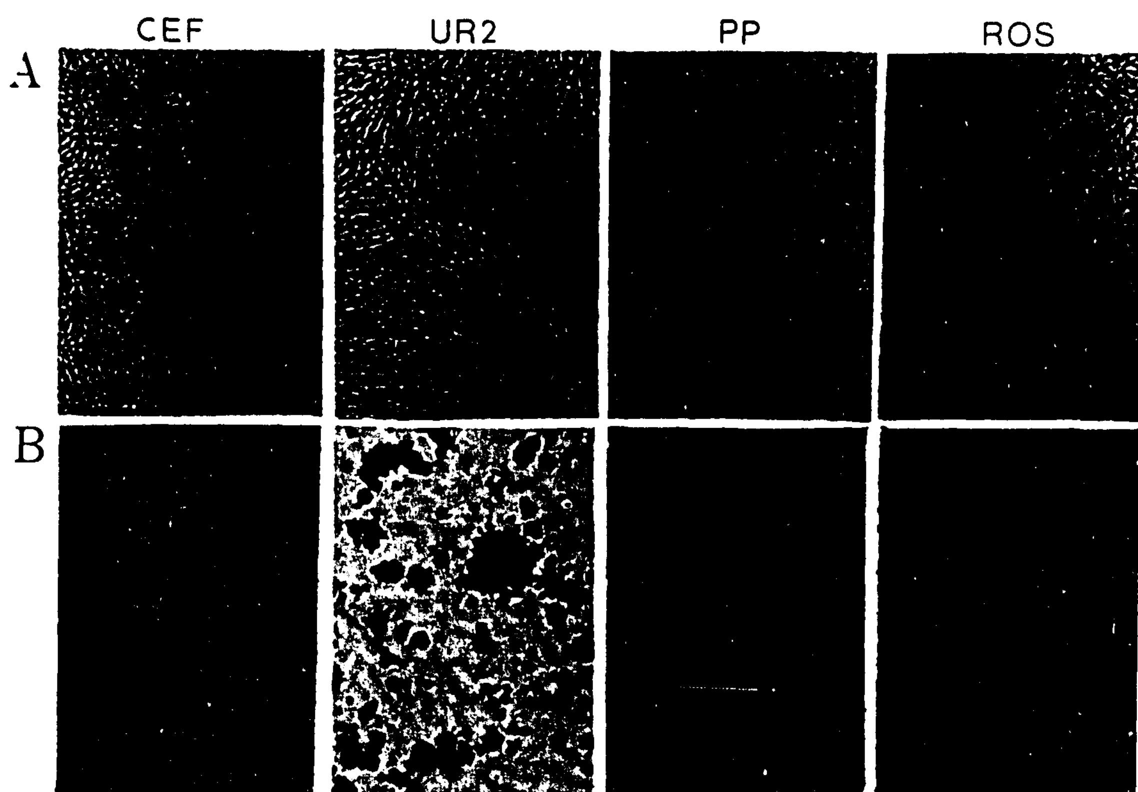


Fig. 4-13 Comparison of phosphotyrosine-containing proteins in UR2 and VCros transformed CEF and CEF transiently expressing ppRos. the cells were treated with 50 μM sodium vanadate overnight and total cell lysates were separated on SDS-PAGE, transferred to NC paper, blotted with an α-pTyr Ab RC20 and developed as described in Material and Methods (A). (B) Same amounts of the cell extracts (1.92 mg) were precipitated with Py20 and PT22-1 (a mouse α-pTyr MAb) and half of the immunoprecipitates were subjected to Western blot analysis as described in (A).

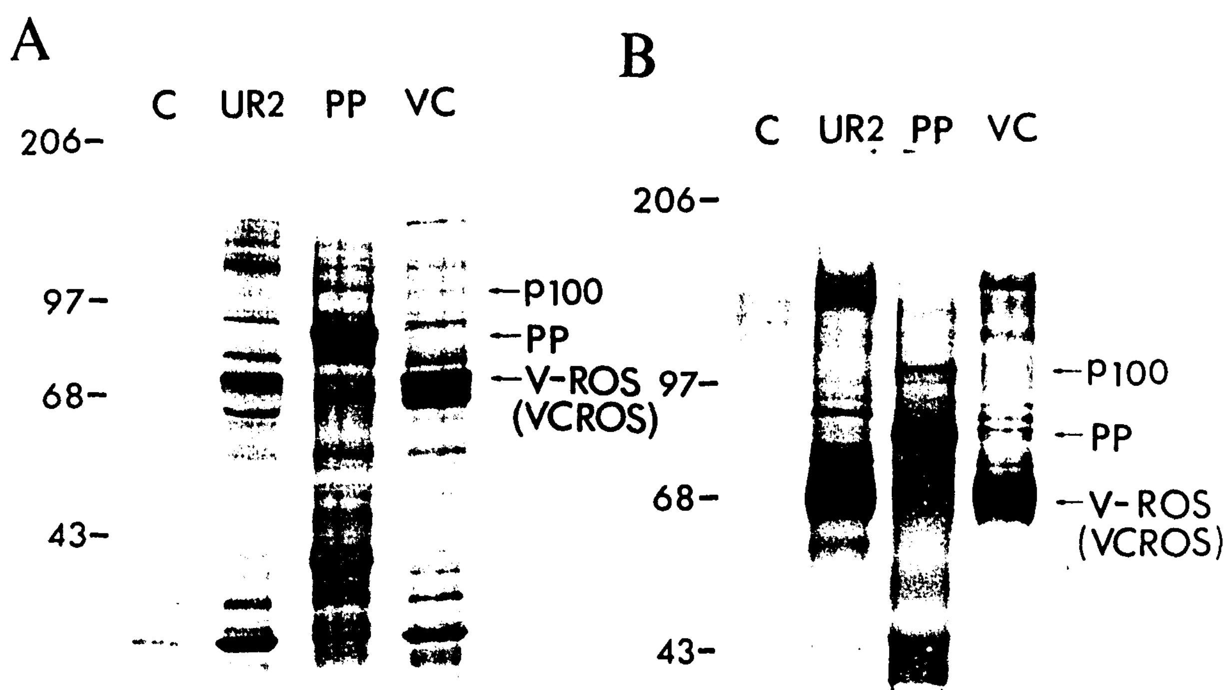


Fig. 4-14 The PI-3-Kinase activity associated with v-Ros and ppRos. Normal CEF, UR2 transformed CEF and CEF expressing ppRos were extracted with NP40 buffer. Same amounts of protein extracts (1.92 mg per assay) were precipitated with α-Ros serum and subjected to PI-3-Kinase assay. (A) The reaction products were separated in TLC plate and exposed onto the X-ray film. (B) The protein A beads from the PI-3-Kinase reaction in (A) was separated on SDS-PAGE and exposed. The v-Ros and ppRos protein bands reflect the autophosphorylation during the PI-3-Kinase assay.

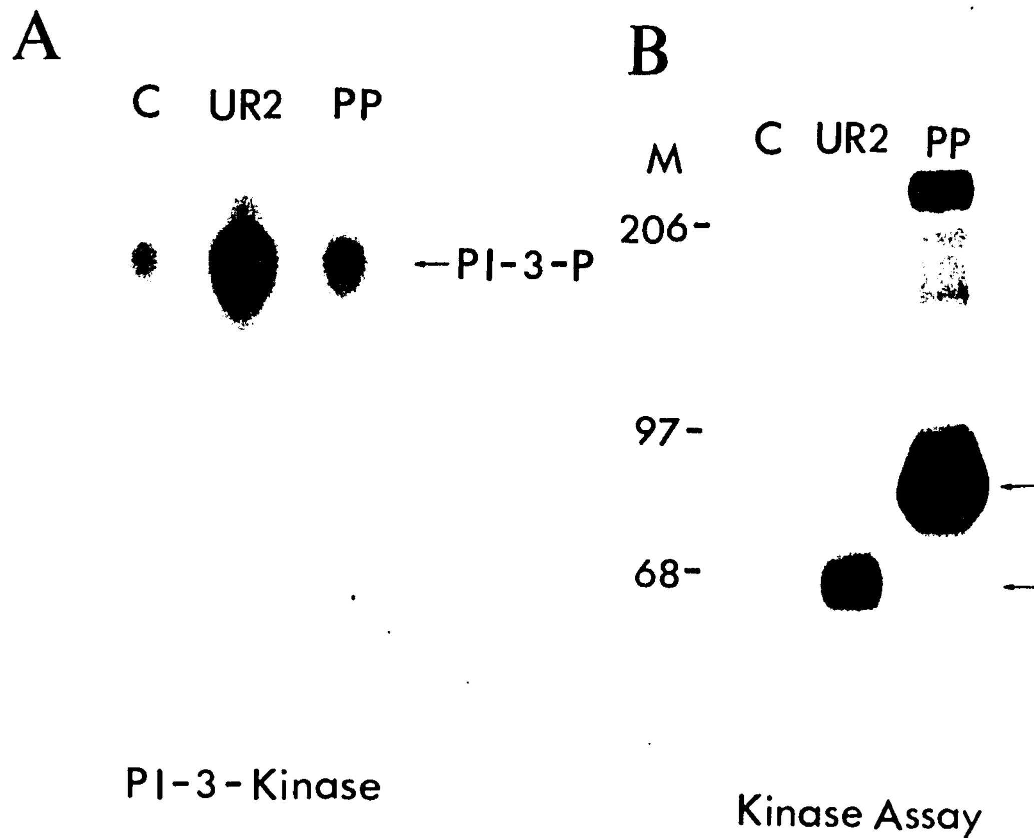


Fig. 4-15 Phosphorylation of Shc was increased in UR2 and VCros transformed CEF but not in CEF expressing ppRos. (A) α-Shc immunoprecipitates from the same amounts of cell lysates were separated on SDS-PAGE, blotted with RC20 (upper panel), and reprobed with the α-Shc antibody after eluting the first Ab (lower panel). (B) The phosphorylation was associated with the kinase activity of v-Ros. CEF transformed with the ts *ros* mutant ts251 were subjected to temperature shift in the presence of 250 μM sodium vanadate for 4 hours and α-Shc immunoprecipitates were blotted with RC20 (upper panel) and α-Shc (lower panel). (C) Tyrosine phosphorylation of ts251 v-Ros is sensitive to the temperature shift. The α-pTyr immunoprecipitates were subjected to the blotting and reaction with α-Ros.

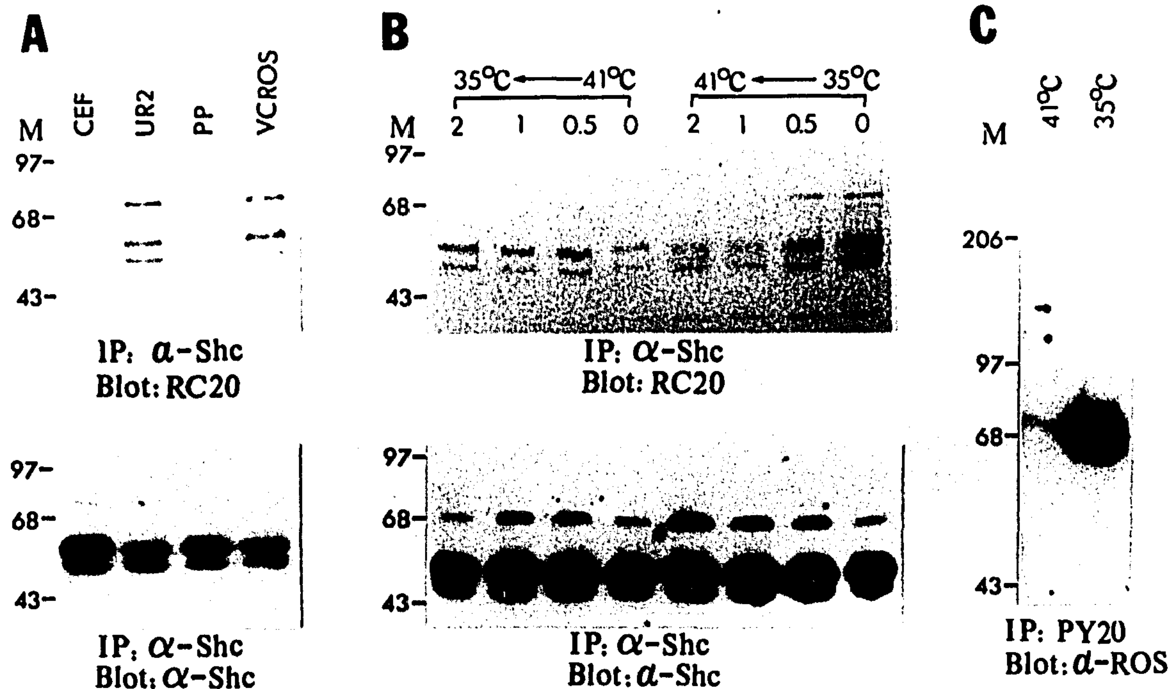


Fig. 4-16 Phosphorylation states of PLCγ1 and a 88-kDa protein. (A) and (B) is similar to Fig. 4-15 (A) and (B) except that the antibody used was α-PLCγ1 monoclonal antibody instead of α-Shc. (C) The 88-kDa protein was phosphorylated in UR2 and VCros transformed CEF.

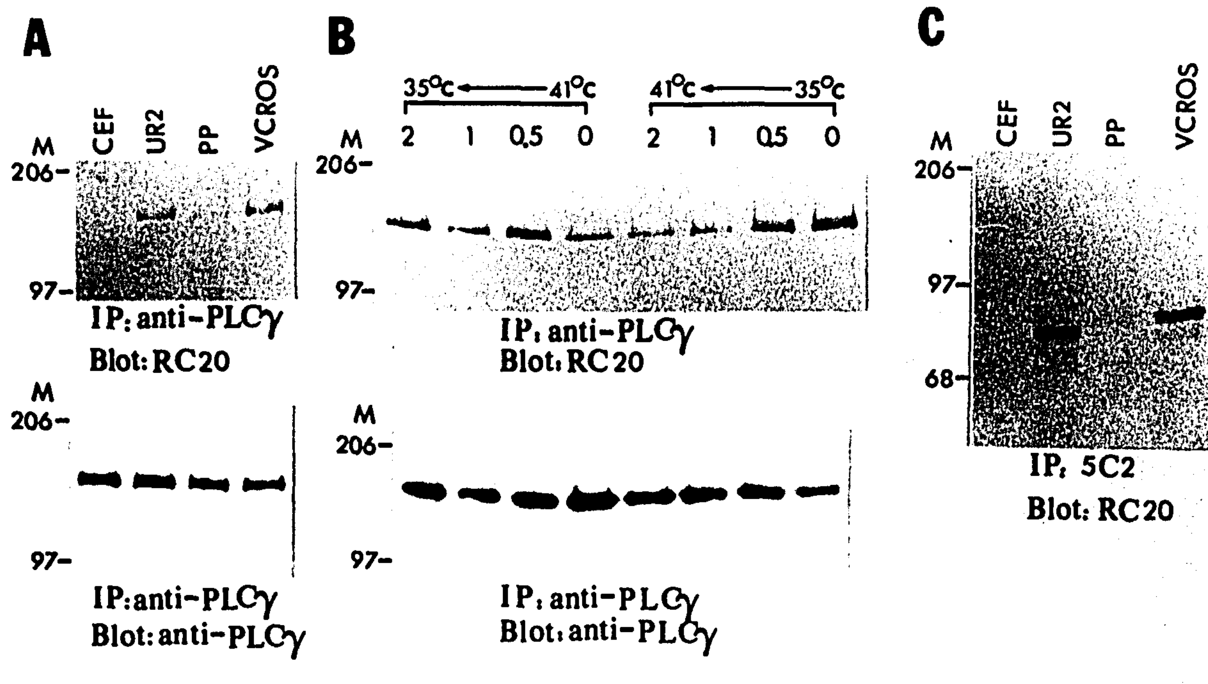


Fig. 4-17. Constructs of pCMVrosAP and pCDMrosFc. (A) A 1.5-kb SnaBI/XhoI DNA fragment encoding the catalytic domain of the human placental alkaline phosphatase was ligated into EcoRV and XhoI sites of pCMVros. The EcoRV site is located on the boundary of the c-*ros* EC and TM domain. The resulting plasmid pCMVrosAP should encode a Ros-AP fusion protein containing the entire EC domain of c-*ros* including its signal peptide but lacking its entire cytoplsmic domain. BGH is the poly A signal from the bovine growth hormone gene. Neo is the Neo^r gene driven by SV40 promoter. (B) The 5.7-kb XhoI-EcoRV fragment encoding the c-*ros* EC domain was inserted before the human IgG Fc genomic DNA in the vector pCDM8 (Kindly provided by Dr. Brian Seed), resulting in the plasmid called pCDMrosFc. The plasmid uses the CMV promoter, SV40 poly A signal and SV40 origin (SVori). SupF is used for the propagation of the plasmid in the host bacterial MC1061/p3.

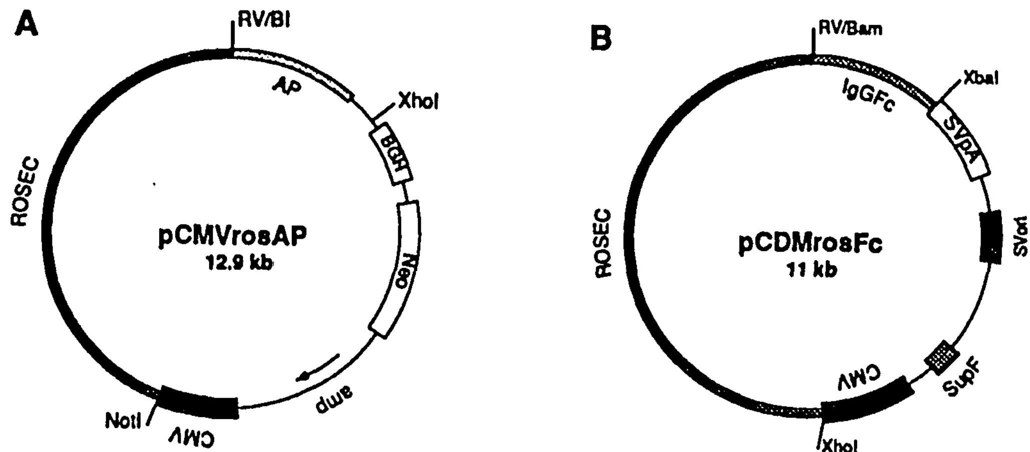
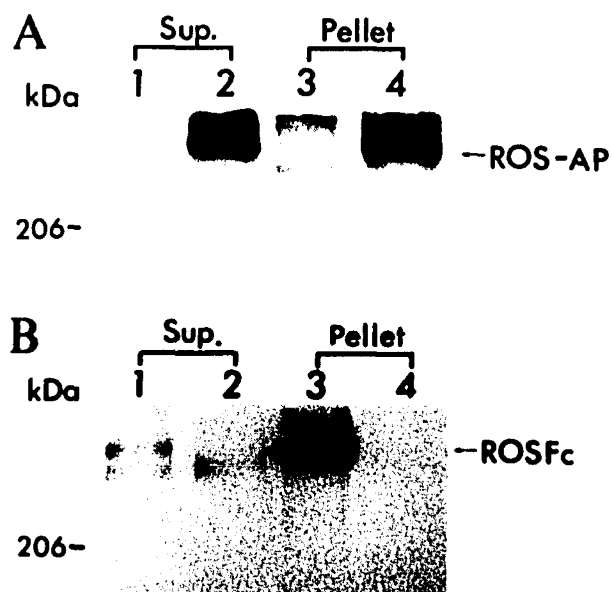


Fig. 4-18 Expression of RosAP and RosFc. (A) Establishment of the cell line expressing RosAP fusion protein. The plasmid pCMVrosAP was transfected into COS-7 cell and G418 resistant clones were purified and assayed for the expression of RosAP by [^{35}S]-Met labeling followed by the α-AP immunoprecipitation. Lanes 1 and 2 are medium supernatant sand 3 and 4 are cell extracts. Lanes 1 and 3 are COS-7 cell and Lanes 2 and 4 are RosAP36. (B) Transient expression of RosFc. pCDMrosFc was transfected into COS-7 cell and [^{35}S]-Met labeling was done after 48 h. Lanes 1 and 3 are supernatant medium and 3 and 4 are cell extracts. Lane 1 and 3 are COS-7 cell transfected with pCDMrosFc and lanes 2 and 4 are COS-7 cell only.



Chapter V. Tissue and epithelial cell-specific expression of chicken proto-oncogene c-*ros* in several organs suggests that it may play roles in their development and mature functions

Jianmin Chen, Cong S. Zong and Lu-Hai Wang

Abstract

Proto-oncogene c-*ros* codes for a RPTK sharing high homology with the *Drosophila sevenless* protein. Recent studies of c-*ros* expression in mouse by *in situ* hybridization showed that c-*ros* was expressed specifically and transiently in the epithelial cells of late embryonic renal collecting duct and intestinal villi. Those investigators suggested that c-*ros* may play a role in the development of those organelles. In the present study, we have examined the expression profile of c-*ros* in chicken by RNase protection and *in situ* hybridization with riboprobes. Our results showed that in addition to kidney and intestine, low levels of c-*ros* mRNA could also be detected in lung, testis, thymus and bursa. Expression of c-*ros* commences during middle to late embryonic stages in those organs and persists into the adult life. *in situ* hybridization revealed that expression of c-*ros* was restricted to the epithelial cells of all the tissues examined including kidney, intestine, bursa, thymus and testis. In kidney c-*ros* was detected in all the epithelial cells of the collecting ducts, in intestine it was detected in the epithelial cells of villi and the underneath crypts. Our finding of c-*ros* expression in chicken differs from those in mouse in 1) instead of transient expression during the embryonic stage, expression of c-*ros* in chicken kidney and intestine persists into the adult life and 2) expression of c-*ros* in renal collecting ducts is not restricted to its growing tips, instead it is expressed in the entire epithelial layer of the ducts. Our results suggest that c-*ros* may play a role not only in the initial induction events in the organogenesis, but also in the mature function of those

organs.

Introduction

c-ros was shown initially to be expressed only in the chicken kidney by Northern blot analysis (Neckameyer et al., 1986). Subsequent analysis indicated that in addition to kidney, low level of *c-ros* expression could also be detected in chicken bursa, thymus, gonad and brain (Chen et al., 1991). In rat, it was shown to be expressed in kidney, lung, heart and testis (Matsushime & Shibuya, 1990). Most recent studies have identified intestine in addition to kidney as the major expressing tissue in mouse (Sonnenberg et al., 1991; Tessarollo et al., 1992). Surprisingly, *c-ros* RNA could not be detected in rat intestine by Northern analysis (Matsushime & Shibuya, 1990). The major *c-ros* mRNA is a 8.3-kb transcript although 4.5-5-kb mRNA presumably resulting from alternative splicing have been observed in rat and mouse testis (Matsushime et al., 1990; Sonnenberg et al., 1991; Tessarollo et al., 1992). Recent studies of *c-ros* expression in mouse by *in situ* hybridization indicated that *c-ros* was mainly expressed in the epithelial cells of the kidney collecting ducts, and the villi and crypts of intestine (Sonnenberg et al., 1991; Tessarollo et al., 1992). The expression of *c-ros* was found to be transient, mainly during embryonic and neonatal stages and restricted to certain epithelial cells in those tissues. The onset of *c-ros* expression in embryonic kidney development is coincident with the occurrence of a well known reciprocal epithelial/mesenchymal induction, a major event in kidney organogenesis. The *c-ros* was suggested to play a role in that event presumably via stimulation by its ligand.

To inquire the function of *c-ros*, we used the RNase protection assay and mRNA *in situ* hybridization techniques to study its tissue and cell type-specific expression in chicken. We found that *c-ros* expression is restricted to the epithelial cells of several organs. However, unlike the situation reported in mouse where *c-ros* is only transiently

expressed during embryonic and neonatal stages, expression of c-*ros* in most of the chicken organs persists into the adult life. In addition, c-*ros* was found to be expressed in immune organs, thymus and bursa, commencing at late embryo and peaking around 10 days after hatching. Our results suggest that in addition to the possible role in embryonic organogenesis of certain organs, c-*ros* may also play a role in the mature physiological functions of those organs in chickens.

Results

Tissue-specific expression of c-*ros* in chicken

Our previous experiments by slot blot hybridization using riboprobes showed that c-*ros* was expressed at the highest level in chicken kidney, and at much lower levels in thymus, bursa, gonad and brain (Chen et al., 1991). In the current study, we used RNase protection assay (RPA), a much more sensitive technique, to reexamine the c-*ros* expression. A 310 nucleotide (nt) ^{32}P-labeled antisense riboprobe containing 3' 231 nt of the cDNA clone 5b was used (Chen et al., 1991; probe B in Fig 1). This probe derived from the EC domain of the c-*ros* was chosen to minimize the possibility of c-*ros* hybridization with other PTK messages. RNA samples prepared from different organs of 1-month old chickens were subjected to RNase protection assay using the probe B. As shown in Fig. 5-2A, an expected 231 nt RNA species was protected by kidney and intestine RNAs, but no protection was detected by RNAs from other tissues including thymus, bursa and lung. Since our previous results by slot blot hybridization were obtained from 10-14 day old chickens, we performed the RPA with RNA samples from chickens of different ages including those from embryos (Fig. 5-2B). Again, kidney and intestine expressed the highest level and a dramatic increase of c-*ros* mRNA in intestine occurred shortly after hatching. Confirming our previous results, much lower amounts of c-*ros* mRNA could be detected in bursa, thymus and lung of 9- to 14-day old chickens, but not in embryonic heart and lung (Fig. 5-2B). However, differing from our previous result (Chen et al., 1991), no c-*ros* expression could be detected in the brain of various stages (data not shown). Since the *Drosophila* homolog of c-*ros*, *sevenless*, is mainly expressed in the fly compound eye (Rubin, 1990), we also examined c-*ros* expression in chicken eyes. No signal could be detected at any stages of chicken eyes (data not

shown.), indicating that the functions of the *c-ros* and *sevenless* genes have diverged even if they have evolved from a common ancestral gene.

Since c-*ros* RNA was detected in chicken gonad and in the mature testis of rat and mouse, where the size of the c-*ros* mRNAs was shown to be 4.5- to 5- kb instead of the 8.3-kb RNA found in other organs, several probes corresponding to various regions of the c-*ros* cDNA (Fig. 5-1) were used to detect c-*ros* mRNA in chicken testis. In contrast to the high level expression of c-*ros* mRNA in mouse testis (Tessarollo et al., 1992), only very low levels of c-*ros* RNA could be detected with probes A, B, C, D (Fig. 5-3, and other data not shown) in 13- and 20-week old chicken testis and the expression level in the latter was higher. Since the testis c-*ros* mRNA contains the sequences represented by probes B, C, D, it must contain EC and cytoplasmic domains.

Northern blot analysis by riboprobe N (Fig. 5-1) showed that similar to kidney and intestine, the size of c-*ros* mRNA in bursa was about 8.3-kb (Chen et al., 1991; and Fig. 5-4). Due to very low level expression, a distinctive c-*ros* RNA species could not be detected in lung, thymus and testis by Northern analysis.

Kinetic analysis of c-*ros* expression in various organs

RNAs from different tissues at different stages of chicken development were analyzed. The kidney c-*ros* mRNA could be detected as early as in 13-day embryo. Its level continued to increase until around 9 to 14 days after hatching when it decreased slightly and lasted throughout the adult life (Fig. 5-5A, 7). This is contrary to the observations in mouse kidney where c-*ros* expression abolished completely within 2 weeks after birth (Sonnenberg et al., 1991; Tessarollo et al., 1992). A previous study reported on the detection of the c-*ros* mRNA in rat kidney up to 15-weeks (Matsushime & Shibuya, 1990).

In intestine, expression of c-*ros* was detected first in 15-day embryos. It increased sharply after hatching and continued to increase for another 2 to 3 weeks when it decreased slightly but maintained at a constant level thereafter (Fig. 5-5B and 5-7). Again this is different from the situation in mouse intestine where c-*ros* expression could not be detected beyond 3 weeks after birth (Sonnenberg et al., 1990; Tessarollo et al., 1991).

In lung, a very low level of c-*ros* RNA was detected in 18-day embryos. c-*ros* RNA expression in lung increased after hatching and persisted into the adulthood (Fig. 5-6A, 5-7). Expression of c-*ros* in bursa was similar to that of lung, but its level peaked within two weeks post hatching and declined somewhat during adulthood (Fig. 5-6B). The highest level of c-*ros* expression in thymus was detected in 9-day old chicks (Fig. 5-6C). c-*ros* could not be detected at various developmental stages of heart (Fig. 5-6C).

Fig. 5-7 shows the result of direct comparison of 9-day old versus 5-month old chickens. The levels of c-*ros* RNAs in kidney and intestine are comparable in both organs from either young or adult chickens whereas in lung the older chicken appears to have a higher level. The c-*ros* RNA expression in various organs was summarized in table 5-1. From the above result, it is clear that the expression of c-*ros* is tissue-specific and under stringent temporal regulation.

Cell type-specific expression of c-*ros*

in situ RNA hybridization was used to study the spatial distribution of c-*ros* mRNA in those positive organs. The ^{35}S-labeled antisense probe E (Fig. 5-1) is 480 nt in length and located about 300 nt upstream from the transmembrane domain (Chen et al., 1991). The sense probe E was used as the negative control for all the experiments. Fig. 5-8 compiles the results of analysis with kidney, intestine and bursa.

For kidney, the earliest time point we examined was 16-day embryo. c-*ros* mRNA

could be clearly detected at the tips of the collecting ducts (Fig. 5-8A). No signal was present when the sense probe was used (Fig. 5-8D). As the kidney development continued and more collecting tubules were branching out, c-*ros* RNA was found to be present in the whole epithelial layers of the collecting duct system (Fig. 5-8B, C and G). Within 2 weeks after hatching, silver grain signals of the c-*ros* mRNA could be seen along the branched tubules and in the epithelial cells of the larger tubules (Fig. 5-8B and C), and even in the epithelium of the proximal part of ureters (Fig. 5-8G). In 10-week and 20-week old chicken kidney, the similar pattern of c-*ros* expression was also observed (data not shown). Therefore, the expression profile of c-*ros* is similar in chicken and mouse kidneys at embryonic stages, but it differs postnatally in the two species (Sonnenberg et al., 1990; Tessarollo et al., 1991).

In intestine, c-*ros* RNA was detected in epithelial cells of villi and crypts which regenerate the villi epithelium (Fig. 5-8E, F). No difference in the level of expression was observed between villus and crypt epithelial layers in all time points examined. Consistent with the RNase protection data, c-*ros* mRNA levels were very low in embryo but increased rapidly after hatching. A similar pattern of expression was observed in embryonic and newborn mouse intestine. While c-*ros* could hardly be detected beyond 3-week old mouse intestine (Sonnenberg et al., 1991; Tessarollo et al., 1992), a similar level of c-*ros* mRNA was observed in the 10-week and 20-week old chicken intestine with a similar pattern of spatial distribution (data not shown).

In bursa and thymus, the two major immune organs of chicken, the c-*ros in situ* hybridization signals were much lower than those in intestine and kidney. In bursa, only the single layers of surface epithelial cells surrounding the lymphoid tissues were positive for c-*ros* expression (Fig. 5-8H, I), while other epithelial cells including the interfollicle

epithelial cells are negative. A similar pattern but even weaker signals were seen for thymus (data not shown). Although lung was also positive for c-*ros* mRNA in RNase protection assay, no signal could be detected in all the time points examined by *in situ* hybridization. Therefore, for all the organs analyzed, c-*ros* expression appears to be restricted to the epithelial cells.

Discussion

In this study, we examined chicken proto-oncogene c-*ros* expression during development in various organs by RNase protection assay and mRNA *in situ* hybridization. Our results clearly show that c-*ros* is specifically expressed in kidney, intestine, lung, bursa, thymus and testis and is mainly restricted to specialized epithelial cells in those organs. The onset as well as tissue- and cell type-specific expression of chicken c-*ros* was similar to that of mouse c-*ros*. However, c-*ros* expression in the two species differs in the following aspects: 1) in contrast to transient and restricted expression of c-*ros* in the embryonic and neonatal mouse kidney and intestine (Sonnenberg et al., 1991; Tessarollo et al., 1992), chicken c-*ros* expression persists in adult kidney and intestine. 2) rather than being restricted to the growing tips of mouse renal collecting ducts, chicken c-*ros* mRNA is present in the tips as well as the whole epithelial layers of the renal collecting tubules. 3) unlike chicken, the c-*ros* mRNA was undetectable in mouse thymus (Sonnenberg et al., 1991). Our observation on the tightly regulated temporal and spatial expression of chicken c-*ros* in various organs suggests that c-*ros* may play an important role not only in the early development but also in the mature function of those organs.

It was previously shown that vertebrate kidney development was characterized by reciprocal epithelial/mesenchymal induction events in which the mesenchymal cells induce the collecting duct of ureters to branch and differentiation of epithelial cells and in turn the epithelial cells of the branched collecting tubules induce the differentiation of mesenchyme into the epithelial cells of renal tubules and glomeruli and the forming of the so called renal units (Gilbert, 1988; Kratockwil, 1983). Studies of c-*ros* expression in mouse kidney have demonstrated that in early embryonic development c-*ros* RNA is localized in the Wolffian duct and the whole ureter, later it becomes restricted to the tips of the collecting

tubules where the inductive events occur (Sonnenberg et al., 1991; Tessarollo et al., 1992). Because of the transient and restricted expression of c-*ros* in mouse embryonic kidney and the demonstration of the expression of a PTK c-*ros* at a place where the inductive events occur, it was proposed that c-*ros* might be the receptor to receive the inductive signals from the surrounding mesenchyme resulting in the induction of branching and differentiation of the epithelial cells of collecting tubules (Sonnenberg et al., 1991; Tessarollo et al., 1992).

Our study of c-*ros* expression in chicken kidney shows that at embryonic stage c-*ros* expression is restricted to the tips of the forming collecting tubules, and may have a similar function as that proposed for the mouse. The onset of c-*ros* expression in chicken kidney coincides with the differentiation of metanephros (Rol'nik, 1970) and may be involved in its development. In contrast, the expression of c-*ros* in chicken kidney increases after hatching and reaches the highest level between 9 and 14 days, it then declined slightly but persists throughout the adulthood. The increased c-*ros* expression after hatching detected by RPA was confirmed by *in situ* hybridization which clearly showed that c-*ros* mRNA was present in the whole epithelial cells of branched collecting ducts and even in large tubules (Fig. 5-8G). It seems that the expression level of c-*ros* per cell was similar and the increased c-*ros* RNA during kidney development is probably due to increased number of expressing cells. A similar pattern of expression was also observed in 10- and 20-week old chicken kidneys (data not shown). Our results suggest that although chicken c-*ros* may have a similar induction and/or differentiation function at embryonic stage as that of mouse c-*ros*, it is likely that it may have some other specialized function in mature kidney in addition. The molecular basis for the vertebrate metanephric induction is largely unknown. In addition to c-*ros*, some potential transcription

factors and nuclear proteins were found to be specifically expressed in cells of collecting tubules. L-myc (Mugrauer & Ekblom, 1991), Hox2.3 (Kress et al., 1990), Pax2 (Dressler et al., 1990) and WT-1 (Pritchard-Jones et al., 1990) are expressed along the ureters, collecting tubules or its surrounding mesenchyme. The possible relationship of c-*ros* with those potential transcription factors remains to be determined.

Our *in situ* hybridization data show that c-*ros* is expressed in the villi epithelial cells and its regenerator, the crypts, in chicken intestine (Fig. 5-8). The pattern of expression is similar to that in mouse intestine (Sonnenberg et al., 1991; Tessarollo et al., 1992). Again, the time course of c-*ros* expression in chicken intestine is different from that in mouse in that c-*ros* RNA can still be detected at a relatively high level in the intestine of 20-week old chicken by RPA and its spatial expression pattern is similar to that of young chicks (data not shown). Although epithelial/mesenchymal induction has also been demonstrated in the development of intestine (LeDouarin et al., 1968), the onset and persistent expression of chicken c-*ros* in intestine does not coincide with this event. Instead, c-*ros* may be involved in the differentiation and/or maintenance of the intestinal villi epithelial cells. It is uncertain whether there is a fundamental difference in the development of kidney and intestine between chicken and mouse. Unlike the mouse kidney, chicken kidney is not divided into medulla and cortex, instead, the urinary tubules make up the bulk of the renal parenchyme. The discrepancy of c-*ros* expression in adult kidney and intestine of chicken and mouse could reflect the species divergence such that c-*ros* is no longer required for the mature functions of those mouse organs. Alternatively, c-*ros* function could be replaced by another mouse gene that can not be detected by the c-*ros* probe under the conditions used. However, the discrepancy of the lack of the c-*ros* expression in rat intestine and expression of c-*ros* in mature rat kidney (Matsushime &

Shibuya, 1990) versus the c-*ros* expression observed for mouse intestine and kidney (Sonnenberg et al., 1991; Tessarollo et al., 1992) is surprising. It is not clear whether those observations reflect the genuine species variation.

The detection of c-*ros* mRNA in the epithelial cells of the two major lymphoid organs of chicken represents another difference among different species because c-*ros* expression was not observed in mouse thymus (Sonnenberg et al., 1991). Interestingly, the interaction of surface epithelial cells with the surrounding mesenchyme was proposed to be important for the generation of the follicle cortex of bursa (Ackerman & Knouff, 1959). The function of c-*ros* in these chicken lymphoid organs remains unknown. However, the epithelial expression of c-*ros* in these organs again underscores the potential functional role of c-*ros* in this particular type of cells.

The function of c-*ros* and *Drosophila sevenless* protein seems to have diverged during evolution because the c-*ros* expression can not be detected in the eyes of chicken and mouse. The ligand of *Drosophila sevenless* protein has recently been identified to be the *boss* gene product (Krämer et al., 1991) which is expressed on the surface of neighboring cells of R7 expressing the *sevenless.* The *boss* protein can bind to the *sevenless* RPTK, activate its kinase and the internalization of the protein complex (Hart et al., 1993). The close sequence and structural relationship between c-*ros* and *sevenless* as well as the implicated role of c-*ros* in epithelial/mesenchymal induction events in organogenesis suggest that c-*ros* might have a ligand similar to the *Drosophila boss* gene product.

Acknowledgement

We are grateful to members of Dr. Lu-Hai Wang's laboratory for their helpful discussions and suggestions. Special thanks goes to Dr. Shudong Zong and Wei Chen for their help in the *in situ* hybridization experiments. This work was supported by NIH grant CA 29339.

Table 5-1. Expression profile of chicken c-*ros* in different tissues

Organ	Embryo	Neonatal	Adult
kidney	+	+++	++
intestine	+	+++	++
lung	+/-	+/-	+
bursa	+/-	+	+
thymus	nt	+	+/-
heart	-	-	-
testis	nt	+/-	+
ovary	nt	-	-
brain	-	-	
eye	-	-	-
liver	-	-	-
spleen	nt	-	-

The table represents summary of the RPA analysis of various chicken tissues. nt, not tested.

Fig. 5-1. Different c-*ros* probes used for RNA analysis. Probes A, B, C, and D were used for RPA and probes N and E were used for Northern Blot analysis and mRNA *in situ* hybridization respectively. Arrowed lines indicate the position and size of individual probes. The translational initiation site (ATG), transmembrane domain (TM), protein tyrosine kinase domain (PTK) and stop

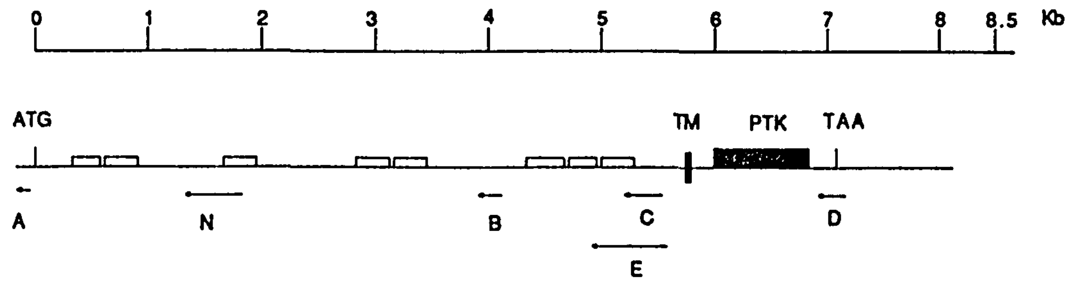


Fig. 5-2. RNase protection analysis of RNA samples from different chicken tissues. For each assay, 30 μg of total RNA and 1X10^5 cpm of probe B were used. (A) RNA samples from various tissues of a 1-month old chicken. (B) RNA samples from various embryonic and chicken tissues. M is molecular weight marker ([γ-^{32}P]-labeled ΦX174+HaeIII). The sizes are in nucleotides. Arrows point to the protected products. Emb. stands for embryo tissues and d stands for days at different stages. The abbreviations for tissues are: T, thymus; St, stomach; M, muscle; Lu, lung; Li, liver; K, kidney; I, intestine; E, eye; Bu, bursa; Br, brain; Sp, spleen. C strands for control yeast RNA (30 μg) and P for the probe only. The exposure time was in day. The quantity and quality of RNA samples were checked by gel electrophoresis and detection of the 28S and 18S rRNAs as shown in Fig. 5-5.

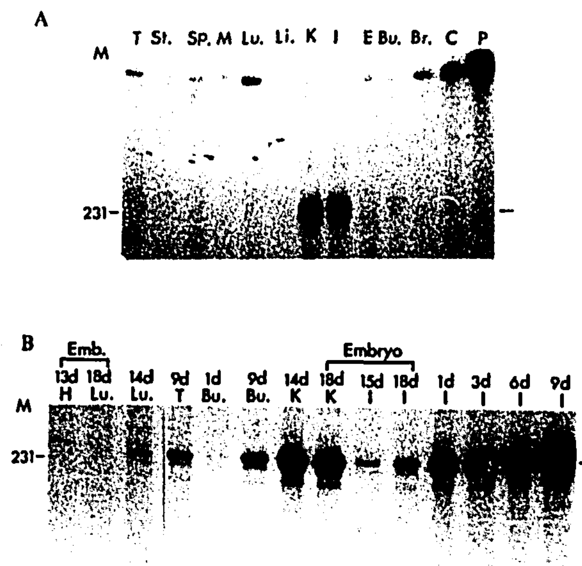


Fig. 5-3. RPA of RNA samples from chicken testis. 30 μg of total RNA was used for each assay. The exposure time was 3 days. RNAs from 10-week and 20-week old chicken testes were used for RPA with probes B, C, and D as indicated.

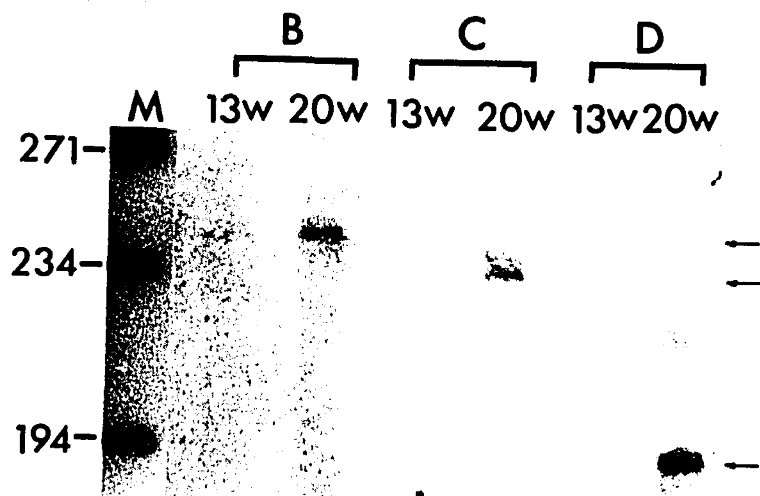


Fig. 5-4. Northern blot Analysis of c-*ros* mRNA in different chicken tissues. 6 μg of poly(A)+ RNA was analyzed using probe N. Abbreviations for tissues are: K, kidney; I, intestine; B, bursa; L, lung; and S, stomach. The positions of 28S and 18S rRNA are marked. The arrow indicates the 8.3-kb c-*ros* transcript. The exposure time was 3 days.

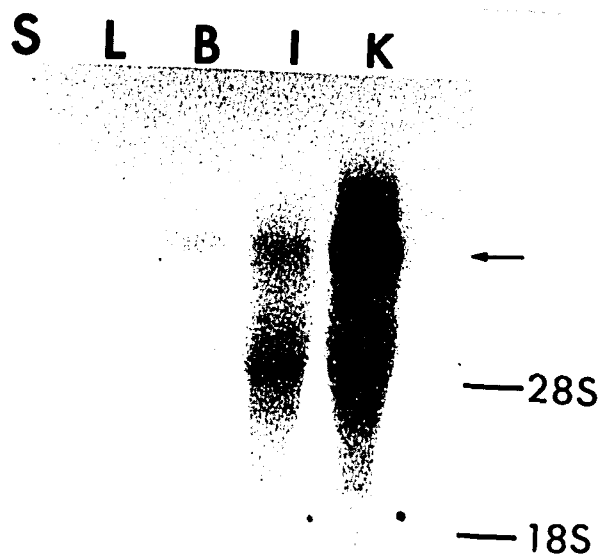


Fig. 5-5. Kinetic study of c-*ros* expression in kidney and intestine. Same amount of RNAs and probe B as in Fig. 5-2 were used. (A) RPA of kidney RNA samples. (B) RPA of intestine RNA samples. Arrows indicate the position of protected products. The abbreviations are: Emb., embryo; ck, chicken; d, days; w, weeks after hatching; c, control yeast RNA; p, probe only. The exposure time was 20 h. 0.5 μg of each of the above RNA samples were separated on 1% agarose/TAE gel , stained with ethedium bromide and photographed as shown under each RPA. The positions of 18S and 28S rRNAs were marked.

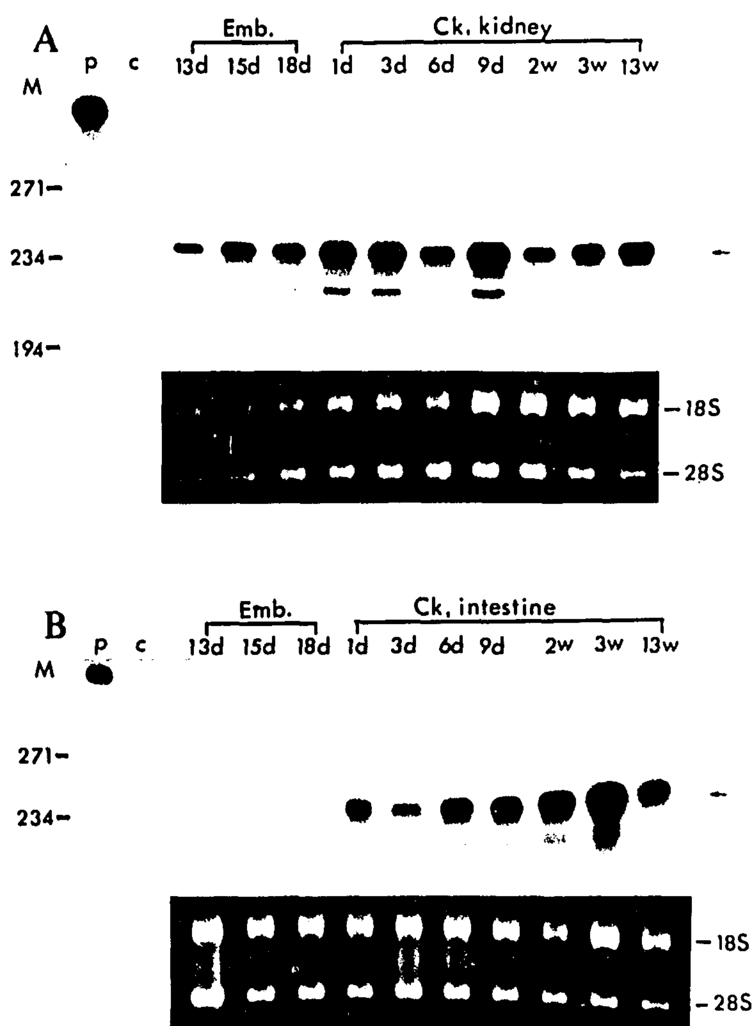


Fig. 5-6. Kinetic study of c-*ros* expression in lung, bursa and thymus. (A) RPA of lung RNA samples. (B) RPA of bursa RNA samples. (C) RPA of thymus and heart samples. E stands for embryo and Ck for chicken. The exposure time was 3 days.

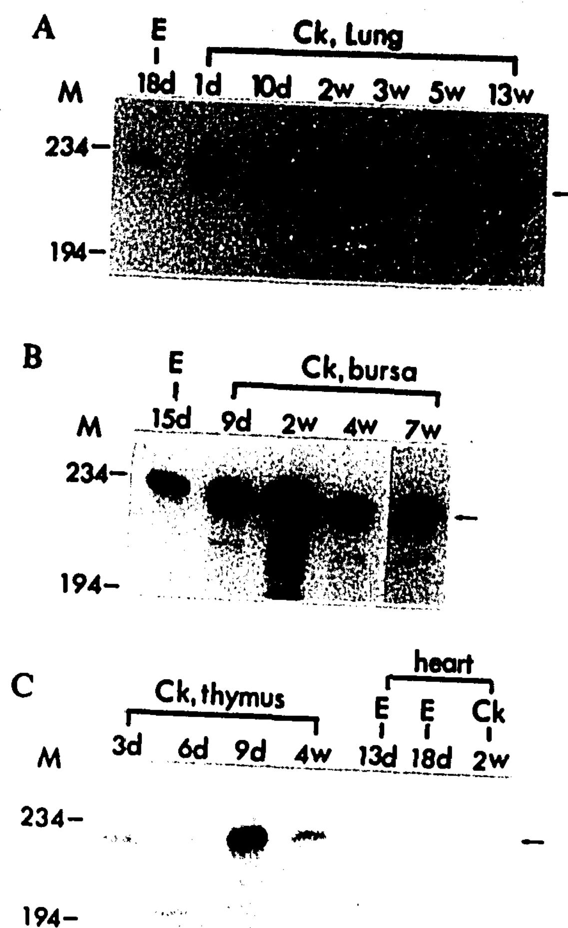


Fig. 5-7. Expression of c-*ros* in young and adult chickens. 30 µg each of the RNA samples (kidney, intestine and lung) from 9-day old chicks and a 20-week old chicken were subjected to RPA using probe B. The exposure time was 1 day. Arrow indicates the position of the protected product.

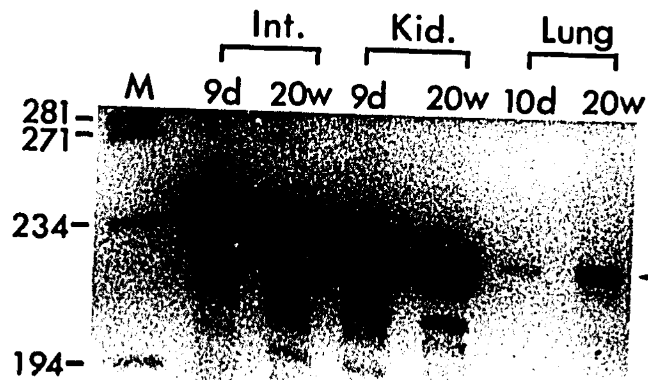
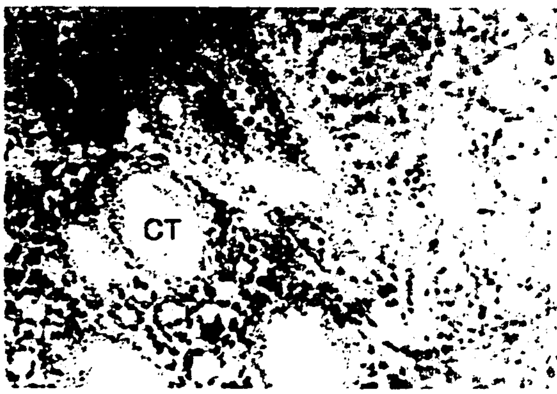


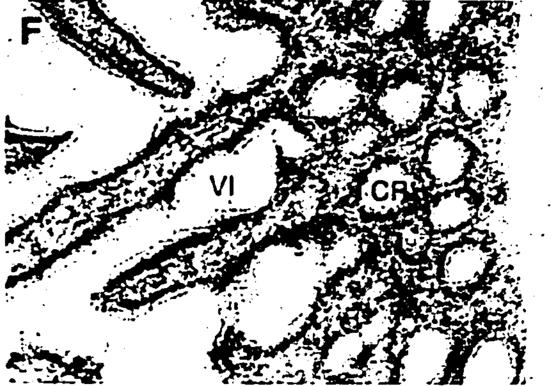
Fig. 5-8. c-*ros* mRNA *in situ* hybridization in chicken kidney, intestine and bursa. 5 micron cryosections of different tissues were subjected to *in situ* hybridization as described in Materials and Methods. Exposure time was 20 days. A. 16-day embryonic kidney, B. 1-day old chick kidney, C. 14-day old chick kidney, D. 16-day embryonic kidney with the sense probe, G. 9-day old chick kidney showing the large tubule. E and F, the dark field and bright field, respectively, of 14-day old chick intestine, H and I. the dark field and bright field, respectively, of 10-day old chick bursa. CT and CD, for collecting tubule and collecting duct; VI for villi and CR for crypts; SE for the surface epithelium and IFE for interfollicle epithelium. Bar size in (I) is 20 μM.

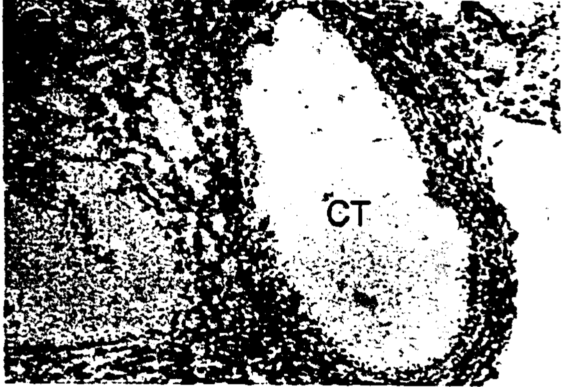
B

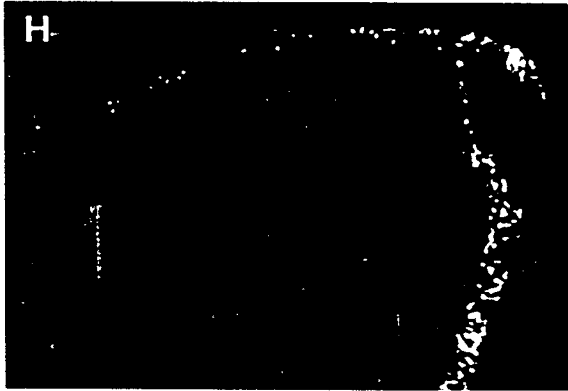

CT

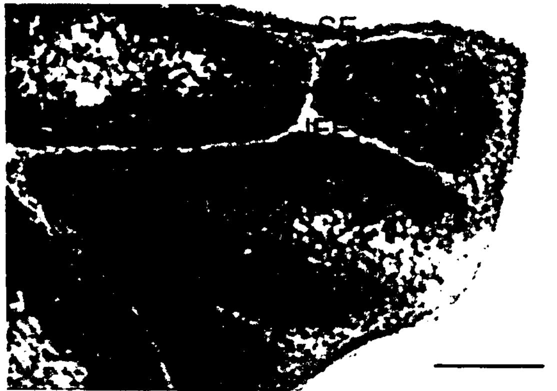
D

E


F
VI
CR


CT


H



Chapter VI. Cloning and characterization of the chicken c-*ros* promoter

Jianmin Chen, Jie Tong and Lu-Hai Wang

Abstract

Our previous results showed that c-*ros* was specifically expressed in chicken kidney, intestine, lung, bursa and thymus, and its expression was detected only in certain epithelial cells in these organs. Furthermore, the expression of c-*ros* in those organs is developmentally regulated, both temporally and spatially. In order to explore the molecular basis for the regulation of c-*ros* expression, we have cloned and characterized the chicken c-*ros* promoter. The 5' non-coding c-*ros* cDNA was isolated by rapid amplification of cDNA ends (RACE) method, resulting in the extension of another 35 nucleotides beyond our previously published cDNA sequence (Chen et al., 1991). The most 5' c-*ros* cDNA was used as a probe to screen a chicken genomic DNA library, 3 clones containing the 5' c-*ros* exon were isolated. A 2.5-kb HindIII fragment in one of those clones was identified to contain the most 5' sequence of the c-*ros* cDNA. Primer extension and RNase protection analyses were used to map the transcription initiation sites for the c-*ros* mRNAs in kidney and intestine. Sequence of the 1.3-kb region upstream of the initiation site revealed the TATA and CAAT boxes at 26 and 54 nucleotides upstream of the initiation site. In addition, potential transcription factor binding sites and several direct and inverted repeats were found further downstream. The 1.3-kb DNA when placed upstream of the CAT gene was shown to be functionally active. Serial deletion of this putative c-*ros* promoter has allowed us to define a minimum c-*ros* promoter and to identify potential positive regulatory regions.

Introduction

The proto-oncogne *c-ros* cDNAs have been molecularly cloned from rat (Matsushime & Shibuya, 1990), chicken (Chen et al., 1991) and a human glioblastoma cell line (Birchmeier et al., 1991). It was recently shown that in rat, *c-ros* RNA could be detected in heart, lung and testis in addition to kidney and alternative splicing of *c-ros* mRNA was observed in different organs (Matsushime & Shibuya). By more sensitive methods of detection such as RNase protection analysis and RNA in situ hybridization, *c-ros* was found to be expressed in kidney, intestine, lung and testis of mouse (Sonnenberg et al., 1991; Tessarollo et al., 1992). The expression was transient and restricted to the epithelial compartments in kidney and intestine, and probably also in lung. Since the time course of *c-ros* expression in mouse kidney and intestine was coincident with the reciprocal epithelial/mesenchymal induction in those organs, a role of induction in their organogenesis was proposed for the mouse *c-ros* (Sonnenberg et al., 1991; Tessarollo et al., 1992).

Our recent study showed that *c-ros* RNA could be detected at considerable levels in chicken kidney and intestine and at much lower levels in lung, thymus and bursa. A low level of mRNA could also be detected in testis at different stages, but whether alternative splicing exists is unknown. Our RPA and mRNA in situ hybridization results confirmed that *c-ros* expression was developmentally regulated and restricted to the epithelial cells of renal collecting tubules and epithelial cells of intestinal villi and crypts. However, its expression profile is different from that in mouse in that *c-ros* is expressed in the entire epithelial layer of renal collecting tubules as well as intestinal villi and crypts commercing at late embryo and persisting thereafter. *c-ros* RNA was also detected in the epithelial cells of the bursa capsule. Thus, we propose that chicken *c-ros* may play a role

in the epithelial differentiation as well as mature function of those organs.

In order to explore the molecular basis for the regulation of the tissue- specific expression of chicken c-*ros*, in this study, we have cloned the complete 5' non-coding sequence of the c-*ros* cDNA and isolated the corresponding genomic DNA. The transcriptional initiation site was mapped by primer extension and RPA. The putative promoter region for chicken c-*ros* contains canonic TATA-like sequence and CAAT box as well as potential transcription factor binding sites. The putative promoter was shown to be active in promoting CAT gene expression. We have defined the basic promoter as a 300 bp sequence upstream of the transcription initiation site. This work provides the basis for further studying the regulation of tissue-specific expression of the proto-oncogene c-*ros*.

Results

Cloning of 5' c-*ros* cDNA

Our previous chicken c-*ros* cDNA clones contain a 158 nucleotide non-coding region at its 5' end (Fig. 3-3). In order to map the mRNA initiation site as well as to map the promoter region, 5' RACE method (Frohman et al., 1988) was used to isolate the most 5' cDNA sequence. The RACE PCR was done as described in the Materials and Methods. 4 positive clones were selected for sequencing analysis. Clone 13 extends 5' cDNA sequence for another 35 nucleotides while clone 20, 26 and 47 extend only 31 nucleotides. The sequence is shown in Fig. 6-2 as the italic and underlined nucleotides. This 5' cDNA sequence was confirmed by sequencing the corresponding genomic region (see below) to exclude the possibility of PCR artifact. Either first nucleotide G in clone 13 or the first nucleotide A in other three clones could be the mRNA initiation site since RNase protection assay and primer extension experiments map the initiation site(s) to these positions (see below).

Isolation of genomic clone corresponding to the most 5' c-*ros* cDNA

Genomic clones corresponding to the 5' c-*ros* cDNA were isolated to identify the potential c-*ros* promoter. The 5' 700 bp PstI fragment of the c-*ros* cDNA was used to screen a chicken genomic DNA library (Vennstron et al., 1980) and three positive clones were isolated. The 15-kb clone 6a positive for the primer JMC4 (Fig. 6-1) was amplified and Southern blot analysis was performed using [^{32}P]-labeled JMC4 as the probe. A JMC4 positive 2.5-kb HindIII fragment was subcloned, restriction mapped and sequenced (Fig. 6-1 and Fig. 6-2). The JMC4 primer happens to be in the middle of the HindIII fragment and the most 5' 35 nt in clone 13 just lies immediate 5' to the JMC4 sequence (as illustrated in Fig. 6-1 & 6-2). Comparison of the genomic sequence and the cDNA

sequence shows that the first intron interrupts nucleotide 273 and 274 in cDNA (Fig. 3-2) and extends to the very end of this 2.5-kb genomic HindIII fragment.

Mapping the transcription initiation site of the c-*ros* mRNA

The primer extension and RNase protection analysis were then performed to map the mRNA initiation site. A 30-nucleotide antisense primer JMC54 which is 147 nucleotide downstream from the 5' end of the RACE clone 13 and starts in the middle of the MscI site (Fig. 6-1 and Fig. 6-2) of the 2.5-kb genomic DNA, was synthesized, labeled with ^{32}P and used to perform the primer extension experiment. The primer extension product by JMC54 shown in Fig. 6-3A is about 150 nucleotide in length and it matched the +1 G residue (Fig, 2) when running together with a sequencing reaction of the 2.5-kb genomic fragment with the same primer JMC54 (data not shown), thus mapping the initiation site to the 5' end of RACE clone 13. The RNase protection assay by probe A was then used to confirm this conclusion (Fig. 6-1). The protected fragment by this riboprobe would have a 5' end right in the middle of the MscI site which matches the 5' end of the primer JMC54 (Fig. 6-1). The protected band shown in Fig. 6-3B was about 150 nucleotides in length. In Fig. 6-3C, JMC54 primer extension product from kidney RNA was co-electrophoresed on a denaturing gel together with the MscI-ClaI riboprobe protected products from kidney, bursa and intestine RNA samples. The size of the fragments from both assays was almost identical. The small difference between the two products was probably due to the fact that in general the RNA transcripts moved 5% slower than their corresponding DNA counterpart (Sharmeen & Taylor, 1987). The protected RNA species appeared as a smear containing multiple bands. Both bands match pretty closely. Appearance of a lower discrete band may suggest that there is an alternative initiation site, which is likely to be the A residue four nucleotides downstream from the +1 G. This is consistent with the 5'

ends of the 3 RACE cDNA clones. However, this could also be due to heterogeneity of the riboprobe or over-digestion of the RNA duplex due to breathing of the ends. It is thus concluded that in kidney and intestine the mRNAs initiate around that G residue in Fig. 6-2 and the proximal promoter for c-*ros* is located upstream of this G residue.

Analysis of the putative c-*ros* promoter

The 1.3-kb region of the genomic DNA upstream of the mRNA initiation site was sequenced to search for the promoter sequence motifs. In this region, some promoter-specific motifs are present. Taking the G residue as +1, a "TATA" like sequence and a "CAAT" motif are present at -26 and -54 respectively (Fig. 6-2), although the "TAATAA" is not a typical one and the "CAAT" motif is in its opposite orientation. There are four AP-1 sites, one AP-2 site and two Oct-1 sites scattered within the upstream 0.9-kb region (Mitchell & Tjian, 1989). Notably, two perfect inverted repeats are present in the middle of this region (marked as repeats 1 and 2 in Fig. 6-2) and some imperfect direct repeats (marked as repeat 3 in Fig. 6-3) are present further downstream. It is possible that these repeats may serve as regulatory sequences or binding sites for transcription-regulatory factors.

Functional Study on the putative c-*ros* promoter

Functional assays are needed to substantiate that this region is indeed the c-*ros* promoter. The 1.3-kb promoter region was inserted before a CAT gene in pCat-enhancer (Promega), and then a SV40 polyadenylation signal (SVpA) was added before the c-*ros* promoter (ROSPRO, Fig. 6-4A) to suppress the non-specific transcription initiated from the upstream sequence in the vector. The resulted plasmid was named pCatpA, from which a serials of deletion mutants were generated (Fig. 6-4B). These plasmids and the pSV2Cat as the positive control (Gorman et al., 1982) were individually transfected into

COS-7 cells together with the pLacZ plasmid serving as an internal marker for normalizing the transfection efficiency. The cell extracts were subjected to CAT assay and galactosidase assay as described in Materials and Methods. Each value of CAT assay was normalized to the co-expressed lacZ activity. The result of one typical Cat assay was shown in Fig. 6-4B. The pCatpAXba containing no promoter sequence but the SVpA has a very low cat activity (2%), indicating that the added SVpA sequence is functioning in suppressing the non-specific transcription initiated from the upstream region. The parental plasmid pCatpA with the undeleted 1-kb putative *c-ros* promoter has a CAT conversion activity of 19%, which is about 10 times higher than that of the basal activity from the plasmid pCatpAXba containing no promoter sequences. Deletion of the sequence between BamHI and ClaI sites seems to increase the CAT activity to some extent (19% to 24%). Deletion of the sequence between ClaI and PacI sites produced little effect, further deletion of 19 nucleotides downstream of the PacI site reduced the CAT activity drastically (pCatpA20) and deletion of another 73 bp completely abolished the CAT activity (pCatpA111). Further deletion from pCatpA111 did not rescue the CAT activity. The above results indicate that the 1-kb genomic sequence upstream of the *c-ros* mRNA initiation site has the transcription-promoting activity and the basic promoter region can be narrowed down to a 300 bp region upstream of the transcription initiation site. It seems that the sequence between BamHI and ClaI sites has some negative regulatory effect on the promoter activity and the 100 bp downstream of the PacI site is essential for the promoter activity. This region may contain positive regulatory sequences or binding sites for the transcription factors. More precise deletion or point mutation can be generated to map those positive and negative regulatory sequences. As an example, the two inverted repeats (1 and 2 in Fig. 6-2) were deleted from the pCatpA, the resulting plasmid

pCatpAdR confers a CAT activity similar to pCatpA, suggesting that those two repeats are not important for the promoter activity in this assay system.

Discussion

Our and others' studies showed that the expression of *c-ros* is under tight regulation, both temporally and spatially (Sonnenberg et al., 1991; Tessarollo et al., 1992; Chen et al., 1993). Its expression is restricted to the epithelial cells of those positive organs. As a first step in exploring the molecular basis for the regulation of the chicken *c-ros* expression, we have mapped the *c-ros* transcription initiation site and isolated and characterized its promoter. This putative promoter was shown to be functional and by deletion study, some positive and negative regulatory regions were mapped.

The RACE method turned out to be useful in mapping the transcription initiation sites since the 5' ends of all the four positive clones matched the product of RPA, an independent method of characterizing the transcript. Both the G at position +1 and the A at +4 (Fig. 2) are all potential transcription initiation sites (Breathnach & Chambon, 1981), especially the sequence around the +4 A matches with the cap consensus sequence which is CA followed by a pyrimidine tract (Bucher & Trifonov, 1986). Since the primer extension product runs closer with the +1 G, we consider this as the major transcription start site. Comparison of genomic sequence with the cDNA sequence reveals that the first exon starts at nucleotide +1 and is interrupted at nucleotide +273 by the first intron (Fig. 3-3).

The putative *c-ros* promoter contains some consensus motifs like TATA and CAAT boxes, but no GC-rich sequences like the binding sites for sp1 are present. The positions of TATA box (-26) and CAAT (-54) are typical of a promoter. The nature of the TATA box is not clear because its sequence TAATAA does not match well with the consensus sequence (Breathnach & Chambon, 1981; Bucher & Trifonov, 1986). If this TAATAA does not work as a TFIID binding sites, the *c-ros* promoter can be considered as the type of the

tissue-specific promoter which does not contain TATA box or sp1 binding sites (Smale & Baltimore, 1989). Some potential transcription factor binding sites are present in this putative region, i.e. AP1, AP2 and Oct1 sites, but none of them has been shown to confer tissue-specific regulation. No sequences homologous to the regulatory sequences identified in the promoters of E-cadherin and epidermal keratin genes which are epithelial-specific (Snape et al., 1990, Leask et al., 1990; Behrens et al., 1991) are present in the c-*ros* promoter.

The putative c-*ros* promoter confers the transcription-promoting activity when placed before a CAT reporter gene. The deletion mutant study defined a potential negative regulatory region between BamHI and ClaI sites, deletion of which increased the CAT activity to some extent. Further deletions downstream of PacI site gradually decreased the CAT activity, indicating that the 90 to 100 bp after the PacI site must contain a positive regulatory sequence. There is one AP1 site and one 14-bp direct repeat within this region. The AP1 site must confer some positive activity because deletion of this site (pCatpA20) resulting in decrease of the CAT activity from 25.8% to 14.4%. The effect of the direct repeat with the region needs to be further determined. Deletion of the two inverted repeats around -590 has no effect on the CAT activity, suggesting that the repeats have no effect on the promoter activity in our assay system. Different AP-1, AP-2 or Oct-1 sites can be mutated or deleted to examine their effect on the promoter activity by CAT assay.

It is widely accepted that DNA methylation is one of the key elements in the control mechanism that governs vertebrate gene function and differentiation (Razin & Riggs, 1980; Bird et al., 1986). A high level of methylation in genomic DNA especially the 5' end of a gene usually indicates that the transcription is not active for this gene. The majority (90%) of the m^5Cytosine (m^5Cyt) residues in eucaryotic DNA are found in the dinucleotide

sequence CpG. It is not clear whether methylation is involved in the regulation of the c-*ros* expression because there are only 6 CpG dinucleotides within the 1290 bp upstream of the transcription start site. Unfortunately, there are no MspI, HapII or other restriction sites that are sensitive to the methylation and are used routinely for dissecting the state of DNA methylation (Bird, 1987).

Since tissue-specifically expressed genes appear to be bound to a complex of tissue-specific factors (Bird et al., 1986), this type of regulation is more likely in the case of c-*ros*. Some tissue-specific factors may bind to the c-*ros* promoter region to regulate its expression. Protein binding and gel shift assay can be performed to identify those transcription factors. Some kidney or intestine epithelial cell lines could be used for c-*ros* promoter driven CAT expression to further analyze the promoter. Furthermore, c-*ros* promoter/lacZ fusion cassette can be used to assay the tissue-specific activity of the promoter in transgenic mice.

Fig. 6-1 Organization of the putative c-*ros* promoter region. Top line represents some restriction sites in the 2.5-kb HindIII fragment from the genomic clone 6a. The solid box indicates the putative *ros* promoter. The arrow 5' to the MscI site indicates the possible transcription initiation site. The arrow 3' to MscI site indicates the position of the beginning of the first intron. JMC4 is a 26-mer primer in the opposite direction of the coding sequence. JMC54 is a 30-mer primer with its 5' end matching the middle of the MscI site. The position of probe A described in materials and Methods is also shown. cDNA clone 5b and RACE clone 13 are also shown in parallel with the genomic clone.

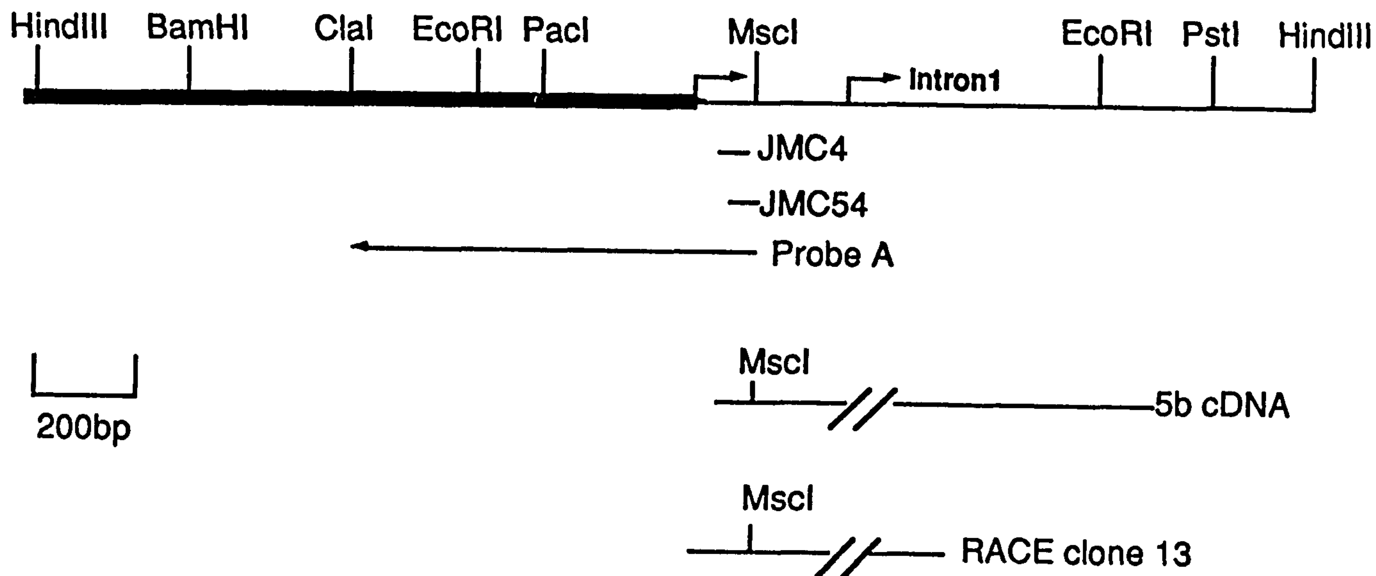


Fig. 6-2 Nucleotide sequence of the putative c-*ros* promoter region. The sequence was derived from the 2.5-kb HindIII fragment genomic clone 6a. Restriction enzyme sites, "TATA" box, "CAAT" box and potential transcription factor binding sites, Ap-1, Ap-2 and Oct1 are underlined. The two inverted repeats and one direct repeat are highlighted by arrowed overlines. The 35 nucleotides which are underlined and in italic are the cDNA sequence derived from the 5' RACE clone 13 and the arrow indicates the possible start site for the c-*ros* transcript. The sequence before translation initiation codon ATG is also shown.

```
                                   AAGCTTCTTCAACCAGGCTGAGCTGGATTA -1261
                                   HindIII
GCAAAGAGTTACTCTTTAAGCAATCAGACTGACCTCGTTTATTTTGGCTGGAGCTAAGGA -1201

GTTTCCTCTTGAGCAGCTTTGCTGGTGGGTTAGAAGGGGGGCAGCAGCTGTCTGGAGTGG -1141

AGAGGTGGCAAAATGTTGAAGACTGACTTCCCTCCCTCCACTGGATAGCTGGAAGCGTAC -1081

AGTCTTACTGTCTGAAGGAGTTCCCTTGGGCAGCAGTGAGAAAGCCCCTCAGGGCTTTCC -1021

AGAGCAAATGGATGAAAAGGAAGGATCCTGTTCCACTAAGCCCTTTCTTCCCCAGGCCAG  -961
                      BamHI                       AP-2
GTGCTTGGCAGTATCCCTCAGATCCCTCCAGCCTGGATTATTCCTTACTCCTGCCTCAGT  -901
                                                    AP-1
AGAGTTGGCTCCTTTCACCCACCACACTTGAGCACAGCAGCACACGGAACTGTGCTGTGA  -841

TCTGACTCAGAGACAACAACTTTGGGTTAATGCTAGCCAGTTTGTTTCTCTTGTTGCTGG  -781
  AP-1
GTTAGCGTTAACCAGATTCTCAGCCTAGTTCATTGCATGAATGCATGGACATTTGCTTGG  -721
                                                    Oct-1
CAAATGAGTTGGGTGTGGGAAGGTGGAGGAGCATTTGAGTATGAGGGGATAATCACATCG  -661
                                    1                  ClaI
ATGGGATGTGACAGTTTGAACTGGTCTTTAAAATGCTATGTTACTTTTGTGCTTGTGTAT  -601
  2                        2              1
ACATGCATACAAAACCACACAAAATGCATGTACAGAAATAACATAGCATCTATTTATACA  -541

TACAAATATGTGAGCTGGATTGATTTCTTATTCCAGTCCACAGAAACTTTGTGGATAGCT  -481
 Oct-1
CTGCTAGTACAATAGAGATATGCACAGGCTGGAAGATTCAAGTATCTCAAGCCTGAATTC  -421
                                                       EcoRI
CCACTGACTTCTTTAATACTGCTGTCCAACCTGATGAAGGACACTTTACCCCTCTCTAAA  -361

TCAGATACATGGTTTATCTATCCACTTTTTTTCCATCATACAGGCATAAATATATTGAAT  -301
                                                      3
AGAATTAATTAATATTTTTATACTGACTGATACGACCTGTAGTGTTCTAACAAAGTCACT  -241
     PacI        3          AP-1
CATAATCATGAACAAATGCACTCAGTTGTAGACATTTGAAAATGCATGGCAGTAGGTAGA  -181

AGTTAAGATTAGGGCACTTTTCTCCATCTTTCCATTTCCGTGTTGTTTGCAGCTCCTGAC  -121
                                                       AP-1
TGAAGCATTGTGCTTAGTTTCCTCTTGACGTTAATGATATTTATGTAAGGAGATACGCTG   -61
      "CAAT"                         "TATA"
ACTTTCATTGGTCACCCCTTTTGTTTCAGTGGTTTAATAAGAAACTGTGAAATAGGTGTT    -1
|--c-ros exon 1
GTGCAATTTCAAAGCAACTGCTCCAGGGAGCAGAAAGGTTGTTCAAGGAATGGATTGTAA   +60
                                   |+Xba
ACAAAGAATTAAAAATATTTAATAAATACTAAAGTAGAAATTTCGTAGTCGTGACTCTGT  +120

CACCTAGACTTGAAGGAGTTTGTATGGCCAAGCCAAGGAAGGAATCAGACTGAAAGATGA  +180
                   Met       MscI
AATTTCTCAAGTTATG                                              +196
```

Fig. 6-3 Primer extension and RNase Protection assay for mapping the initiation site of the c-*ros* transcript. (A) primer extension. 50 μg total tissue RNAs (150 μg for the intestine in 2nd lane) were used for primer extension using c-*ros*-specific primer JMC54 which was labeled with $[\gamma\text{-}^{32}P]$ATP to high specificity. The primer extension products were resolved on a denaturing gel. The abbreviations for tissues are: i, intestine; k, kidney; c, yeast RNA as negative control; M, molecular weight marker (ΦX174 DNA digested with HaeIII, dephosphorylated and labeled with $[\gamma\text{-}^{32}P]$ATP, using T4 polynucleotide Kinase). Sizes are indicated in number of nucleotides. (B) RNase protection assay (RPA). For each assay 30 μg of total RNA from different tissues was hybridized to $[\alpha\text{-}^{32}P]$UTP-labeled antisense transcripts (see text). The hybrid molecules were subjected to RNA digestion, and protected fragments were resolved on a denaturing gel. (C) RNase protected products as in B run together with primer extension products as in A. b stands for bursa and PE stands for primer extension.

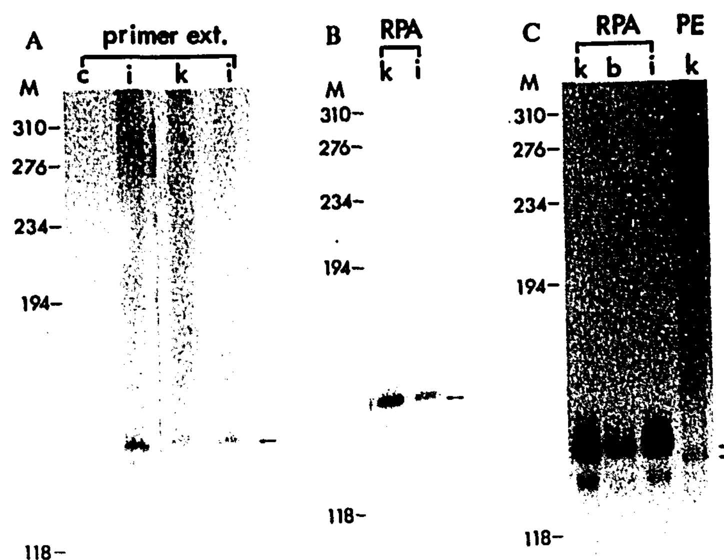
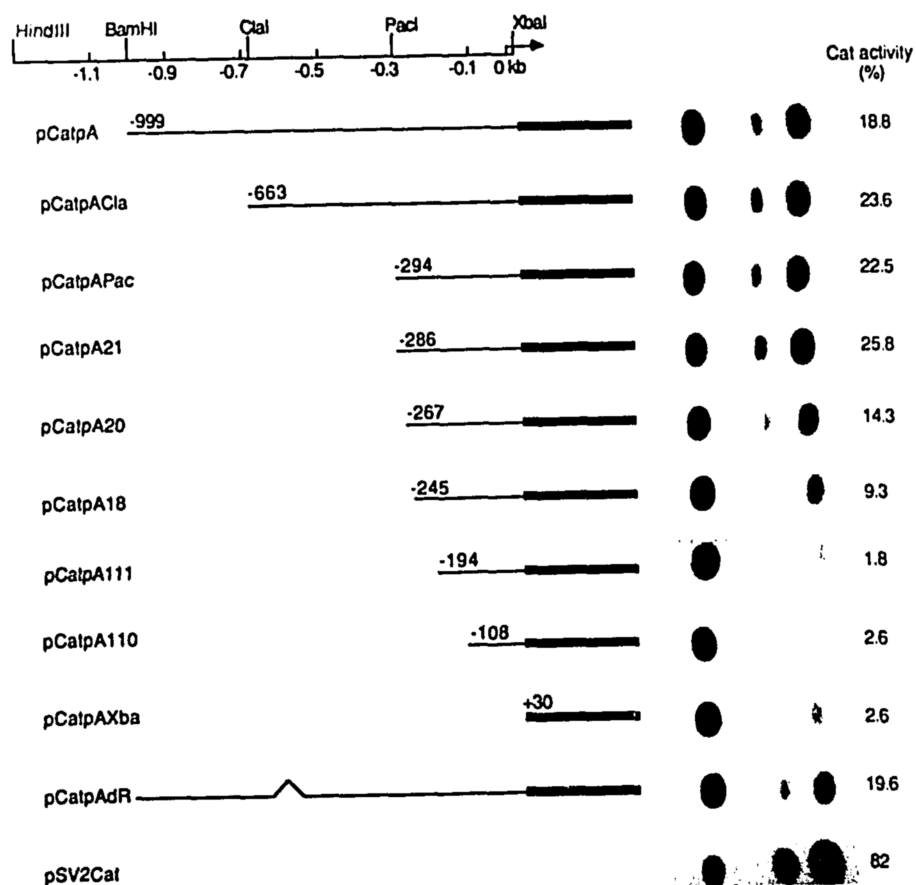


Fig. 6-4 The putative c-*ros* promoter is functionally active. (A) Plasmid map of pCatpA. (B) Deletion mutants of pSVpA and their CAT activity. 20 μg each of the plasmids was transfected together with 2 μg of pLacZ into COS-7 cells. 48 hrs later, cell lysate was subjected to CAT assay and β-galactosidase assay as described in Materials and Methods. Conversion represented by percentages was derived from the scintillation counting of those spots. The result represents a typical one of the experiemnt repeated threee times.



Chapter VII. General Discussion

The retrovirus UR2 was isolated back in 1982 from a avian sarcoma virus together with its helper virus UR2AV (Balduzzi et al., 1981). Later, the oncogene transduced in this virus was identified to be a new PTK gene named *ros* (Wang et al., 1982) encoding a gag fusion protein called the $P68^{gag\text{-}ros}$ (Neckameyer et al., 1984). cDNA and genomic clones for chicken c-*ros* were isolated and the sequence analysis suggested that c-*ros* codes for a receptor-like protein tyrosine kinase (Neckameyer et al., 1985; Podell & Sefton, 1987). Significant progress in the understanding of the possible function of c-*ros* has been made in the past years. cDNAs for the human, rat and partially for the mouse c-*ros* have been cloned (Matsushime & Shibuya, 1990; Birchmeier et al., 1990; Sonnenberg et al., 1991; Tessarollo et al., 1992), and the tissue-specific expression pattern of c-*ros* in mouse was elucidated (Sonnenberg et al., 1991; Tessarollo et al., 1992). For my Ph.D. thesis study, efforts were made to understand the molecular structure, biochemical and biological properties, the tissue-specific expression and the molecular basis for the regulated expression of the ckicken c-*ros*. I will give a general discussion on each part of my work, followed by a perspective of future study on the c-*ros*.

The full length cDNA for the chicken c-*ros* was molecularly cloned and sequenced and the sequence clearly shows that it codes for a RPTK (Chen et al., 1991). It shares great homology both in sequence and structure with the product of the *Drosophila sevenless* gene (Chen et al., 1991). Despite the difference in the structure and processing of the final protein product, c-*ros* and *sevenless* share the following features both in the EC domain and kinase domain: 1) high homology in the kinase domain (75%); 2) the characteristic 6 aa insertion and the cluster of three major tyrosine phosphorylation sites instead of one as in other PTKs; 3) a large EC domain with a 43% similarity and the

8 conserved type III fibronectin repeats. It is very likely the c-*ros* is the vetebrate homolog of the *Drosophila sevenless* gene. However, the function of c-*ros* and *sevenless* obviously diverged. While *sevenless* is important for the differentiation of the photoreceptor R7 in the *Drosophila* ommatidia, c-*ros* is mainly expressed in the epithelial cells of kidney, intestine, bursa and possibly lung and thymus. Although the ligand for the *sevenless* was identified and downstream signaling pathway for the *sevenless* is much clearer than before, almost nothing is known about the signal transduction of c-*ros*. It is possible that *sevenless* and c-*ros* share certain common components in their signal transduction pathways and function similarly in terms of their roles in the induction and differentiation.

Little was known about the structure and biochemical properties of the c-*ros* proten product. With full length c-*ros* cDNA, I constructed different plasmids for the expression of the c-*ros* cDNA in mammalian cells as well as in CEF. Both transient and permanent expression of c-*ros* shows that c-*ros* is mainly synthesized as a polypeptide of 260-280-kDa with low kinase activity in the absence of ligand stimulation. The two different species of the c-*ros* protein probably represent differential glycosylation because tunicamycin treatment prevented the synthesis of both species. Instead, a smaller product was observed. No similar cleavage in the EC domain as seen for Sevenless, IR and IGFR was observed for the c-*ros* product. Proteins of a similar size were observed in the tissue extracts of the chichen kidney and intestine, indicating that the c-*ros* protein expressed in cultured cells is similar to that *in vivo* although the glycosylation or other modification might be different. Furthermore, there is no evidence for covalent dimerization of c-Ros but whether non-covalent dimerization exists remains to be determined. Since the protein size of the c-Ros is very large and its ligand was unknown, a mutant protein ppRos was generated by deleting most of the EC domain to facilitate the study of the biochemistry

and signal transduction of c-Ros. ppRos has an elevated kinase activity and a much longer half life in comparison to c-Ros. The failure of c-Ros to transform the CEF can be explained by its inactive PTK and low level expression in CEF, but surprisingly, the fully activated kinase ppRos can not transform CEF either. Further study shows that although ppRos is tyrosine phosphorylated *in vivo* and can phosphorylate an exogenous substrate such as lysozyme *in vitro*, the pattern of substrate phosphorylation in CEF over-expressing ppRos is different from those of UR2 and VCros transformed CEF. The signal transduction of v-Ros and VCRos versus ppRos was studied by examing the phosphorylation of specific substrates and activation of PI-3-Kinase. In UR2 and VCros transformed CEF, PI-3-Kinase is associated with the *ros* proteins and the phosphorylation of Shc, PLCγ1 and an unknown 88-kDa protein is elevated. Since PLCγ1 and PI-3-Kinase can mediate the mitogenicity of PDGF receptor and over-expression of Shc can cause cell transformation through the activation of the *ras* signaling pathway, it is very likely the PI-3-kinase, PLCγ1 and Shc are also important mediators of the transforming and mitogenic signals of v-Ros and VCRos. The inability of ppRos to transform CEF is probably due to the fact that the ppRos can neither associate with the PI-3-kinase activity nor phosphorylate Shc and PLCγ1. The phosphorylation of a 100-kDa protein by ppRos but not by v-Ros or VCRos suggests an alternative possibility that p100 may be a substrate of c-Ros that mediate an anti-mitogenic signal. Whether the autophosphorylation sites of the ppRos are different from those of the VCRos and v-Ros remains to be defined. Since the major difference between VCRos and ppRos is the EC domain and the three aa insertion in the TM domain of VCRos, it will be interesting to see whether the EC domain or the TM domain can exert some influence on both the kinase activity and on the interaction between the RPTK and its substrates.

Increasingly, RPTKs have been shown to be important in the development and differentiation. The finding that c-Ros is specifically expressed in the epithelial cells of the renal collecting tubules, intestinal villi and crypts of, and capsule of the bursa suggests that c-*ros* may play roles in the differentiation or maintenance of those specific epithelial cells. The restricted expression of c-Ros at the tip of the collecting duct in the embryo correlates with the epithelial/mesenchymal inductive events, suggesting that c-Ros might be involved in the organogenesis. The major difference of c-Ros expression between chicken and mouse is that in chicken c-Ros was expressed at high levels in kidney and intestine even at adulthood. This might reflect species differences or in adult mouse the function of another gene can take place c-Ros. The finding of the expression of the c-Ros in the chicken bursa and thymus raises the possibility of the c-Ros expression in mouse thymus. Further proof of the roles of the c-*ros* in the organogenesis awaits the genetic evidence from transgeneic mice with dominant negative *ros* or *ros*-null mice.

Since c-*ros* is specifically expressed in epithelial cells and the expression is under tight temporal and spatial regulation, the putative c-*ros* promoter in chicken was isolated to elucidate the molecular basis for the regulation. The transcription start site was mapped using RPA and primer extension. This putative promoter region contains some potential transcription factor binding sites and some inverted and direct repeats. This putative promoter confers transcription promoting activity when placed before a Cat reporter gene. Serial deletion mutants were made and a potential positive regulatory region was defined. However, the region confering the tissue-specific expression is not identified due to the lack of proper *in vitro* system in which the c-*ros* is specifically expressed. Alternatively, a transgenic model with the c-*ros* promoter/lacZ reporter cassette should be imformative.

With c-*ros* molecularly cloned and its tissue-specific expression pattern established, attention should now shift to signal transduction, regulation of the expression, identification of the c-*ros* ligand and eventually understanding the biological function of the c-*ros*. Since a cell line expressing c-*ros* at high levels is not available and the ligand of the c-*ros* is not yet identified, one can use the activated c-*ros* mutant ppRos which might represent the activated c-*ros* quantitatively and qualitatively. Otherwise the activation of the c-*ros* kinase awaits the identification of the c-*ros* ligand. Alternatively, a chimeric c-*ros* RPTK with the EC domain from a receptor of known ligand can be constructed for activation of c-*ros*. To study the signal transduction of v-*ros*, a series of mutants could be made with each of the tyrosine residue in the cytoplasmic tail mutated to phenylalanine. The goal is to see which tyrosine residue confers the ability to associate with Shc, PLCγ1 or PI-3-Kinase and thus activate the downstream signalling. The positive and negative regulatory elements in the putative promoter should be more precisely mapped by either deletion or site directed mutagenesis. The tissue-specific regulatory region can be mapped by performing the transfection and CAT assay in some kidney or intestine epithelial cells although COS-7 cell is a epithelial-like cell line. Most importantly, a genetic model is needed to eventually understand the function of c-*ros* protein. Transgenic mice with the activated ppRos will be imformative in understanding the function of c-*ros*, so will the c-*ros* knock-out mice.

Bibliography

Ackerman, G. A. and Knouff, R. A. (1959). Lymphocytopoiesis in the bursa of Fabricius. Ameri. J. Anat. 104, 163-205.

Adari, H., Lowy, D. R., Willumsen, B. M., Der, C. J. and McCormick, F. (1988). Guanosine triphosphatase activating protein (GAP) interact with the p21 *ras* effector binding domain. Science 240, 518-521.

Alessandrini, A., Crews, C. M., and Erikson R. L. (1992). Phorbol ester stimulates a protein-tyrosine/threonine kinase that phosphorylates and activates the ERK-1 gene product. Proc. Natl. Acad. Sci. USA 89:8200-8204.

Ambrosio, L., Mahowald, A. P. and Perrimon, N. (1989). Requirement of the *Drosophila raf* homolog for *torso* function. Nature 342, 288-290.

Anderson, D., Koch, C. A., Grey, L., Ellis, C., Moran, M. F. and Pawson, T. (1990). Binding of SH2 domains of phospholipase $C_{\gamma}1$, GAP, and *src* to activated growth factor receptors. Science 250, 979-982.

Aroian, R. V., Koga, M., Mendel, J. E., Oshima, Y. and Sternberg, P. W. (1990). The let-23 gene necessary for *C. elegans* vulval induction encodes a tyrosine kinase of the EGF receptor subfamily. Nature 348, 6939-699.

Aruffo, A., and Seed, B. (1987). Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. Proc. Natl. Acad. Sci. USA 84, 8573-8577.

Bagchi, S., Weinmann, R., and Raychaudhuri, P. (1991). The retinoblastoma protein copurifies with E2F-I, an E1A-regulated inhibitor of the transcription factor E2F. Cell 65, 1063-1072.

Balduzzi, P. C., Notter, M. F. D., Morgan, H. R. and Shibuya, M. (1981). Some biological properties of two new avian sarcoma viruses. J. Virol. 40, 268-275.

Bargonetti, J., Friedman, P.N., Kern, S.E., Vogelstein, B., and Prives, C. (1991). Wild-type but not mutant p53 immunopurified proteins bind to sequences adjacent to the SV40 origin of replication. Cell 65, 1083-1091.

Basler, K., Christen, B. and Hafen, E. (1991). Ligand-independent activation of the *sevenless* receptor tyrosine kinase changes the fate of cells in the developing *Drosophila* eye. Cell 64, 1069-1081.

Basler, K. and Hafen, E. (1988). Control of photoreceptor cell fate by the *sevenless* protein requires a functional tyrosine kinase domain. Cell 54, 299-311.

Behrens, J., Lorick, O., Klein-Hitpass, L., and Birchmeier, W. (1991). The E-cadherin promoter: Functional analysis of a G.C-rich region and an epithelial cell-specific palindromic regulatory element. Proc. Nat. Acad. Sci. USA 88, 11495-11499.

Beitel, G.J., Clark, S.G. and Horvitz, H.R. (1990). C. elegans *ras* gene *let*-60 acts as a switch in the pathway of vulval induction. Nature 348, 503-509.

Berridge, M. J. (1984). Inositol triphosphate and diacylglycerol as second messengers. Biochem.J. 220, 345-360.

Berridge, M. J. and Irvine, R.F. (1989). Inositol phosphate and cell signalling. Nature 341, 197-205.

Besmer, P., Murphy, J. E., George, P. C., Qiu, F., Bergold, P. J., Lederman, L., Snyder, H. W., Jr., Brodeur, D., Zuckerman, E. E. and Hardy, W. D. (1986). A new acute transforming feline retrovirus and relationship of its oncogene v-*kit* with the protein kinase gene family. Nature 320, 415-421.

Birchmeier, C., Birnbaum, D., Waitches, G., Fasano, O. and Wigler, M. (1986). Charaterization of an activated human c-*ros* gene. Mol. Cell. Biol., 6:1122-1129.

Birchmeier, C., O'Neill, K., Riggs, M. and Wigler, M. (1990). Characterization of ROS1 cDNA from a human glioblastoma cell line. Proc. Natl. Acad. Sci. USA 87, 4799-4803.

Birchmeier, C., Sharma, S., and Wigler, M. (1987). Expression and rearrangement of the *ros*1 gene in human glioblastoma cells. Proc. Natl. Acad. Sci. USA 84, 9270-9274.

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. Nature 321, 209-213.

Bird, A. P. (1987). CpG islands as gene marker in the vertibrate nucleus. Trends in Genetics 3, 342-347.

Bishop, J.M. (1991). Molecular themes in oncogenesis. Cell 64, 235-248.Blenis, J. (1991). Growth-rigulated signal transduction by the MAP kinases and RSKs. Cancer Cells 3:445-449.

Blenis, J. 1991. Growth-regulated signal transduction by the MAP kinases and RSKs. Cancer Cells 3, 445-449.

Bonfini, L., Karlocick, C. A., Dasgupta, C. and Banerjee U. (1992). The *son of sevenless* gene product: a putative activator of *ras*. Science 255, 603-606.

Boulton, T. G., Nye, S. H., Robbins, D. J., Ip, N. Y., Radziejiewska, E., Morgenbesser, S. D., Depinho, D. A., Panayotatos, N., Cobb, M. H. and Yancopoulos, G. D. (1991). ERKs: a family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF. Cell 65, 663-675.

Bowtell, D. D. L., Simon, M. A. and Rubin, G. M. (1988). Nucleotide sequence and structure of the *sevenless* gene of *Drosophila melanogaster*. Genes Dev. 2, 620-634.

Breathnach, R., and Chambon, P. (1981). Organization and expression of eucaryotic split

genes coding for proteins. Ann. Rev. Biochem. 50, 349-383.

Broach, J.R. (1991). RAS genes in S. cerevisiae: signal transduction in search of a pathway. Trends in Genetics 7, 28-32.

Bucher, P., and Trifonov, E.N. (1986). Compilation and analysis of eucaryotic POL II promoter sequences. Nuc. Acid Res. 14, 10009-10026.

Cantley, L. C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. and Soltoff, S. (1991). Oncogenes and signal transduction. Cell 64, 281-302.

Chellappan, S.P., Hiebert, S., Mudryl, M., Horowitz, J.M., and Nevins, J.R. (1991). The E2F transcription factor is a cellular target for the Rb protein. Cell 65, 1053-1061.

Chen, R. H., C. Sarnecki, and J. Blenis. (1992). Nuclear localization and regulation of erk- and rsk- encoded protein kinases. Mol. Cell. Biol. 12:915-927.

Chen, J., Heller, M., Poon, B., Kang, L. and Wang, L.-H. (1991). The proto-oncogene c-*ros* codes for a transmembrane tyrosine protein kinase sharing sequence and structural homology with *sevenless* protein of *Drosophila melanogaster*. Oncogene 6, 257-264.

Chen, J., Zhu, X., and Silerstein, S. (1991). Mutational analysis of the sequence encoding ICP0 from herpessimplex virus type 1. Virol 180, 207-220.

Clark, S.G., Stern, M.J. and horvitz, H.R. (1992). C. elegans cell-signalling gene sem-5 encodes a protein with SH2 and SH3 domains. Nature 356, 340-344.

Clarke, A.R., Robanus Maandag, E., van Roon, M., van der Valk, M., Hooper, M.L., Berns, A., and te Riele, H. (1992). Requirement for a functional Rb-1 gene in murine development. Nature 359, 328-330.

Cooper, G. M. (1990). Oncogenes, Jones and Bartlett, Publishers, Inc.

Crews, C. M., and Erikson, R. L. (1992). Purification of a murine protein-tyrosine/threonine kinase that phosphorylates and activates the ERK-1 gene product: relationship to the fission yeast byr1 gene product. Proc. Natl. Acad. Sci. USA 89:8205-8209.

Crews, C. M., Alessandrini, A. and Erikson, R. L. (1992). The primary structure of MEK, a protein kinase that phosphorylates the ERK gene product. Science 258:478-480.

Cross, M. and Dexter, T. M. (1991). Growth factors in development, transformation, and tumorigenesis. Cell 64, 271-280.

Denohower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Butel, J.S., Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumors. Nature 356, 215-221.

Dickson, B., Sprenger, F., Morrison, D. Hafen, E. (1992). raf functions downstream of

*ras*1 in the sevenless signal transduction pathway. Nature 360, 600-603.

Dorai, T., Levy, J. B., Kang, L., Brugge, J. S. and Wang L.-H. (1991). Analysis of cDNAs of the proto-oncogene c-*src*: heterogeneity in 5' exons and possible mechanism for the genesis of the 3' end of v-*src*. Mol. Cell Biol. 11, 4165-6176.

Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. and Waterfield, M. D. (1984). Close similarity of epidermal growth factor and v-*erb*B oncogene protein sequences. Nature 307, 521-527.

Doolittle, R. F., Hunkapiller, M. W., Hood, L. E., Devare, S. G., Robbins, K. C., Aaronson, S. A. and Antoniades, H. N. (1983). Simian sarcoma virus oncogene, v-*sis*, is derived from the gene encoding a platelet growth factor. Science 222, 275-277.

Dorai, T. and Wang, L.-H. (1990) An alternative non-tyrosine protein product of the c-*src* gene in chicken skeletal muscle. Mol. Cell. Biol. 10, 4068-4079.

Dressler, G.R., Deutsch, U., Chowdhury, K., Nornes, H.O., and Gruss, P. (1990). Pax2, a new murine paired-box-containing gene and its expression in the developing excretory system. Development 109, 787-795.

Ebina, Y., Ellis, L., Jarnajin, K., Edery, M., G*raf*, L., Clauser, E., Qu, J.-H., Masiarz, F. Kan, Y. W., Goldfine, I. D., Roth, R. A. and Rutter, W. J. (1985). The human insulin receptor cDNA: the structural basis for hormone-activated transmembrane signalling. Cell 40, 747-758.

Egan, S. E., Giddings, B. W., Brooks, M. W., Duday, L., Sizeland, A. M. and Weinberg, R. A. (1993). Association of Sos *ras* exchange protein with Grb2 is implicated in tyrosine kinase signal transduction and transformation. Nature 363, 45-51.

Ellis, L., Clauser, E., Morgan, D. O., Edery, M., Roth, R. A., and Rutter, W. J. (1986). Replacement of insulin receptor tyrosine residues 1162 and 1163 compromises insulin-stimulated kinase activity and uptake of 2-deoxyglucose. Cell 45, 721-732.

Ellis, C., Moran, M., McCormick, F. and Pawson, T. (1990). Phosphorylation of GAP and GAP-associated proteins by transforming and mitogenic tyrosine kinases. Nature 343, 377-383.

Eliyahu, D., Goldfinger, N., Pinhasi-Kimhi, O., Shaulsky, G., Skurnik, Y., Arai, N., Rotter, V., and Oren, M. (1988). Meth A fibrosarcoma cells express two transforming mutant p53 species. Oncogene 3, 313-321.

Endemann, G., Yonezawa, K. and Roth, R. (1990). Phosphatidylinositol kinase or an associated protein is a substrate for the insulin receptor tyrosine kinase. J. Biol. Chem. 265, 396-400.

Escobedo, J. A., Navankasattusas, R., Kavanaugh, W. M., Milfay, D., Fried, V. A., and L. T. Williams. (1991). cDNA cloning of a novel 85 kd protein that has SH2 domains and

regulates binding of PI3-kinase to the PDGF ß-receptor. Cell 65,75-82.

Fantl, W.J., Escobedo,J.A, Martin, G. A., Turck, C. W., Rosario, M., McCormick, F., and Williams, L. T. (1992). Distinct phosphotyrosines on a growth factor receptor bind to specific molecules that mediate different signaling pathways. Cell 69, 413-423.

Farmer, G., Bargonetti, J., Zhu, H., Friedman, P., Prywes, R., and Prives, C. (1992). Wild-type p53 activates transcription in vitro. Nature 358, 83-86.

Fasano, O., Birnbaun, D., Edlund, L., Fogh, J. and Wigler, M. (1984). New human transforming genes detected by tumorigenicity assay. Mol. Cell. Biol. 4, 1695-1705.

Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell 61, 759-767.

Fearon, E.R., Cho, K.R., Nigro, J.M., Kern, S.E., Simons, J.W., Ruppert, J.M., Hamilton, S.R., Preisinger, A.C., Thomas, G., Kinzler, K.W., and Volgestein, B. (1990). Identification of a chromosome 18q gene that is altered in colerectal cancers. Science 247, 49-56.

Feldman, R.A., Wang, L.-H., Hanafusa, H. and Balduzzi, P. C. (1982). Avian sarcoma virus UR2 encodes a transforming protein which is associated with a unique protein kinase activity. J. Viol. 42, 228-236.

Feng, G.-S., Hui, C.-C., Pawson, T. (1993). SH2-containing phophotyrosine phosphatase as a target of tyrosine protein kinases. Science 259, 1607-1610.

Fields, S., and Jang, S.K. (1990). Presence of a potent transcription activating sequence in the p53 protein. Science 249, 1046-1049.

Flanagan, J. G. and Leder, P. (1990). The *kit* ligand; a cell surface molecule altered in steel mutant fibroblast. Cell 63, 183-194.

Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986). A human DNA segment with property of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 323, 643-646.

Frohman, M. A., Dush, M. K. and Martin, G. R. (1988). Rapid production of full length cDNA from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl Acad. Sci. USA 85, 8998-9002.

Fukui, Y., Kornbluth, S., Jong, S.-M., Wang, L.-H. and Hanafusa, H. (1989). Phosphatidylinositol kinase type I activity associates with virus oncogene products. Oncogene Research 4, 283-292.

Gale N. W., Kaplan, S., Lowenstein, E. J., Schlessinger, J. and Bar-Sagi, D. (1993). Grb2 mediates the EGF-dependent activation of guanine nucleotide exchange on *ras*. Nature 363, 88-92.

Garber, E. A., Hanafusa, T. and Hanafusa, H. (1985). Membrane association of the transforming protein of avian sarcoma virus UR2 and mutants temperature sensitive for cellular transformation and protein kinase activity. J. Viol. 56, 790-797.

Gilbert, S.F. (1988). Developmental Biology. (Sinauer associates, Sunderland, Massachussettes).

Glass, D. J., Nye, S., Tantzopoulos, P., Macchi, M. J., Squinto, S., Goldfarb, M. and Yancopoulos, G. D. (1991). *trk*B mediates BDNF/NT-3-dependent survival and proliferation in fibroblast lacking the low affinity NGF receptor. Cell 66, 405-413.

Glisin, V., Crkvenjakov, R., and Byus, C. (1974). Ribonucleic acid isolated by cesium chloride centrifugation. Biochemistry 13, 2633-2637.

Gluzman, Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. Cell 23, 175-182.

Gorman, C.M., Morffat, L.F., and Howard, B.H. (1982). Recombinant genomes which express chlorophenicol acetyltransferase in mammalian cells. Mol. Cell Biol. 2, 1044-1051.

Gomez, N. and Cohen, P. (1991). Dissection of the protein kinase cascade by which nerve growth factor activates MAP kinases. Nature 353:170-173.

Gulbins, E., Coggeshall, K.M., Baier, G., Katzav, S., Burn, P. and Altman, A. (1993) Tyrosine kinase-stimulated guanine nucleotide exchange activity of Vav in T cell activation. Science 260, 822-825.

Hall, A. (1992). Signal transduction through small GTPase-a tale of two GAPs. Cell 69, 389-391

Hamaguchi, M., Grandori, C., and Hanafusa, H. (1988). Phosphorylation of cellular proteins in Rous sarcoma virus-infected cells: analysis by use of anti-phosphotyrosine antibodies. Mol. Cell. Biol. 8, 3035-3042.

Hampe, A., Gobert, M., Shewrr, C. J. and Galibert, F. (1984). Nucleotide sequence of the feline retroviral oncogene v-*fms* shows unexpected homology with oncogenes encoding tyrosine-specific protein kinase. Proc. Natl. Acad. Sci. USA 81, 85-89.

Han, M. and Sternberg, P.W. (1990). *let*-60, a gene that specifies cell fates during c. elegans vulval induction, encodes a *ras* protein.development. Cell 63, 921-931.

Hanafusa, H. (1969). Rapid transformation of cells by Rous sarcoma virus. Proc. Natl. Acad. Sci. USA 63, 318-325.

Hanks, S. K., Quinn, A. M. and Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science 241, 42-52.

Harris, H. (1988). The analysis of malignancy by cell fusion: the position in 1988. Cancer

Res. 48, 3302-3306.

Hart, A.C., Krämer, H.K., and Zipursky, S.L. (193). Extracellular domain of the boss tansmembrane ligand acts as an antagonist of the sev receptor. Nature 361, 732-736.

Hempatead, B., Martin-Zanca, D., Kaplan, D. R., Parada, L. F. and Chao, M. V. (1991). High-affinity NGF binding requires coexpression of the *trk* proto-oncogene and the low-affinity NGF receptor. Nature 350, 678-683.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28, 351-359.

Hiles, I. D., Otsu, M., Volinia, S., Fry, M. J., Gout,I., Dhand, R., Panayotou, G., Ruiz-Larrea, F., Thompson, A., Totty, N. F., Hsuan, J. J., Courtneidge, S. A., Parker, P. J., and Waterfield, M.D. (1992). Phophatidylinositol-3-Kinase: structure and expression of the 110 kd catalytic subunit. Cell 70, 419-429.

Holmes, W.E., Sliwkowski, M.X., Akita, R.W.,Henzel, W.J.,Lee, J., Park, J.W., Kuang, W.-J., Wood, W.I., Goeddel, D.V., and Vandlen, R.L. (1992). Identification of Heregulin, a specific activator of p185^{erbB2}. Science 256, 1205-1210.

Huang, E., Nocka, K., Beier, D. R., Chu, T.-Y., Buck, J., Lahm, H.-W., Wellner, D., Leder, P. and Besmer, P. (1990). The hematopoietic growth factor KL is encoded at the Sl locus and is the ligand of the *c-kit* receptor, the gene product of the W locus. Cell 63, 225-233.

Hunter, T. (1991). Cooperation between oncogenes. Cell 64, 249-270.

Li, N., Batzer, R., Daly, R., Yajnik, V., Skolnik, E., Chardin, P., Bar-Sagi, D., Margolis, B. and Schlessiger, J. (1993). Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras signalling. Nature 363, 85-88.

Jacks, T., Fazeli, A., Schmitt, E.M., Bronson, R.T., Goodell, M.A., Weiberg, R.A. (1992). Effects of an Rb mutation in the mouse. Nature 359, 295-230.

Jong, S.-M. and Wang, L.-H. (1987). The transforming protein p68$^{gag\text{-}ros}$ of avian sarcoma virus UR2 is a transmembrane protein with the *gag* portion protruding extracellularly. Oncogene Res. 1, 7-21.

Jong, S.-M. and Wang, L.-H. (1990). Role of *gag* sequences in the biochemical properties and transforming activity of the avian sarcoma virus UR2-encoded $^{gag\text{-}ros}$ fusion protein. J. Viol. 64, 5997-6009.

Jong, S.-M., and Wang, L.-H. (1991). Two point mutations in the transmembrane domain of P68$^{gag\text{-}ros}$ inactivate its transforming activity and cause a delay in embrane association. J.Viol. 65, 180-189.

Jong, S.-M. and Wang, L-H. (1992). Transforming properties and substrate specificities of the protein tyrosine oncogenes ros and src and their recombinants. J. Virol. 66, 4909-

4918.

Joslyn, G., Carlson, M., Thliverris, A., Albertsen, H., Gelbert, L., Samowitz, W., Groden, J., Stevens, J., Spirio, L., Robertson, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, R., Hughes, J.P., Warrington, J., McPherson,J, Wasmuth, J., Le Paslier, D., Abderrahim, H., Cohen, D., Leppert, M., and White, R. (1991). Identification of deletion mutations and three new genes at the familial polyposis locus. Cell 66, 601-613.

Kaleko, M., Rutter, W. J. and Miller, A. D. (1990). Overexpression of the human IGFI receptor promotes ligand-independent neoplastic transformation. Mol. Cell. Biol. 10, 464-473.

Kaplan, D. R., Hempstead, B. L., Martin-Zanca, D., Chao, M. V. and Parada, L. F. (1991). The *trk* proto-oncogene product: a signal transducing receptor for nerve growth factor. Science 252, 554-558.

Kaplan, D. R., Morrison, D. K., Wong, G., McCormick, F. and Williams, L. T. (1990). PDGF ß-receptor stimulates tyrosine phosphorylation of GAP and association of GAP with a signaling complex. Cell 61, 125-133.

Kastan, M.B., Zhan, Q., El._Deiry, W.S., Carrier, F., Jacks, T., Walsh, W.V., Plunkett, B.S., Vogelstein, B., and Fornace, A.J. (1992). A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in Ataxia-Telangiectasia. Cell 71, 589-597.

Kern, S.E., Kinzler, K.W., Bruskin, A., Jarosz, D., Friedman, P., Prives, C., and Vogelstein, B. (1991). Identification of p53 as asequence-specific DNA-binding protein. Science 252, 1708-1711.

Kim, H.K, Kim, J. M., Zilberstein, A., Margolis, B., Kim, J.G., Schlessinger, J., and Rhee, S. G. (1991). PDGF stimulation of inositol phospholipid hydrolysis requires PLC-γ1 phosphorylation on tyrosine residues 783 and 1254. Cell 65, 435-441.

Klein, R., Jing, S., Nanduri, V., O'Rourke, E. and Barbacid, M. (1991). The *trk* proto-oncogene encodes a receptor for nerve growth factor. Cell 65, 189-197.

Klein, R., Nanduri, V., Jing, S., Lamballe, F., Tapley, P., Bryant, S., Reichardt, L. F. and Barbacid, M. (1991). The *trk*B tyrosine protein kinase is a receptor for brain-derived neurotrophic factor and neurotrophin-3. Cell 66, 395-403.

Koch, C. A., Anderson, D., Moran, M. F., Ellis, C. and Pawson, T. (1991). SH2 and SH3 domains: elements that control interactions of cytoplasmic signalling proteins. Science 252, 668-674.

Kolch, W., Heidecker, G., Lloyd, P., and Rapp, P. (1991). Raf-1 protein kinase is required for growth of induced NIH/3T3 cells. Nature 349:426-428.

Kornfeld, R. and Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. Annu. Rev. Biochem. 52, 631-664.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucl. Acids Res. 15, 8125-8131.

Krämer, H., Cagan, R. L. and Zipursky, L. (1991). Interaction of bride of *sevenless* membrane-bound ligand and the *sevenless* tyrosine-kinase receptor. Nature 352, 207-212.

Kress, C., Vogels, R., Graff, W.D., Bonnerot, C., Meijlink, F., Nicolas, J.-F., and Deschamps, J. (1990). Hox-2.3 upstream sequences mediate lacZ expression in intermediate mesoderm derivatives of transgenic mice. Development 109, 775-786.

Kratochwil, K. (1983). In Yamada, K.M. (ed). Cell interaction and development. John Wiley & Sons, NY, pp. 99-122.

Kypta, R. M., Goldberg, Y., Ulug, E. T. and Courtneidge, S. A. (1990). Association between the PDGF receptor and members of the *src* family of tyrosine kinases. Cell 62, 481-492.

Laemmli, U. K. (1970). Cleavage of structural proteins duing the assembly of the head of bacteriophage T4. Nature (London) 227,680-685.

Lamballe, F., Klein, R. and Barbacid, M. (1991). *trk*C, a new member of the *trk* family of tyrosine protein kinase, is a receptor for neurotrophin-3. Cell 66, 967-979.

Latif, F., and 32 others. (1993). Identification of the von Hippel-Lindau disease tumor suppressor gene. Science 260, 1317-1320.

LeDouarin, N., Bussonet, C., and Chaumont, F. (1968). Annu. Embryol. Morphol. 1, 29-39.

Leask, A., Rosenberg, M., Vassar, R., and Fuchs, E. (1990). Regulation of a human epidermal keratin gene: sequences and nuclear factors involved in keratinocyte-specific transcription. Genes Dev. 4, 1985-1998.

Lee, E. Y.-H.P., Chang, C.-Y., Hu, N., Wang, Y.-C. J., Lai, C.-C., Herrup, K., Lee, W.-H., and Bradley, A. (1992). Mice deficient for Rb are nonviable and shoe defects in neurogenesis and haematopoiesis. Nature 359, 288--294.

Lee, W.-H., Bookstein, R., Hong, F., Young, L.-J., Shew, J.-Y., and Lee, E.Y.-H.P. (1987). Human retinoblastoma susceptibility gene: cloning, identification, and sequence. Science 235, 1394-1399.

Liu, D., Rutter, W. J., and Wang, L.-H. (1993). Modulating effects of extracellular sequence of human insulinlike growth factor I receptor on its transforming and tumorigenic potential. J. Virol. 67, 9-18.

Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., Margolis, B., Lammers, R., Ullrich, R., Skolnik, J. Y., Bar-Sagi, D., and Schlessiger, J. (1992). The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to *ras* signaling. Cell 70, 431-442.

Margolis, B., Rhee, S. G., Felder, S., Mervic, M., Lyall, R., Loviteks, A., Ullrich, A. Zilberstein, A. and Schlessinger, J. (1989). EGF induces tyrosine phosphorylation of phospholipase c-II: a potential mechanism for EGF receptor signalling. Cell 57, 1101-1107.

Majerus, P. W., Ross, T. S., Cunningham, T. W., Caldwell, K. K., Jefferson, A. B. and Bansal, V. S. (1990). Recent insights in phosphatidylinositol signalling. Cell 63, 459-465.

Marshall, C.J. (1991). Tumor suppressor genes. Cell 64, 313-326.

Matsushime, H., Wang, L.-H. and Shibuya, M. (1986). Human c-*ros*-1 gene homologous to the v-*ros* sequence of UR2 sarcoma virus encodes a transmembrane receptor-like molecule. Mol. Cell. Biol. 6, 3000-3004.

Matsushime, H., and Shibuya, M. (1990). Tissue-specific expression of rat c-*ros*-1 gene and partial structural similarity of its predicted products with *sev* protein of *Drosophila melanogaster*. J. Viol. 64, 2117-2125.

Matthews, R. J., Bowen, D. B., Flores, E. and Thomas, M. L. (1992). Characterization of hematopoietic intracellular protein tyrosine phosphatases: Description of a phosphatase containing an SH2 domain and another enriched in proline-, glutamic acid-, serine-, and threonine-rich sequences. Mol. Cell. Bio. 12, 2396-2405.

Mayer, B.J. and Baltimore, D. (1993). Signalling through SH2 and SH3 domains. Trends in cell biology 3,8-13.

Mayer, B. J. and Hanafusa, H. (1990). Association of the v-crk oncogene product with phosphotyrosine-containing proteins and protein kinase activity. Proc. Natl. Acad. Sci. USA 87, 2638-2642.

Meisenhelder, J., Suh, P. G., Rhee, S. G. and Hunter, T. (1989). Phospholipase C-γ is a substrate for the PDGF and EGF receptor proteintyrosine kinases *in vivo* and *in vitro*. Cell 57, 1109-1122.

Millauer, B., Wizigmann-Voos, S., Schnurch, H., Martinez, R., Muller, N. P. H., Risau, W. and Ullrich, A. (1993). High affinity VEGF binding and developmental expression suggest Flk-1 as a major regulator of Vasculogenesis and angiogenesis. Cell 72, 835-846.

Mitchell, P.J. and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 245, 371-378.

Montesano, R., Matsmoto, K., Nakamura, T., and Orci, L. (1991). Identification of a fibroblast-derived epithelial morphgen as hepatocyte growth factor. Cell 67, 901-908.

Morrison, D. K., Kaplan, D.R., Escobedo, J. A., Rapp, U. R. Roberts, T. M., and Williams, L. T. (1989). Direct activation of the serine/threonine kinase activity of Raf-1 through tyrosine phosphorylation by the PDGF ß-receptor. Cell 58,649-657.

Morrison, D. K., Kaplan, D. R., Rapp, U. and Roberts, T. M. (1988). Signal transduction from membrane to cytosol: growth factors and membrane-bound oncogene products increase *raf*-1 phosphorylation and associated protein kinase activity. Proc. Natl. Acad. Sci. 85, 8855-8859.

Morrison, D. K., Kaplan, D. R., Rhee, R. G., and Williams, L. T. (1990). PDGF- dependent association of PLCγ with the PDGF receptor signaling complex. Mol. Cell. Biol. 10, 2359-2366.

Mugrauer, G. and Ekblom, P. (1991). Contrasting expression patterns of three members of the myc family of proto-oncogenes in the developing and adult mouse kidney. J. Cell Biol. 112, 13-25.

Nagarajan, L., Louie, E., Tsujimoto, Y., Balduzzi, P., Huebner, K. and Croce, C. M. (1986). The human c-*ros* gene (ROS) is located at chromosome region 6q16-6q22. Proc. Natl. Acad. Sci. USA. 83, 6568-6572.

Naldini, I., Vigna, E., Narsimhan, R. P., Gaudino, G., Zarnigar, R., Mickalopoulos, G. K. and Comoglio, P. M. (1991). Hepatocyte growth factor (HGF) stimulates the tyrosine kinase activity of the receptor encoded by the proto-oncogene c-MET. Oncogene 6, 501-504.

Nechameyer, W. S. and Wang, L.-H. (1984). Molecular cloning and characterization of avian sarcoma virus UR2 and comparison of its transforming sequence with those of other avian sarcoma viruses. J. Virol. 50, 914-921.

Neckameyer, W. S. and Wang, L.-H. (1985). Nucleotide sequence of avian sarcoma virus UR2 and comparison of its transforming gene with other members of the tyrosine protein kinase oncogene family. J. Viol. 53, 879-884.

Neckameyer, W. S., Shibuya, M., Hsu, M. T. and Wang, L.-H. (1986). Proto-oncogene c-*ros* codes for a molecule with structural features common to those of growth factor receptors and displays tissue-specific and developmentally regulated expression. Mol. Cell. Biol. 6, 1478-1486.

Olivier, J. P., Raabe, T., Henkemeyer, M., Dickson, B., Mbamalu, G., Margolis, B., Schlessinger, J., Hafen, E. and Pawson, T. (1993). A drosophila SH2-SH3 adaptor protein implicated in coupling the sevenless tyrosine kinase to an activator of *ras* guanine nucleotide exchange, sos. Cell 73, 179-191.

Otsu, M., I. Hiles, I. Gout, M. J. Fry, F. Ruiz-Larrea, G. Panayotou, A. Thompson, R. Dhand, J. Hsuan, N. Totty, A. D. Smith, S. J. Morgan, S. A. Courtneicge, P. J. Parker and M. D. Waterfield. (1991). Characterization of two 85-kd proteins that associate with receptor tyrosine kinases, middle-T/pp60^{v-src} complexes, and PI3-kinase. Cell 65, 91-104.

Peles, E., Bacus, S.S., Koshi, R.A., Lu, H.S., Wen, D., Ogden, S.G., Ben Levy, R., and Yarden, Y. (1992). Isolation of the neu/Her-2 stimulatory ligand: a 44kd glycoprotein that induces differentiation of mammary tumor cells. Cell 69, 205-216.

Perkins, L.A., Larsen, I., and Perrimon, N. (1992). *corkscrew* encodes a putative protein tyrosine phosphatase that functions to transduce the terminal signal from the receptor tyrosine kinase torso. cell 70, 225-236.

Plutzky, J., Neel, B. G. and Rosenberg, R. D. (1992). Isolation of a src homology 2-containing tyrosine phosphatase. Proc. Natl. Acad. Sci. USA 89, 1123-1127.

Podell, S. B. and Sefton, B. M. (1987). Chicken proto-oncogene c-*ros* cDNA clones: identification of a c-*ros* RNA transcript and deduction of the amino acid sequence of the carboxyl terminus of the c-*ros* product. Oncogene 2, 9-14.

Poon, B., Dixon, D., Ellis, L., Roth, R. A. and Rutter, W. J. (1991). Molecular basis of the activation of the tumorigenic potential of *gag*-insulin receptor chimeras. Proc. Natl. Acad. Sci. USA 88, 877-881.

Price, J. V., Clifford, R. J. and Schupbach, T. (1989) The maternal ventralizing locus *torpedo* is allelic to *faint little ball,* an embryonic lethal and encodes the *Drosophila* EGF receptor homolog. Cell 56, 1085-1092.

Pritchard-Jones, K., Fleming, S., Davidson, D., Bickmore, W., Porteous, D., Gosden, c., Bard, J., Buckler, A., Pelletier, J., Housman, D., van Heyningen, V., and Hastie, N. (1990). The candidate Wilm's tumor gene is involved in genitourinary development. Nature 346, 194-197.

Raycroft, L., Wu, H., and Lozano, G. (199). Transcriptional by wild type but not transforming mutants of the p53 anti-oncogene. Science, 249, 1049-1051.

Reddy, E. P., Skalka, A. M. and Curran, T. (Eds) (1988). The oncogene Handbook, Elsevier Science, Publishers.

Reinke, R. and Zipursky, S. L. (1988). Cell-cell interaction in the *Drosophila* retina: the bride of *sevenless* gene is required in photoreceptor cell R8 for R7 cell development. Cell 55, 321-330.

Rogge, R.D., Karlovich, C.A. and Banerjee, U. (1991). Genetic dissection of a neurodevelopmental pathway: *Son of sevenless* functions downstream of the *sevenless* and EGF receptor. tyrosine kinases. Cell 64, 39-48..

Rol'nik, V.V. (1970). Bird Embryology. Academy of Sciences of the USSR.

Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T. and Bowtell D. (1993). The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. Nature 363, 83-85)

Rubin, G. M. (1990). Development of the *Drosophila* retina: inductive events studied at single cell resolution. Cell 57, 519-520.

Ruderman, N. B., Kapeller, R., White, M. F. and Cantley, L. C. (1990). Activation of phosphotidylinositol 3-kinase by insulin. Proc. Natl. Acad. Sci. USA 87, 1411-1415.

Sadowski, I., Stone, J. C. and Pawson, T. (1986). A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinase modifies the kinase function and transforming activity of fujinami sarcoma virus P130$^{gag\text{-}fps}$. Mol. Cell. Biol. 6, 4396-4408.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular cloning, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. sci. USA 74, 5463-5467.

Schejter, E. D. and Shilo, B. Z. (1989). The Drosophila EGF receptor homolog (DER) gene is allelic to *faint little ball,* a locus essential for embryonic development. Cell 56, 1093-1104.

Schu, P.V., Takegawa, K., Fry, M.J., Stack, J.H., Waterfield, M.D., and Emr, S.D. (1993). Phosphotidylinositol 3-kinase encoded by yeast *vps*34 gene essential for protein sorting. Science 260, 88-91.

Settleman, J., V. Narasimhan, L.C Foster, and R.A. Weinberg. (1992). Molecular Cloning of cDNAs encoding the GAP-associated protein p190: implcations for a signaling pathway from ras to the nucleus. Cell 69, 539-549.

Sharma, S., Birchmeier, C., Nikawa, J., O'neill, K., Rodgers, L. and Wigler, M. (1989). Characterization of the *ros*-1-gene products expressed in human glioblastoma cell lines. Oncogene Res. 5, 91-100.

Sharmeen, L., and Taylor, J. (1987). Enzymatic synthesis of RNA oligonucleotides. Nucl. Acid. Res. 15, 6705-6711.

Shaw, G. and Kamen, R. (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46, 659-667.

Shen, S.-H., Bastein, L., Posner, B. I. and Chretien, P. (1991). A protein-tyrosine phosphatase with sequence similarity to the SH2 domain of the protein-tyrosine kinases. Nature 352, 736-739.

Sherr, C. J., Rettenmier, C. W., Sacca, R., Roussel, M. F., Lock, A. T. and Stanley, E. R. (1985). The *c-fms* proto-oncogene product is related to the receptor for the mononuclear phagocyte growth factor CSF-1. Cell 41, 665-676.

Shibuya, M., Hanafusa, H. and Balduzzi, P. C. (1982). Cellular sequences related to three new onc genes of avian sarcoma virus (fps, yes and ros) and their expression in normal and transformed cells. J. Virol. 42, 143-152.

Simon, M. A., Dodson, G. S. and Rubin, G. M. (1993). An SH3-SH2-SH3 protein is required for p21 activation and binds to sevenless and sos proteins in vitro. Cell 73, 169-177.

Simon, M. A., Bowtell, D. L. and Rubin G. R. (1989). Structure and activity of the *sevenless* protein. Proc. Natl. Acad. Sci. USA 86, 8333-8337.

Simon, M.A., Bowtell, D.D.L., Dodson, G.S., Laverty, T.R. and Rubin G.M. (1991). Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the *sevenless* protein tyrosine kinase. Cell 67, 701-716.

Skolnik, E. Y., Margolis, B., Mohammadi, M., Lowenstein, E., Fischer, R., Drepps, D., Ullrich A., and Schlessinger, J. (1991). Cloning of P13 kinase-associated p85 utilizling a novel method for expression/cloning of target proteins for receptor tyrosine kinases. Cell 65, 83-90.

Smale, S. T., and Baltimore, D. (1989). The "initiator" as a transcription control element. Cell 57, 103-113.

Snape, A. M., Jonas, E. A., and Sargent, T. D. (1990). KTF-1, a transcrptional activator of *Xenopus* embryonic keratin expression. Vevelopment 109, 157-165.

Sonnenberg, E., Godecke, A., Walter, B., Bladt, F. and Birchmeier, C. (1991). Transient and locally restricted expression of the *ros*1 protooncogene during mouse development. EMBO J. 10, 3693-3702.

Soppet, D., Escandon, E., Maragos, J., Middlemas, D. S., Reid, S. W., Blair, J., Burton, L. E., Stanton, B. R., Kaplan, D. R., Hunter, T., Nikolics, K. and Parada, L. F. (1991). The neurotrophic factors brain-derived neurotrophic factor and neurotrophin-3 are ligands for the *trk*B tyrosine kinase receptor. cell 65, 895-903.

Sprenger, F., Stevens, L. M. and Nusslein-Volhard, C. (1989). The *Drosophila torso* gene ennncodes a putative receptor tyrosine kinase. Nature 338, 478-483.

Squinto, S. P., Stitt, T. N., Aldrich, T. H., Davis, S., Bianco, S. M., Radziejewski, C., Glass, D. J., Masiakowski, P., Furth, M. E., Valenzuela, D. M., Distefano, P. S. and Yancopoulos, G. D. (1991). *trk*B encodes a functional receptor for brain-derived neurotrophic factor and neurotrophic-3 but not nerve growth factor. Cell 65, 885-893.

Storms, R. W. and Bose, H. R., Jr. (1989). Viral oncogenes and signal transduction. Virus Research 12, 251-282.

Sugawara, O., Oshimura, M., Koi, M., Annab, L. A. and Barret, J. C. (1990). Induction of cellular senescence in immortalized cells by human chromosome 1. Science 247, 707-710.

Sun, X. J., Rothenberg, P., Kahn, C. R., Backer, J. M., Araki, E., Wilden, P. A., Cahill, D. A., Goldstein, B. J., and White, M. F. (1991). Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. Nature 352, 73-77.

Sturgill, T. W., Ray, L. B., Erikson, E., and Maller, J. L. (1988). Insulin-stimulated MAP-2 kinase phosphorylates and activates ribosomal protein S6 kinase II. Nature 334, 715-718.

Sturgill, T. W., and Wu, J. (1991). Recent progress in characterization of protein kinase cascades for phosphorylation of ribosomal S6. Biochim. Biophys. Acta. 1092, 350-357.

Takeya, T. and Hanafusa, H. (1983). Structure and sequence of the cellular gene homologous to the RSV *src* gene and the mechanism for generating the transforming virus. Cell 32, 881-890.

Tessarollo, L., Nagarajan, L. and Parada, L. F. (1992). c-ros: the vertebrate homolog of the sevenless tyrosine kinase receptor is tightly regulated during organogenesis in mouse embryonic devlopment. Devlopment 115, 11-20.

Tiedge, H. (1991). The use of UV light as a Cross-linking agent for cells and tissue sections in *in situ* hybridization. DNA & Cell Biol. 10, 143-147.

Trofatter, J. A., MacCollin, M. M., Rutter, J. L., Murrell, J.R., Duyao,M. P., Parry, D. M., Eldridge, R., Kley, N., Menon, A. G., Pulaski, K., Haase,V. H., Ambrose, C. M, Munroe, D., Bove, C., Haines, J. L., Martuza, R. L., MacDonald, M. E., Seizinger, B. R., Short, M. P., Buckler,A. J., and Gusella, J. F. (1993). A novel moesin-, Ezrin-, Radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor.Cell 72, 791-800.

Ullrich, A., Bell, J. R., Chen, E. Y., Herrera, R., Petruzzelli, L. M., Dull, T. J., Gray, A., Coussens, L., Liao, Y.-C., Tsubokawa, M., Mason, A., Seeburg, P. H., Grumfeld, C., Rosen, O. M. and Ramachandran, J. R. (1985). Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. Nature 313, 756-761.

Ullrich, A., Coussens, L., Hayfield, J. S., Dull, J. S., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Mayes E. I., Whittle, N., Waterfield, M. D. and Seeburg, P. H. (1984). Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature 309, 418-425.

Ullrich, A., Gray, A., Tam, A. W., Yang-Feng, T., Tsubokawa, M., Jacobs, S., Francke, U., Ramachandran, J., Fujita-Yamaguchi, Y. (1986). Insulin-like growth factor I receptor primary structure: Comparison with insulin receptor suggests structural determinants that define functional specificity. EMBO J. 5, 2503-2512.

Ullrich, A., Shine, J., Chirgwin, J., Pietet, R., Tischer, E., Rutter, W. J. and Goodman, H. M. (1977). Rat insulin genes: construction of plasmids containing the coding sequences. Science 196, 1313-1319.

Ullrich, A., and Schlessinger, J. (1990). Signal transduction by receptors with tyrosine kinase activity. Cell 61, 202-212.

Valius M. and Kazlauskas, A. (1993). Phospholipase C-r1 and phosphotidylinositol 3 kinase are the downstream mediators of the PDGF receptor's mitogenic signal. Cell 73, 321-334.

Vennstrom, B. and Bishop, J. M. (1982). Isolation and characterization of chicken DNA homologous to the two putative oncogenes of avian erythroblastosis virus. Cell 28, 135-143.

Vogel, W., Lammers, R., Huang, J. and Ullrich, A. (1993). Activation of a phosphotyrosine phophatase by tyrosine phophorylation. Science 259, 1611-1614.

Wang, L.-H., Lin, B., Jong, S.-M., Dixon, D., Ellis, L., Roth, R. and Rutter, W. J. (1987). Activation of transforming potential of the human insulin receptor gene. Proc. Natl. Aca. Sci. USA 84, 5725-5729.

Wang, L.-H., and Duesberg, P. (1974). Properties and function of poly(A) in Rous sarcoma virus RNA. J. Virol. 14, 1515-1529.

Wang, L.-H., Feldman, F., Shibuya, M., Hanafusa, H., Notter, M. F. D. and Balduzzi, P. (1981). Genetic structure, transforming sequence and gene product of avian sarcoma virus UR1. J. Virol. 40, 258-267.

Wang, L.-H., Hanafusa, H., Noter, M. F. D. and Balduzzi, P. C. (1982). Genetic structure and transforming sequence of avian sarcoma virus UR2. J. Viol. 41, 833-841.

Wang, X.-F., Lin, H., Ng-Eaton, E., Downward, J., Lodish, H., and Weinberg, R. (1991). Expression cloning and characterization of the TGF-β Type III receptor. Cell 67, 797-805.

Waterfield, M. D., Huang, J. S. and Deuel, T. F. (1983). Platelet-derived growth factor is structurally related to the putative transforming protein $p28^{sis}$ of simian sarcoma virus. Nature 304, 35-39.

Watson, C. J. and Jackson, F. F. (1985). An alternative procedure for the synthesis of double-stranded for cloning in phage and plasmid vectors, p79-88. In D. M. Glover (ed.), DNA cloning, vol. 1. IRL Press, New York.

Wen, D., Peles, E., Cupples, R., Suggs, S.V., Bacus, S.S., Luo, Y., Trail, G., Hu, S., Silbiger, S.M., Ben Levy, R., Koski, R.A., Lu, H.S., and Yarden, Y. (1992). Neu differentiation factor: a transmemberane glycoprotein containing an EGF domain and an immunoglobulin containing unit. Cell 69, 559-572.

Williams, L. T (1989). Signal transduction by the platelet-derived growth factor receptor. Science 243, 1564-1570.

Woolford, J., McAuliffe, A. and Rohrschneider, L. R. (1988). Activation of feline c-*fms* proto-oncogene: multiple alterations are required to generate a fully transformed phenotype. Cell 55, 965-977.

Xu, G., O'Connel, P., Vischochil, D., Cawthon, R., Robertson, M., Culver, M., Dunn, D., Stevens, J., Gesteland, R., and White, R. (1990a). The neurofibromatosis type 1 gene encodes a protein related to GAP. Cell 62, 599-608.

Xu, G., Lin, B., Tanaka, K., Dunn, D., Gesteland, R., White, R., and Weiss, R. (1990b). The catalytic domain of the neurofibromatosis type 1 gene product stimulates *ras* GAPase and complements *ira* mutants of S. cerevisiae. Cell 63, 835-841.

Yarden, Y., and Ullrich, A. (1988). Growth factor receptor tyrosine kinase. Annu. Rev. Biochem. 57, 443-478.

Yi, T. L., Cleveland, J. L. and Ihle, J. N. (1992). Protein tyrosine phosphatase containing SH2 domains: Characterization, preferential expression in hematopoietic cells, and localization to human chromosome 12p12-p13. Mol. Cell. Bio. 12, 836-846.

Zong, S. Z., Chen, J., and Wang, L.-H. (1993). Molecular and biochemical basis for activation of the transforming potential of the transforming gene v-*ros*. Submitted to J. Virol.

Zsebo, K. M., Williams, D. A. and Suggs, S. V. (1990). Stem cell factor is encoded at the Sl locus of the mouse and is the ligand for the c-*kit* tyrosine kinase receptor. Cell 63, 213-224.