THE UNIVERSITY OF CALGARY

Cloning and Characterization of Proteins Antigenically Related to Sam68 Including a Putative Golgi Protein that Interacts with Src

by

Eko Wargonyo Raharjo

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF
BIOCHEMISTRY AND MOLECULAR BIOLOGY

CALGARY, ALBERTA
SEPTEMBER, 1998

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

0-612-34989-6

Canada

ABSTRACT

A number of proteins with diverse functions have been identified as substrates of the tyrosine kinase, Src. One nuclear protein known as Sam68 (Src associated in mitosis) has also been identified as a specific target of Src in mitosis in fibroblasts. The interaction of Sam68 with Src is thought to occur upon nuclear envelope breakdown during mitosis. However, the exact function of this interaction is currently unknown. A number of possible roles have been proposed, which include regulation of RNA processing, RNA splicing or possibly RNA trafficking through the nuclear membrane. Another protein with a molecular weight of 70 kDa, p70, has been observed to be phosphorylated on tyrosine and associated with SH2 domain of Src family member $p56^{lck}$ in T-cells stimulated with IL-2 or PHA. This latter protein has been shown to have sequence identity to the published sequence of Sam68.

In our attempt to isolate novel Sam68-related proteins and gain a better understanding of the function of Sam68 and its regulation by Src, two antibodies raised against the N- and C-terminal region of Sam68 were used separately to screen a human T-lymphocyte cDNA expression library. A number of cDNA clones were isolated and found to encode novel proteins which were antigenically related to Sam68. One cDNA clone, called C-4.3, possesses coiled-coil and proline-rich domains and exhibits a high degree of homology to golgin-95 and the *cis*-Golgi matrix protein GM130, proteins that are part of the Golgi complex. A protein with an apparent molecular mass of approximately 63 kDa was identified in Western blot analyses of NIH 3T3 whole cell lysates using a polyclonal antibody raised against a GST-C-4.3 fusion protein. C-4.3 was also observed to be localized to the Golgi complex of NIH 3T3 fibroblasts and co-localized with Src in some of these cells. Additionally, the GST-C-4.3 fusion protein associated with and was phosphorylated by purified baculovirus expressed c-Src *in vitro*. In summary, a number of cDNA clones encoding novel proteins and exhibiting antigenic relationship to Sam68 were isolated. Further protein characterization of a cDNA clone, C-4.3, suggests that it may be a putative target of Src that is involved in regulating certain events related to Golgi functions.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. D.J. Fujita for his academical advice and guidance. I would like to thank the members of my supervisory committee, Dr. P.W.K. Lee and Dr. S.M. Robbins, for their critical advice. I would like to thank Dr. B. Burke and Dr. W.H. Raharjo for their valuable technical advice especially in the producing and purifying antibodies and in the immunofluorescence experiments. I would like to thank Dr. J.B. Rattner for his valuable contribution in the localization and co-localization experiments. I would like to thank Dr. E. Chan at the Scripps Research Institute, La Jolla, California for providing valuable information relating to our research project. I would like to thank my colleagues Dr. Jeff Bjorge, Dr. Bob Qi and Dr. Mike Treuner for providing valuable technical advice; and Dr. Cay Egan, Dr. Yamini Achari, Raymond Lee, Dongping Ma, Paul Enarson and especially Andrew Jakymiw for their valuable suggestions and critical review in preparing this thesis. Finally, I would like to express my personal thank you to my wife, Indra, my son, Tio, and my friend, Otto, for their continuous support and encouragement during my time in pursuing the M.Sc. degree.

Dedicated to

my wife, Indra,

my son, Tio,

and my friend, Otto

TABLE OF CONTENTS

Approval page ii

Abstract iii

Acknowledgements iv

Dedication v

Table of Contents vi

List of Figures ix

List of Abbreviations xi

CHAPTER ONE: INTRODUCTION 1

A. Src tyrosine kinase and some of its activation mechanisms 1

B. Localization of Src and the regulation mechanisms 3

C. Src substrates and the role of SH2 and SH3 domains 4

D. Sam68, a Src target in mitosis 9

D.1. Sam68-related proteins 9

D.2. Structure and possible functions of Sam68 11

E. The objective of the study 14

CHAPTER TWO: MATERIALS AND METHODS 15

A. Immunological screening of λgt11 cDNA library 15

A.1. Plating bacteriophage 15

A.2. Transferring bacteriophage and antibody screening 15

B. DNA hybridization screening 16

C. Southern blot analysis 17

D. Cloning and subcloning 18

D.1. Nested deletion 18

D.2. Sequencing 19

E. Glutathione S-transferase (GST) fusion constructs 19

F. Expression and purification of GST-fusion proteins..... 20
G. Production of polyclonal antibodies..... 21
H. Purification of polyclonal antibodies..... 22
I. Western blot analysis..... 22
J. Localization and co-localization experiments..... 23
K. *In vitro* binding assay..... 24
L. *In vitro* kinase assay..... 24
M. Immunoprecipitation..... 24

CHAPTER THREE: RESULTS..... 26
A. Polyclonal antibodies raised against specific regions of Sam68 reacted with several proteins in Western blots of NIH 3T3 fibroblasts..... 26
B. Library screening using an antiserum raised against an N-terminal region of Sam68 resulted in the isolation of clones that have homology to RNA binding proteins..... 28
B.1. Cloning and sequence analysis of N-1.1, N-1.2 and N-5..... 28
B.2. Further cDNA library screening in order to obtain the full transcript of clone N-5..... 36
C. Library screening using an antiserum raised against C-terminal region of Sam68 resulted in the isolation of a clone, C-4.3, that exhibits proline-rich regions and a coiled-coil structure..... 36
C.1. Southern blot hybridization analysis of the isolated clones..... 37
C.2. Cloning, subcloning and sequencing of C-4.3, C-2.4 and C-4.1 42
C.3. Structural and sequence analyses of C-4.3..... 42
C.4. Structural and sequence analyses of clones C-2.4 and C-4.1..... 49
C.5. Further library screening in order to obtain the full transcript of clones C-2.4 and C-4.1..... 54
D. Characterization of C-4.3 protein..... 59
D.1. Expression and purification of the GST-C-4.3 protein..... 59

D.2. Anti-C-terminal Sam68 polyclonal antibody reacted with GST-C-4.3 and GST-C-2.4 fusion proteins 62
D.3. Production and purification of anti-C-4.3 polyclonal antibody 66
D.4. The C-4.3 protein exhibited a molecular weight of 63 kDa and was localized in the Golgi complex 66
D.5. C-4.3 associated with and was phosphorylated by c-Src *in vitro* 68
D.6. C-4.3 did not co-immunoprecipitate with Src in NIH 3T3 fibroblasts 75
D.7. Src was co-localized with C-4.3 in the perinuclear region of NIH 3T3 fibroblasts 75
E. Summary of the results 78

CHAPTER FOUR: DISCUSSION 80
A. Structural and antigenic relationship between the isolated clones and Sam68 80
B. Coiled-coil motif in C-4.3 and its possible functions 82
C. Proline-rich regions of C-4.3 and its possible functions 85
D. C-4.3, a putative substrate of Src tyrosine kinase 86
E. Future perspective on C-4.3 88

BIBLIOGRAPHY 90

APPENDIX 103
The complete sequence of C-4.3 103
Sequence of N-1.1 107
Sequence of N-1.2 108

LIST OF FIGURES

Figure 1.	Schematic representation of Src protein	2
Figure 2.	Schematic representation of Src activation and some responses that are elicited by Src tyrosine kinase action	5
Figure 3.	Schematic representation of Sam68 structure	12
Figure 4.	Anti-C-Sam68 polyclonal antibody reacted with several proteins in a Western blot from NIH 3T3 fibroblasts	27
Figure 5.	Schematic representation of the screening process of a T-lymphocyte cDNA expression library using an anti-N-terminal specific Sam68 antibody	29
Figure 6.	Sequence analysis of clone N-5	32
Figure 7.	Alignments between N-5, human ribosomal protein S5 and several RNA binding proteins	34
Figure 8.	Schematic representation of the screening process of a T-lymphocyte cDNA expression library using an anti-C-terminal region specific Sam68 antibody	38
Figure 9.	Southern blot analyses of the cDNA clones isolated by screening of a T-lymphocyte cDNA expression library using an anti-C-terminal specific Sam68 antibody	40
Figure 10.	Nested deletion products of the cDNA clone C-4.3	43
Figure 11.	Schematic representation of amino acid sequence and structural analysis of clone C-4.3	44
Figure 12.	Schematic representation of sequence homology between C-4.3 and golgin-95	47
Figure 13.	Sequence analysis of clone C-2.4	50
Figure 14.	PC GENE alignment between C-2.4 and the sequence of an annexin V binding protein and a hydropathy analysis of C-2.4 sequence	52

Figure 15. Sequence analysis of clone C-4.1 55
Figure 16. Southern blot analyses of the clones isolated from a DNA hybridization screening using C-4.1 and C-4.2 probes 57
Figure 17. Glutathione S-transferase (GST) fusion protein constructs 60
Figure 18. Purification of the expressed GST-fusion proteins 63
Figure 19. Anti-C-terminal Sam68 polyclonal antibody reacted with GST-C-4.3 and GST-C-2.4 fusion proteins 65
Figure 20. A polyclonal antibody raised against a GST-C-4.3 fusion protein recognized a 63 kDa protein in Western blot analysis of NIH 3T3 fibroblasts 67
Figure 21. C-4.3 localized to the Golgi and co-localized with Src at the perinuclear membrane in NIH 3T3 fibroblasts 69
Figure 22. GST-C-4.3 associated with purified baculovirus expressed c-Src *in vitro* 72
Figure 23. GST-C-4.3 phosphorylation by purified baculovirus expressed c-Src *in vitro* 74
Figure 24. Immunoprecipitations of C-4.3 and Src proteins from NIH 3T3 fibroblast lysates 76

LIST OF ABBREVIATIONS

ABP7	Annexin binding protein 7
ATP	Adenosine tri phosphate
BFA	Brefeldin A
BSA	Bovine serum albumin
CENP	Centromeric motor protein
CI-MPR	Cation-dependent mannose-6-phosphate receptor
CNBr	Cyanogen bromide
CSK	Carboxyl-terminal of Src kinase
DNA	Deoxyribonucleic acid
cDNA	Complementary DNA
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethilene diamine tetra-acetic acid
EGF	Epidermal growth factor
ER	Endoplasmic reticulum
FMR-1	Fragile X mental retardation-1
GAP	GTP-activating protein
GST	Glutathione S-transferase
GTP	Guanosine triphosphate
IgG	Immunoglobulin G
IIF	Indirect immunofluorescence
IL-2	Interleukin-2
IP	Immunoprecipitation
IPTG	Isopropyl-1-thio-β-D-galactopyranoside
KH	hnRNP K homology
LB	Luria bertani
MTOC	Microtubule organizing centre

NLS	Nuclear localization signal
NRK	Newborn rat kidney
NZW	New Zealand white
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDGF	Platelet-derived growth factor
PHA	Phytohemagglutinin
PI3	Phosphatidylinositol 3
PLCγ	Phospholipase C gamma
PMSF	Phenylmethylsulfonyl fluoride
TBS	Tris-buffered saline
RNA	Ribonucleic acid
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SH1/2/3/4	Src homology one/ two/ three/ four
SLE	Systemic lupus erythematosus
SNRP	Small nuclear ribo-nucleoprotein
XNP	X-linked nuclear protein

CHAPTER ONE

INTRODUCTION

A. Src tyrosine kinase and some of its activation mechanisms

The cellular Src gene was the first molecularly-defined proto-oncogene, and its product, c-Src (Figure 1), is known to be the source of the first detected tyrosine phosphorylation event. As a model tyrosine kinase and because of its unique regulation and functions, Src remains a primary focus of many investigations (Brown and Cooper, 1996). Studies of the Src tyrosine kinase and its subtrates represent a major contribution to our understanding of the role of protein phosphorylation in the control of normal cell growth and differentiation as well as in the development of neoplastic disease (Cooper, 1990).

Phosphorylation of a regulatory site in the carboxy-terminal tail of Src at residues Tyr 527 in chicken Src, or Tyr 530 in human Src, by an enzyme known as carboxy-terminal Src kinase (CSK) is thought to render Src to be inactive (Kmiecik and Shalloway, 1987; Cooper *et al.*, 1986; Tanaka and Fujita, 1986). This inactivation of Src is due to an intramolecular interaction between the phosphotyrosine and a Src homology two (SH2) domain within the amino-terminal half of the molecule (Superti-Furga *et al.*, 1993). On the other hand, dephosphorylation of the phosphorylated tyrosine 527 or 530, by tyrosine phosphatases is thought to result in the activation of Src. The tyrosine phosphatases responsible for this activation, however, have not yet been identified (Zheng *et al.*, 1992).

Elevation of Src activity has been observed in cells undergoing mitosis. This activation of Src is attributed to serine and threonine phosphorylations within the amino-

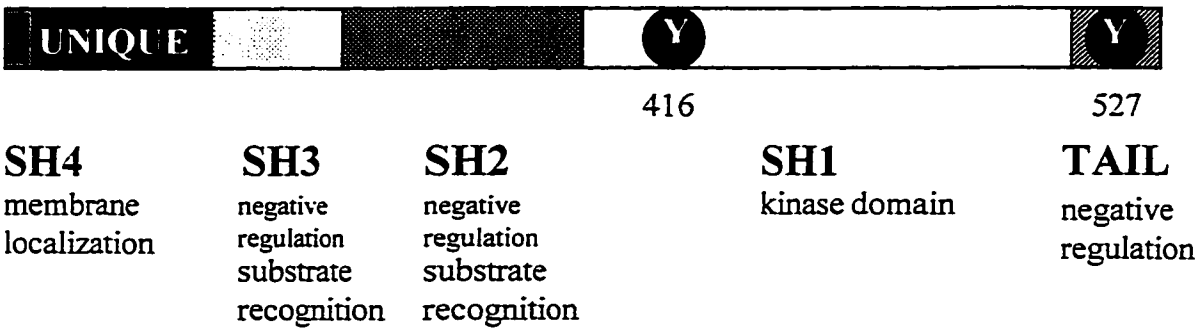


Figure 1. Schematic representation of Src protein. The protein of chicken Src is drawn approximately to scale, showing the positions and functions of the SH4 domain, unique region, SH3 domain, SH1 domain (kinase) which contains the autophosphorylation site (Y416) and the tail region which contains the CSK phosphorylation site (Y527). (Based on Brown and Cooper, 1996. Biochim. Biophys. Acta 1287, 121-149).

terminal portion of Src, which is thought to enhance the dephosphorylation of Src carboxy-terminal regulatory site (Morgan *et al.*, 1989; Chackalaparampil and Shalloway, 1988). Additionally, Src tyrosine kinase activitiy can also be elevated by growth factor stimulation, such as by platelet derived growth factor (PDGF) (Kypta *et al.*, 1990). This activation of Src is due to the binding of the SH2 domain of Src to a phosphotyrosine in PDGF receptors, which displaces the intramolecular interaction of the SH2 domain of Src with its regulatory carboxy-terminal phosphotyrosine (Alonso *et al.*, 1995). The binding of phosphotyrosine-containing proteins to the Src SH2 domain and displacement of the regulatory carboxy-terminus may be a general mechanism for Src activation (Bjorge *et al.*, 1996) (Figure 2).

B. Localization of Src and the regulation mechanisms

Immunofluorescence and biochemical fractionation studies in fibroblast cells have suggested that both c-Src and v-Src are localized to perinuclear and plasma membranes (Courtneidge *et al.*, 1980; Rohrschneider, 1979). In some other cell types, Src and its related proteins, Fgr and Hck, were observed to be localized to the nucleus and cytoplasm respectively (David-Pfeuty and Nouvian-Dooghe, 1995; Lowell *et al.*, 1994). Kaplan *et al.*, (1992) demonstrated that in mammalian fibroblasts, a significant proportion of Src protein was found at the microtubule organizing centre (MTOC) and co-localized with cation-dependent mannose-6-phosphate receptor (CI-MPR), a marker of late endosomes. Furthermore, during mitosis Src was observed to cluster at dividing centrosomes. These distinct localization pattern of Src and its related kinases in different cells suggest that localization is more complicated than simply targeting by the SH4 domain (an N-terminal

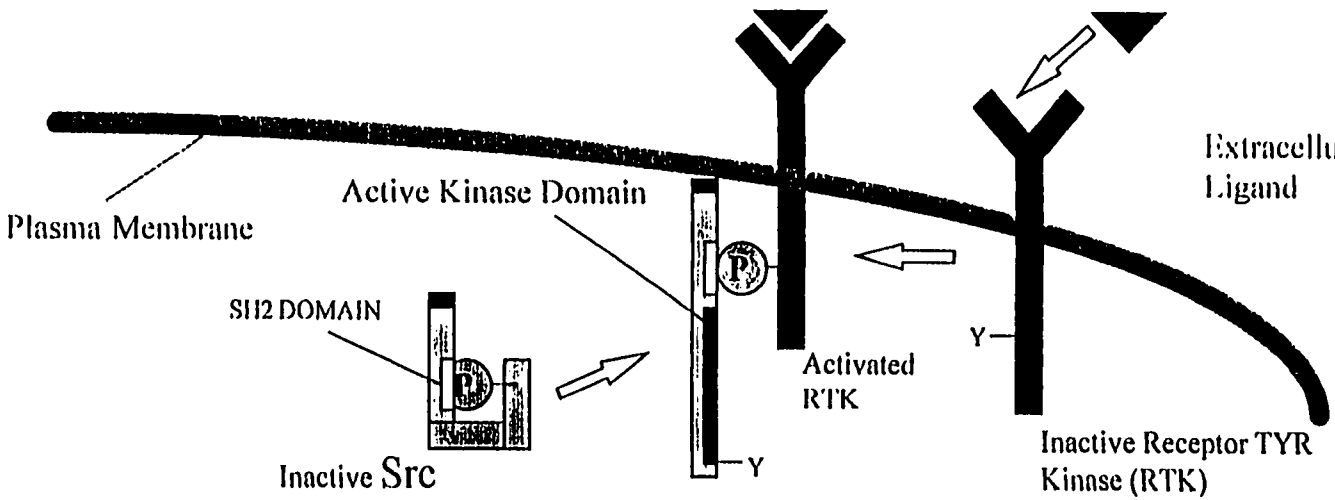
region of Src that is thought to be critical for membrane localization) to a generic membrane environment (Brown and Cooper, 1996).

Localization of Src family kinases seems to be regulated not only by SH4, but also by SH2 and SH3 domains. Mutation of SH2 and SH3 domains differentially affect localization to perinuclear and peripheral membranes (Kaplan *et al.*, 1990). However, the conformation of the Src molecule is also critical for its localization. It has been observed that the SH2 domain of Src which binds to phosphotyrosines is unlikely to be needed for localization of kinase-defective Src to adhesion plaques. This suggests that an activated-'open' conformation of Src, is necessary for proper Src localization to adhesion plaques (Kaplan *et al.*, 1994). Lastly, Liebl and Martin (1992) reported that the translocalization of activated Src to focal adhesion sites was critical for transforming activities. In contrast, directing Src to the nucleus or perinuclear membranes (e.g. endoplasmic reticulum) prevented transformation (Liebl and Martin, 1992).

C. Src substrates and the role of SH2 and SH3 domains

A number of proteins with diverse functions have been identified as subtrates of Src. Some have been shown to associate with and become phosphorylated by Src and others only display increased tyrosine phoshorylation levels in Src transformed cells (Hunter and Cooper, 1986). In a review by Cooper and Brown (1996), Src substrates are grouped into two classes. One class comprises the proteins that are also phosphorylated in cells stimulated through receptor tyrosine kinases, such as PDGF and EGF receptors, and are implicated in mitogenic signalling pathways. This class of substrates includes phospholipase Cγ (PLCγ),

Figure 2. Schematic representation of Src activation and some responses that are elicited by Src tyrosine kinase action. Binding of the SH2 domain of Src to a phosphotyrosine in growth factor receptors results in activation of Src tyrosine kinase. The activated Src mediates cell signaling of many cellular pathways leading to diverse biological functions.



Activated Src

- PI 3-kinase activation
- Ras/MAPK activation
- Mitotic functions
- Cytoskeletal reorganization
- Other functions
  - Protein trafficking

RasGAP, the tyrosine phosphatase SHPTP2 or syp, the adaptor protein Shc, and the p85 regulatory subunit of PI3 kinase. The other one comprises the proteins that have been studied primarily because they are phosphorylated in Src-transformed cells. Many Src substrates in this class have also been shown to bind to Src through the Src SH2 and SH3 domains. Therefore, these proteins may likely be direct substrates in the cell. This class of Src substrates includes p110 (AFAP110), cortactin, Fak, paxillin, tensin, talin, vinculin, $p130^{cas}$ (Flynn *et al.*, 1993; Kanner *et al.*, 1991; Wu *et al.*, 1991; Kanner *et al.*, 1990; Reynolds *et al.*, 1990a, 1990b) and Sam68, a new class of Src substrate that has the ability to bind to RNA (Fumagalli *et al.*, 1994; Taylor and Shalloway, 1994). These diverse substrates of Src reflect its involvement in multiple functions (Brown and Cooper, 1996).

The selectivity of Src substrates, especially class two substrates is directed by SH2 and SH3 domains of Src (Brown and Cooper, 1996). SH2 domains have the ability to interact with specific tyrosine-phosphorylated proteins (Anderson *et al.*, 1990; Moran *et al.*, 1990). According to Songyang *et al.* (1993) the optimal peptide sequence that is preferred by Src and its related kinases, Fyn and Lck, to bind to is pYEEI. This *in vitro* prediction of the binding specificity is supported by the molecular nature of the two molecules. Src SH2 possesses deep hydrophobic pockets positioned to accept the I (isoleucine) at position pY + 3. Similarly, the acidic residues at pY + 1 and pY + 2 are spread apart on the surface of the SH2 and point towards basic residues on the otherwise neutral surface. Indeed, some Src - binding proteins contain the predicted optimal binding sequence pYEEI, or similar sequences. Hamster polyoma virus middle T and the focal adhesion kinase, Fak, are phosphorylated and bind to SH2 domains of Src through Tyr 324 (pYEEI) (Songyang *et al.*,

1993) and Tyr 397 (pYAEI) respectively (Schaller *et al.*, 1994; Mori *et al.*, 1993).

Association of Src with the p85 subunit of PI3 kinase has been demonstrated to be mediated by its SH3 domain (Liu *et al.*, 1993; Vogel and Fujita, 1993). The sites responsible for the binding to Src SH3 domains has been identified to be proline-rich sequences on p85 subunit of PI3 kinase (Kappeler *et al.*, 1994; Liu *et al.*, 1993; Prasad *et al.*, 1993; Vogel and Fujita, 1993). The structural analysis of Src SH3 revealed a globular domain, one side of which has a slightly depressed hydrophobic surface with an acidic cluster at one end (Yu *et al.*, 1994). Two consensus sequences known as class I and II binding sites were identified to bind with high affinity to SH3 domains. They are R-P-x-q-P-x-q and q-P-x-q-P-x-R, where x can be any amino acid, but most often is a P (proline), V (valine) or L (leucine). These two ligands have different polarity in their binding to SH3 domains, in which class I ligands have the opposite orientation with the respect to class II ligands (Mayer and Eck, 1995; Feng *et al.*, 1994). While the proline stretch contacts the hydrophobic surface of the SH3 domain, the first arginine contacts the acidic patch. The prolines are separated by two amino acids, such that it enables them to reside on the same face of the helix, which is important for binding. Furthermore, specificity and affinity may be enhanced by the presence of repeated SH3-binding sites in a single protein, since it raises the local concentration of binding sites.

In the case of PI3 kinase, there are two SH3 binding sites in its p85 regulatory subunit and the association complex is stable as shown by immunoprecipitation methods, implying a slow disssociation rate (Liu *et al.*, 1993). However, the SH3-polyproline-mediated association of Src and its substrate, Shc, can only be proven *in vitro*, suggesting that this

association may not be as stable. Since tyrosine phosphorylation of Shc in *v-Src*-transformed cells is a likely cause of its activation, there is no doubt that Shc is a substrate of Src (Egan *et al.*, 1993; McGlade *et al.*, 1992; Rozakis-Adcock *et al.*, 1992). Presumably, there may be many more Src substrates with this type of unstable binding that have not yet been identified.

D. Sam68, a Src target in mitosis

In fibroblasts, Src family kinases are activated not only at the G0-G1 transition but also at the onset of mitosis. Additionally, in the G2 phase, Src is observed to be translocated from its plasma membrane location and is present in diffuse and patchy structure thought to be endocytic vesicles, throughout the cytoplasm in mitosis (David-Pfeuty and Nouvian-Dooghe, 1990). In a search for mitotic target of Src two groups identified a protein of approximately 68 kDa, later known as Sam68, that was specifically phosphorylated on tyrosine during mitosis, both in normal and in *v-Src*-transformed cells (Fumagalli *et al.*, 1994; Taylor and Shalloway, 1994). Not only was Sam68 shown to be a substrate of Src kinase *in vitro*, it also associated physically with Src derived from mitotic lysate *in vivo*. The nature of binding was shown by the use of GST fusion proteins containing parts of Src, in which Sam68 bind to both the SH3 domain of Src and, when phosphorylated, to the SH2 domain of Src. Since Sam68 was identified as a nuclear protein, it was suggested that the association with Src occurs when the nuclear envelope breakdowns during mitosis (Courtneidge and Fumagalli, 1994; Fumagalli *et al.*, 1994; Taylor and Shalloway, 1994).

D.1. Sam68-related proteins

Based on antigenic cross-reactivity and sequence homology, Sam68 was thought to be related to GAP-p62, or possibly a product of alternative splicing from the same gene (Fumagalli *et al.*, 1994; Taylor and Shalloway, 1994). GAP-p62 was first identified by Ellis *et al.* (1990) as a protein that was phosphorylated on tyrosine in *v-src*-transformed cells and several other types of transformed cells, and that associated with the Ras GTPase-activating protein (RasGAP). However, both Sam68 and GAP-p62 showed distinct features; for example, GAP-p62 was phosphorylated in the G1 phase during the response to particular growth factors (also in cells transformed by several oncogenes) and associated with the SH2 domain of RasGAP in a phosphotyrosine-specific fashion (Moran *et al.*, 1991). In contrast, Sam68 was phosphorylated on tyrosine only during mitosis, and associated with Src through presumably an SH3-domain specific mechanism. It was finally confirmed that there was no antigenic relation between the two proteins. Furthermore, Lock *et al.* found that GAP-p62 cDNA (Wong *et al.*, 1992) encoded the Sam68 protein, not the GAP-p62 protein (Lock *et al.*, 1996; Courtneidge and Fumagalli, 1994). Therefore, Sam68 and GAP-p62 have been characterized as unrelated proteins.

It is possible that there may be other proteins that are structurally or functionally related to Sam68. Vogel and Fujita (1995) observed that a tyrosine-phosphorylated 70 kDa protein associated with the SH2 domain of the Src family member $p56^{lck}$ in peripheral blood lymphocytes stimulated by interleukin-2 (IL-2) or phytohemagglutinin (PHA), and partial peptide sequencing of the isolated protein revealed a 100% identity to Sam68 in two peptides of 20 and 15 amino acids. It is not clear whether p70 and Sam68 are the same protein, since

unlike Sam68 which is phosphorylated only in mitosis, p70 becomes phosphorylated just 5 minutes after IL-2 stimulation (Vogel and Fujita, 1995). Furthermore, a natural isoform of Sam68, known as Sam68 Δ KH, has been found in human tissues. This isoform contains a deletion of 38 amino acid residues within the KH domain, suggesting that this isoform arises from alternative splicing of a single pre-mRNA species (Barlat *et al.*, 1997). Finally, our laboratory observed that both monoclonal and polyclonal antibodies to Sam68 recognized proteins of many sizes other than 68 kDa in Western blots of lysate from several cell lines (unpublished data). These data suggest that several proteins antigenically related to Sam68 may exist.

D.2. Structure and possible functions of Sam68.

Sam68 exhibits a very interesting basic structure (Figure 3) which may reflect its biological functions within the cell. First of all, it has RNA binding regions known as the KH domain and RGG boxes within the N- terminal half of the protein (Gibson *et al.*, 1993).This RNA binding domain is an evolutionarily conserved sequence, which has been shown to be important in the RNA binding ability of FMR-1 and hnRNP K (Siomi *et al.*, 1994). A single point mutation in a conserved residue of the FMR-1 KH domain has been reported to be associated with severe fragile X syndrome (Siomi *et al.*, 1993; Verkerk *et al.*, 1993). The ability of Sam68 to bind to RNA has been shown by Shalloway and coworkers, in which it can bind to poly (U) RNA homopolymers *in vitro* (Taylor *et al.*, 1995). Additionally, the SELEX method was used to show that Sam68 binds with high affinity to specific RNA sequences containing a UAAA motif, which is similar to the mammalian polyadenylation

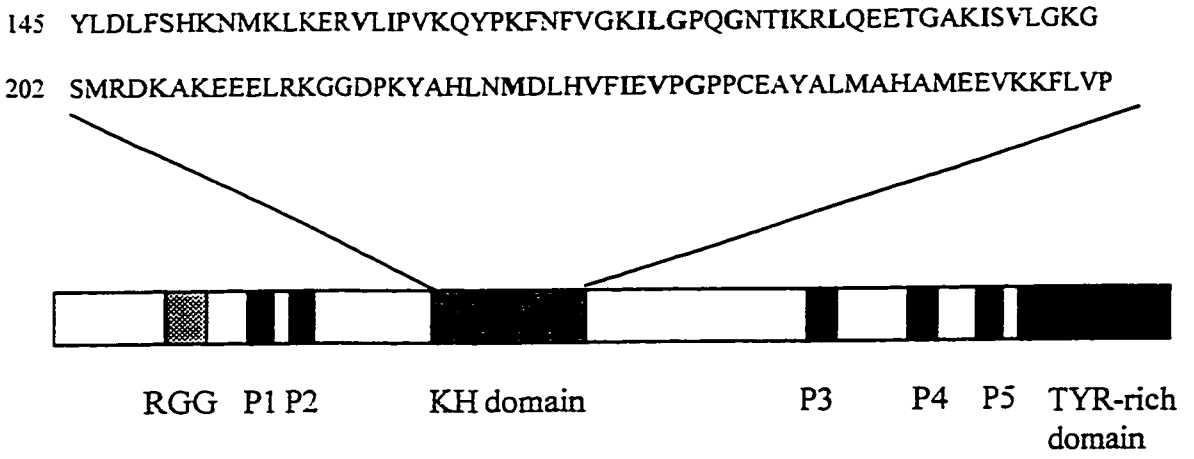


Figure 3. Schematic representation of Sam68 structure. Sam68 is composed of five consensus proline-rich motifs (P1, P2, P3, P4 and P5), RGG boxes, a tyrosine-rich domain and a KH homology domain (dotted box). Consensus residues in KH domain are indicated in bold (based on Barlat *et al.*, 1997. J. Biol. Chem. 272, 3129-3132).

signal. These data suggest that Sam68 plays a specific role in post-transcriptional regulation of RNA processing and gene expression. The fact that Sam68 is a substrate of Src tyrosine kinase raises the possibility that that mechanisms may also be regulated by Src tyrosine kinase or other related kinases (Lin *et al.*, 1997). Indeed, *in vitro* data suggest that the binding of Sam68 to RNA sequences is regulated by tyrosine phosphorylation, in which phosphorylation of this protein inhibits its binding to RNA sequences (Wang *et al.*, 1995).

The predicted sequence of Sam68 also has at least five proline-rich regions. The proline-rich motif is known to have a potential to be a docking site for SH3 domain-containing proteins (Pawson and Schlessinger, 1993). *In vitro* translated Sam68 was shown to bind selectively to recombinant SH3 domains, with the highest affinity for the Src and p85 SH3 domains. Additionally, through peptide competition as well as deletion analyses it was shown that proline-rich regions, especially at residues 289-306, were the binding sites for Src (Taylor *et al.*, 1995). However, the work of Taylor and Shalloway also suggests that once Sam68 is phosphorylated, its association with c-Src is mediated largely by the SH2 domain (Taylor and Shalloway, 1994). Several other SH2/SH3 proteins have been shown to associate with Sam68 such as the adaptor protein Grb2, the p85 subunit of PI3 kinase and PLCγ. This suggests another possible function of Sam68, in which it may act as scaffolding protein for c-Src during mitosis (Richard *et al.*, 1995; Taylor *et al.*, 1995).

A possible role of Sam68 in cell cycle progression is demonstrated by a natural isoform of Sam68 lacking of 38 amino acids within the KH domain, known as Sam68 ΔKH. This isoform was shown to inhibit serum-induced DNA synthesis and cyclin D1 expression, which could be overcome by Sam68 (Barlat *et al.*, 1997). Other evidence that Sam68 may

play a role in cell cycle regulation was shown in experiments using a Src tyrosine kinase inhibitor, known as radicicol. Mitosis-specific tyrosine phosphorylation inhibition of Sam68 by radicicol caused severe retardation of the exit of cells from mitosis (Pillay *et al.*, 1996). Recently, Sam68 was also reported to be a direct substrate of Cdc-2 kinase, raising the possibility that it may serve to integrate signals generated by both Cdc-2 and Src kinases during mitosis. However, the exact mechanisms of these events need to be further identified (Resnick *et al.*, 1997).

E. The objective of the study

The discussion above suggests that, firstly, Sam68 may have related proteins; secondly, the biological functions of Sam68 need to be further clarified; and, finally, the mechanism through which Src tyrosine kinase regulates Sam68 functions also needs further investigation. This research was primarily intended to isolate possible novel Sam68-related proteins. It was also assumed that the identification and characterization of Sam68-related proteins would further contribute toward the understanding of Sam68, in term of its functions, and its regulation by Src tyrosine kinase. The approach chosen in this study was by immunological screening of a cDNA expression library. Further studies involved the use of a Glutathione S-transferase (GST) fusion-bacterial expression system (Pharmacia) to express the proteins from isolated cDNA clones, and biochemical and immunohistochemical analyses to examine their association with Src tyrosine kinase.

CHAPTER TWO

MATERIALS AND METHODS

A. Immunological screening of λgt11 cDNA library

A λgt11 cDNA expression library derived from human T-lymphocytes (provided by S. Orkin, Yale) was screened with two polyclonal antibodies of Sam68. They were a polyclonal antibody raised against the N-terminal region corresponding to amino acids 103-281 of Sam68 (Santa Cruz) and a polyclonal antibody raised against the C-terminal region corresponding to amino acids 331-443 of Sam68 (Santa Cruz).

A.1. Plating Bacteriophage

Approximately 6 x 10^5 λgt11 pfu's for the screening with N-terminal Sam68 antibody and 1.25 x 10^6 of λgt11 pfu's for the screening with C-terminal Sam68 antibody were plated. The *E. coli* strain Y1090*hsdR* bacterial host was inoculated in LB medium pH 7.4-7.5 (Luria Bertani: 10 g bactotrypsin, 5 g yeast, 10 g NaCl in 1 L) containing 10 mM $MgSO_4$ and 0.2% maltose, and incubated overnight at 37°C in a shaker incubator to OD_{600} ~ 1.5. The λgt11 library with the required pfu's was then mixed with 500 μL of bacterial host. After 20 minutes incubation at 37°C, 8 mL of 0.75% top agar (7.5 g agarose in 1 L LB) was added, and finally the mixture was plated on a 150 mm petri dish containing 1.5% bottom agar (15 g bactoagar in 1 L LB medium).

A.2. Transferring bacteriophage and antibody screening

The plated library was incubated at 42°C for 6 hours and then a nitrocellulose membrane (Amersham) impregnated with 10 mM Isopropyl-1-β-D-galactopyranoside (IPTG) was laid on top of it. After 4 hours incubation at 37°C, the membrane was removed from the plate, rinsed once in Tris-buffered saline (TBS), and incubated with blocking solution (1% BSA and 0.05% Tween-20 in TBS) for 1 hour at room temperature. The membrane was then incubated with 0.1 μg/mL of Sam68 antibody for 45 min at room temperature. After 3 x 10 min washing in 0.1 % Tween-20 in TBS, the membrane was incubated for 30 min at room temperature with the secondary antibody, a 1:2,000 dilution (in TBS containing 0.1% BSA and 0.05% Tween-20) of donkey anti-rabbit Immunoglobulin G (IgG) conjugated with horse radish peroxidase. An enhanced chemiluminescence (ECL, Amersham) kit was used for detection of positive plaques. The membrane was incubated in ECL solution for 1 min, dried with filter paper, and then exposed to X-ray film for approximately 30 seconds.

B. DNA hybridization screening

In order to obtain the full length of cDNAs isolated from the immunological screening, a DNA hybridization screening was carried out. The same cDNA library as above was used for this purpose and approximately 1.25 x 10^6 pfu's were plated. As a probe, 25 ng DNA obtained by PCR-amplification of the phage clone was radioactively labelled with [γ-^{32}P] ATP using a random hexamer priming method (Megaprime, Amersham) following the protocol provided by the manufacturer. Plating of the bacteriophage was carried out as

described previously. After overnight incubation at 37°C, the plate was chilled at 4°C for 1 hour. A nylon membrane (Hybond-N, Amersham) was then laid on top of the agar for 1 min at room temperature. The membrane was then peeled off the agar and incubated for 5 min in denaturation solution (0.5 N NaOH, 1.5 M NaCl) and dipped into neutralization solution (1.5 M NaCl, 0.5 M Tris-HCl pH 7.4). After washing in 2x SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0 in 20x SSC), the filter was then dried at room temperature for 30 min. Cross-linking of the denatured DNA was achieved by baking the membrane in an 80°C oven for 1 hour. The filter was washed in 2x SSC for 5 min and then incubated in prewashing solution (5x SSC, 0.5% SDS, 1 mM EDTA, pH 8.0) for 1 hour at 37°C, and followed by incubation in prehybridization solution (6x SSC, 0.5% SDS, 100 μg/mL denatured Salmon sperm DNA, 50% Formamide and 5x Denhardt's reagent) for 1 hour at 37°C. Subsequently, a ^{32}P-labelled DNA probe was added and the hybridization mixture was incubated for 16 hour at 37°C. The membrane was then washed 2 x 5 min at room temperature with 2x SSC and 0.1% SDS. Finally, after washing once for 20 min in 1x SSC and 0.1% SDS, the membrane was dried and exposed to an autoradiography film.

C. Southern blot analysis

The DNAs subjected to Southern blot analyses were resolved in 1% agarose gel electrophoresis. The resolved DNAs were transfered onto a nylon membrane (Hybond-N, Amersham) following the protocol described in the Current Protocols in Molecular Biology (Ausubel *et al.*, 1990). The DNA probes were synthesized using the random hexamer priming method (Megaprime, Amersham) following the protocol provided by the

manufacturer. DNA probe hybridization was carried out with the same conditions as described previously.

Removal of probes from the hybridized membranes was carried out using moderate treatment (Ausubel *et al.*, 1990). The hybridized membranes were washed in 0.4 M NaOH for 30 min at 45°C. Then the membranes were rinsed twice in several hundred millilitres of moderate stripping solution (200 mM Tris-HCl pH 7.0, 0.1% SDS) for 10 min at room temperature. The result was monitored by autoradiography of the stripped membrane overnight.

D. Cloning and subcloning

DNA cloning of the isolated λgt11 phages was achieved by two ways. Firstly, by PCR amplification the cDNA inserts and subsequent digestion of the PCR products with the restriction enzyme Eco RI. The digested PCR products were then cloned into the Eco RI restriction site of the plasmid pBluescript. Secondly, by PCR amplification of the cDNA inserts and followed by cloning of the PCR products directly to a PCR cloning plasmids, TA-2.1. The primers used for PCR amplification were λgt11 forward and reverse primers. Each PCR reaction contained 1x PCR buffer (Gibco), 0.2 mM dNTP (Gibco), 1.5 mM $MgCl_2$ (Gibco), 4% Dimethylsulfoxide (DMSO), 20 pmol of each primer, 2 μL DNA template and 1U of Taq DNA Polymerase (Pharmacia). The Taq DNA polymerase was added after the PCR reaction was heated on the thermocycler (Perkin Elmer) at 97°C for 5 min. The PCR was programmed for 30 cycles, each cycle consisted of 1 min at 95°C, 3 min at 65°C and 3 min at 72°C.

D.1. Nested deletion

A nested deletion approach was applied to a clone, C-4.3, that exhibited a molecular size of 1.9 kb. This procedure was carried out using a double-stranded nested deletion kit (Pharmacia Biotech) following the protocol provided by the manufacturer. A TA-2.1 plasmid bearing the C-4.3 insert was digested with Bam HI to create the nuclease-susceptible end and with Kpn I to generate the nuclease-resistant end. Based on the manual provided by Pharmacia, the C-4.3 insert was fragmented into twelve pieces with intervals of 150-200 bp and subcloned back into the TA-2.1 plasmid.

D.2. Sequencing

Manual sequencing was carried out using the double-stranded dideoxy termination method (Sanger *et al.*, 1977). For that purpose, a T7 sequencing kit (Pharmacia) was used and sequencing of the clones was performed following the manufacturer's protocol.

E. Glutathione S-transferase (GST) fusion constructs

In order to be expressed as GST fusion proteins, the cDNA clones were recloned in to pGEX- 4T-3 plasmid (Pharmacia). The full length of C-4.3 was inserted in-frame into the Eco RI site of plasmid PGEX-4T-3. Similarly, the full length of clone C-2.4 was inserted in-frame into the Bam HI and Xho I sites of plasmid PGEX-4T-3. Additionally, a cDNA fragment of the N-terminal region of Sam68, corresponding to nucleotides 148-582, was inserted into the Bam HI and Eco RI sites of pGEX-4T-3. This latter cDNA was obtained by PCR amplification of a λgt11 cDNA library using internal primers designed from the

published sequence of Sam68 (Wong *et al.*, 1992). They were primer Sam68-1 consisted of 5'CGCTTTCTCGCTCCTTGGAT3' and primer Sam68-D, consisted of 5'GCACCAGTCTCTTCCTGCA3'. All constructs were then partially sequenced in the junction regions in order to verify the proper reading frame.

F. Expression and purification of GST-fusion proteins

XL1B or BL21(DE3)pLysS *E. coli* strains containing GST constructs were inoculated into 500 mL 2x TY medium, pH 7.3 (16 g tryptone, 10 g yeast extract, 5 g NaCl and 2.5 g disodium phosphate in 1 L). For constructs in XL1B, 100 μg/mL Ampicillin was added into the medium, whereas for constructs in BL21(DE3)pLysS, 100 μg/mL Ampicillin and 25 μg/mL Chloramphenicol were added. The cultures were incubated overnight at 30°C in a shaker incubator. Insoluble proteins were expressed upon induction with 0.2 mM Isopropyl-1-thio-β-D-galactopyranoside (IPTG) and incubation for 2 hour at 30°C. Soluble proteins were expressed upon 0.4 mM IPTG induction followed by 5 hour incubation at 30°C.

Purification of insoluble GST-fusion proteins was carried out following the protocol described by Frangioni and Neel (1993). The bacterial culture was centrifuged for 20 min at 3,000 rpm at 4°C. The pellet was washed once with ice cold STE buffer (10 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA) then redissolved in 5 mL of STE buffer containing 100 μg/mL Lysozyme. After 15 min incubation on ice, 25 mL STE containing 10 μg/mL Leupeptin, 5 μg/mL Aprotinin, 5 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF) and 1.5% sarkosyl were added and the suspension was vortexed for 5 seconds. The cells were then sonicated for 3x 1 min with 50% output control. 1% Triton-X

100 was then added into the suspension, followed by centrifugation at 15,000 rpm at 4°C for 20 min in an SS-34 rotor. The supernatant containing the GST-fusion protein was then mixed with 5 mL of 50% suspension of glutathione agarose beads (Sigma) in PBS. After overnight incubation at 4°C, the agarose was transferred into a column and washed with 100 mL cold PBS. The GST-fusion protein was eluted using 10 mM reduced glutathione (Sigma), 75 mM HEPES pH 7.4, 150 mM NaCl, 5 mM DTT and 2% N-octyl glucoside.

Soluble proteins were purified using a different method. After the cultures were centrifuged, the pellet was resuspended in 30 mL of 1x PBS containing 1 mM PMSF, 1 mM DTT and 2 mM EDTA. The cells were then sonicated for 3 x 30 sec with 50% output control. Following addition of 0.2% Triton-X 100, the cell suspension was centrifuged for 20 min at 4°C at 15,000 rpm in an SS-34 rotor. The supernatant was incubated with 5 mL of pretreated glutathione agarose overnight at 4°C. After transferring into a column, the protein was washed with 100 mL of cold PBS containing 2 mM EDTA. The GST-fusion proteins were eluted using 10 mM reduced glutathione and 2 mM EDTA in PBS.

G. Production of polyclonal antibodies

The production of polyclonal antibodies was carried out by immunizing male New Zealand white (NZW) rabbits with the GST-fusion proteins. For the first immunization, 500 μg protein mixed with Freund's complete adjuvant (Sigma) was given intra-muscularly, whereas, for the second and third immunizations, 250 μg protein was mixed with Freund's incomplete adjuvant and was given sub-cutaneously. The interval between immunizations was 3 weeks. To obtain the antibodies, blood was taken from the rabbit and the serum was

separated by centrifugation at 3,000 rpm at 4°C for 10 min.

H. Purification of polyclonal antibodies

Rabbit antibody anti GST-fusion protein was purified using GST-fusion protein bound to cyanogen bromide (CNBr)-activated sepharose 6MB (Sigma). To bind the GST-fusion protein to the sepharose, the protein was dialyzed against coupling buffer (0.5 M NaCl, 0.2 M $NaHCO_3$, pH 8.5) overnight at 4°C. The protein was then mixed with pretreated CNBr-activated sepharose for 2 hours at room temperature followed by overnight incubation at 4°C. Following centrifugation, the supernatant was taken and the sepharose was washed with coupling buffer. After incubation of the sepharose in 0.2 M glycine pH 8.5 at room temperature for 2 hours, the sepharose was washed 3x using coupling buffer then 3x using 0.1 M sodium acetate and 0.5 M NaCl, pH 4.0. To purify the antibody, the sepharose containing GST-fusion protein was incubated overnight at 4°C with rabbit antiserum. The sepharose was then washed with 100 mL PBS. The antibody was eluted using 100 mM glycine pH 2.8, and neutralized with 120 mM Tris pH 8.8.

I. Western blot analysis

To carry out the Western blotting analysis, purified proteins were resolved in 10% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The electrophoresis was performed on Bio-Rad minigel apparatus with a current of 20 mA/gel in a running buffer (2 L of 5x running buffer contains 30.25 g Tris, 142 g glycine and 10 g SDS). Proteins were then transferred onto a 0.45 μm nitrocellulose membrane (Bio-Rad)

using a Bio-Rad transfer apparatus in a transfer buffer (50 mL of 1 M Tris-HCl pH 8.3, 28.83 g glycine and 400 mL methanol in 2 L buffer) with a voltage of 50 V for 2 hours.

After proteins were transferred, membranes were incubated in blocking solution (1% BSA and 0.05% Tween-20 in TBS) for 1 hour at room temperature. As a control, a duplicate membrane was stained with 0.2% Ponceau S and 3% trichloro-acetic acid. A Polyclonal antibody against the C-terminal region of Sam68 was used at 0.1μg/mL concentration as the primary antibody. For the secondary antibody, 1:2,000 dilution of donkey anti-rabbit IgG conjugated with horse radish peroxidase was used. ECL (Amersham) was used to detect the reactivity of the antibody to the proteins.

J. Localization and co-localization experiments

To examine the localization and co-localization of c-Src and C-4.3, indirect immunofluorescence was carried out following methods described by Kaplan *et al.*, 1992. Cells were grown on a cover slip, fixed with 3% paraformaldehyde, and permeabilized with 0.2% Triton-X 100. As primary antibodies, anti-Src 2-17 monoclonal antibody (1:50 dilution in PBS containing 0.2% gelatin) and rabbit antiserum anti-C-4.3 (1:5 dilution) were used. The cells were incubated in the antibodies for 30 min at room temperature. After 3x washing with PBS, the cells were incubated in secondary antibodies, goat anti-mouse IgG labelled with rhodamine (1:100 dilution in PBS) and goat anti-rabbit IgG labelled with fluorescein (1:100 dilution in PBS), for 30 min at room temperature. Cells were then stained with 1 μg/mL bisbenzamide (Sigma) for 5 min (to stain the DNA) and washed 3x with PBS. The cover slip was then mounted on a slide using mowiol (Calbiochem) as the mounting medium

and was examined under a fluorescence microscope.

K. *In vitro* binding assay

A 100 ng purified baculovirus-expressed c-Src (provided by Dr. J. Bjorge) was used for the *in vitro* binding experiment. 10 μg GST-C-4.3 fusion protein and a control of 10 μg GST protein were bound to glutathione agarose beads. After incubation for 30 min at 4°C with purified baculovirus-expressed c-Src in binding buffer (25 mM Tris pH 7.5, 1 mM DTT, 1 mM EDTA, 0.2 mM PMSF, 50 mM NaCl), the beads were washed 3 times with washing buffer (0.1% Triton X-100 in binding buffer) and the proteins were resolved on 10% SDS-PAGE gel, and were transferred onto a nitrocellulose membrane (Bio-Rad). The membrane was stained with Ponceau S to visualize the transferred proteins. Subsequently, the membrane was immunoblotted with 0.1 μg/mL of the anti-Src-327 monoclonal antibody as described above in the western blot analyses section.

L. *In vitro* kinase assay

10 μL (10 μg) GST-C-4.3, GST-C-2.4 fusion protein and control GST were separately incubated for 30 min at 30°C with 10 μL (25 ng) purified baculovirus-expressed c-Src in 25 μL kinase buffer (100 mM HEPES, 10 mM $MgCl_2$, 30 mM NaCl, 2 mM DTT) and 30 μM cold ATP as well as 25 μCi [γ-^{32}P] ATP. The solution was then resolved on a 10% SDS-PAGE gel and the gel was dried using a gel drier. The dried gel was then exposed to an autoradiography film overnight.

M. Immunoprecipitation

To carry out the *in vivo* binding assay, NIH 3T3 fibroblasts were grown on a 100 mm tissue culture dish. Cells were washed once with 5 mL of cold PBS and lysed with 600 μL of RIPA buffer (25 mM Tris pH 7.2, 0.1% SDS, 1% Triton X-100, 1% Nadeoxycholate, 150 mM NaCl, 1 mM EDTA and 100 μg/mL Leupeptin). Following incubation on ice for 10 min, the lysates was then centrifuged at 10,000 rpm at 4°C for 15 min. Two hundred μg of supernatant (quantitated using a Bio-Rad protein assay) was mixed with 2 μL (100 μg/mL) of anti-Src 327 and the mixture was incubated on ice for 1 hour. After addition of 2 μL of rabbit-anti mouse IgG, the lysates were incubated on ice for 30 min, and 20 μL of protein A sepharose was added. Following incubation on a rotator at 4°C for 30 min, the lysate was centrifuged for 5 min at 4°C. The pellet was dissolved in SDS sample buffer and the immunoprecipitates were resolved on a SDS gel and transferred onto nitrocellulose membrane. Western blotting was performed using anti-C-4.3 polyclonal antibody. Detection of the positive signal was done using an ECL kit.

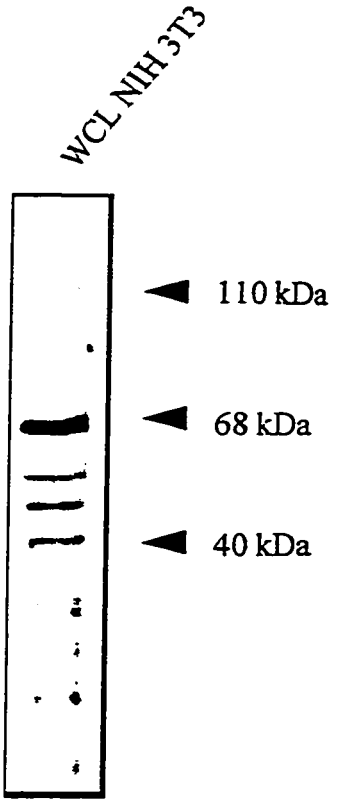
CHAPTER THREE

RESULTS

A. Polyclonal antibodies raised against specific regions of Sam68 reacted with several proteins in Western blots from NIH 3T3 fibroblasts.

Our laboratory has previously demonstrated that a monoclonal antibody raised against an N-terminal region of Sam68 (formerly known as GAP-associated p62) recognized several protein bands in Western blots of various cell types. The size differences between Sam68 and these protein bands, in some cases, were too large to be characterized as phosphorylated forms of Sam68. Additionally, a polyclonal antibody raised against the same region of Sam68 recognized multiple bands in Western blot analyses of various cell types (unpublished data). These results indicated that Sam68 might be expressed as different isoforms in other cell types and that it might be a member of a family of related proteins. A polyclonal antibody raised against a peptide corresponding to amino acids 331-443 of the Sam68 (Santa Cruz) was used to immunoblot lysates of NIH 3T3 fibroblasts. Five bands with molecular mass ranging between 40 kDa and 110 kDa were observed (Figure 4). The major band, exhibiting a molecular mass of 68 kDa, was apparently the Sam68 protein. Together, these data support the hypothesis that Sam68 might have antigenically related proteins.

In an attempt to isolate novel Sam68-related proteins and possible novel substrates of Src tyrosine kinase, two polyclonal antibodies raised against both N- and C-terminal regions of Sam68 (Santa Cruz) were used to screen a human T-lymphocyte λgt11 cDNA



Blot with anti-C-Sam68

Figure 4. Anti-C-Sam68 polyclonal antibody reacted with several proteins in a Western blot from NIH 3T3 fibroblasts. Whole cell lysates of NIH 3T3 fibroblasts were resolved on a 10% SDS-PAGE gel and transferred onto a nitrocellulose membrane. Upon immunoblotting the membrane with an anti-C-terminal specific Sam68 polyclonal antibody, several bands with molecular masses ranging between 40 kDa and 110 kDa were identified. The major band with a molecular mass of 68 kDa was apparently the Sam68 protein.

expression library. The main focus of this project concentrated on screening a library using the polyclonal antibody raised against the C-terminal region of Sam68, because the polyclonal antibody raised against the N-terminal region of Sam68 was no longer commercially available and an attempt to reproduce the antibody was unsuccessful. In addition, the isolated cDNA clones obtained during the screening process with the N-terminal Sam68 antibody only represented partial cDNAs, and the library used during this screening process did not appear to content full length cDNAs of these clones. As a result, a polyclonal antibody raised against the C-terminal region of Sam68 was used instead.

B. Library screening using an antiserum raised against an N-terminal region of Sam68 resulted in the isolation of clones that have homology to RNA binding proteins.

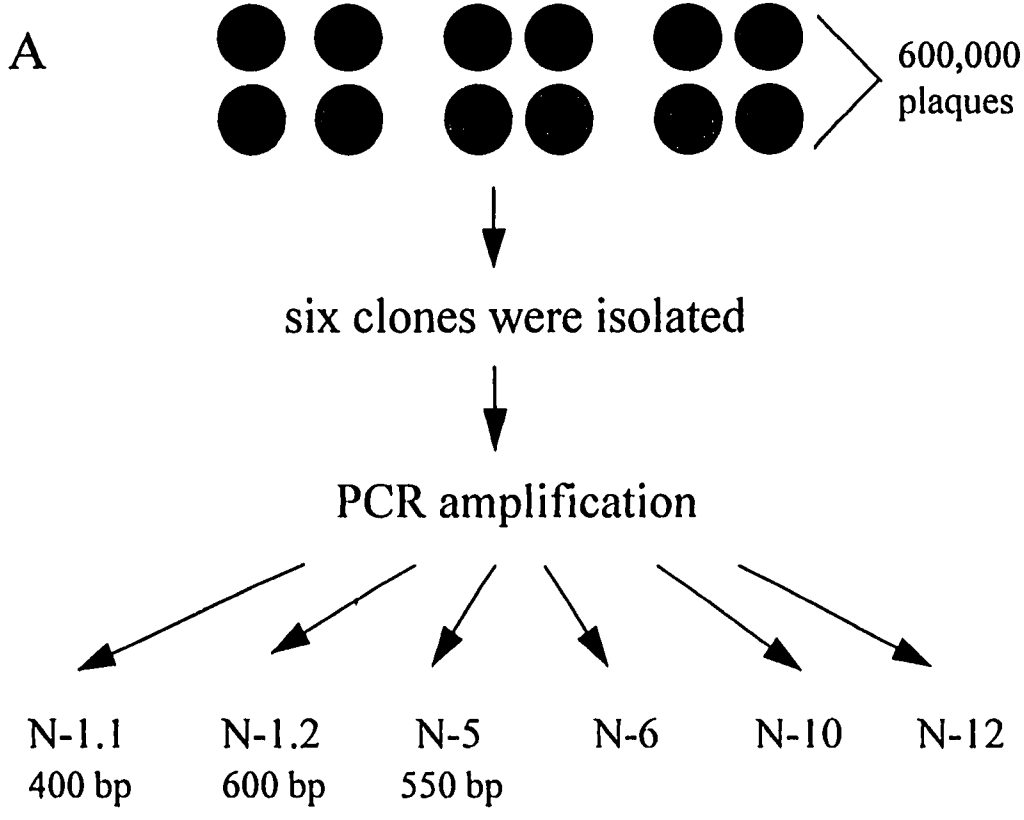
A polyclonal antibody raised against an N-terminal region of Sam68 corresponding to amino acids 103-281 (Santa Cruz) was used to screen 7.5 X 10^5 plaques from a human T-lymphocytes λgt11 cDNA expression library. Six positive plaques were isolated during the first screening. Secondary and tertiary screenings were performed in order to purify the positive plaques. Finally, the remaining positive six λ phage clones were isolated from the agar plate. They were designated as N-1.1, N-1.2, N-5. N-6, N-10 and N-12 (Figure 5).

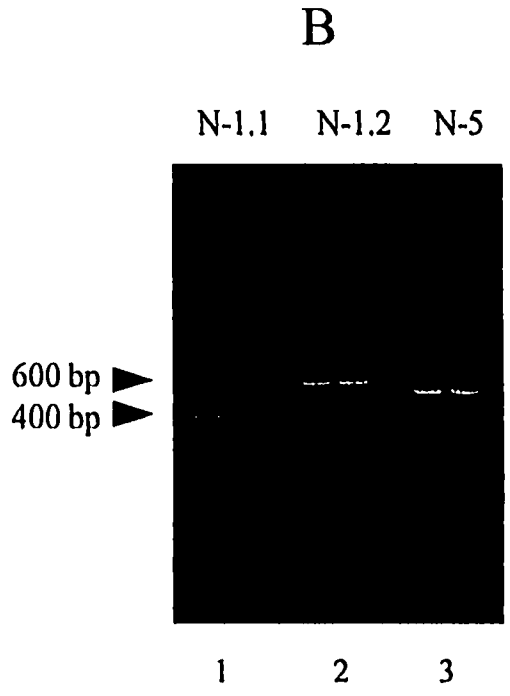
B.1. Cloning and sequence analysis of N-1.1, N-1.2, and N-5

Molecular cloning of the isolated clones was achieved through PCR-amplification techniques, as well as restriction endonuclease digestions. PCR-amplification of the cDNA

Figure 5. Schematic representation of the screening process of a T-lymphocyte cDNA expression library using an anti-N-terminal specific Sam68 antibody. Approximately 600,000 λgt11 pfu's were plated. After tertiary screening six purified phage clones were isolated (N-1.1, N-1.2, N-5, N-6, N-10 and N-12) (A). PCR amplification of N-1.1, N-1.2 and N-5 using λgt11 reverse and forward primers resulted in identification of DNA products with a molecular size of approximately 400 bp, 600 bp and 550 bp respectively (B).

Screening of a T-lymphocytes cDNA expression library using an anti-N-terminal specific Sam68 antibody





inserts was carried out using λgtl1-forward and reverse primers. Subsequently, the PCR products were digested with Eco RI restriction enzyme and were subcloned into pBluescript vectors as described in materials and methods. The approximate molecular sizes of the cDNA clones N-1.1, N-1.2, N-5 were 400 bp, 600 bp, and 550 bp respectively (Figure 5B). These clones were then sequenced using the double-stranded dideoxy chain termination method according to Sanger *et al.* (1977).

Clone N-5 contained 534 nucleotides that encoded 167 amino acids (Figure 6). A start codon, ATG, was found at nucleotide positions 137-139. However, no stop codon was found in the C-terminal region of this sequence, indicating that this cDNA clone did not represent a full length transcript. Interestingly, analysis of the N-5 amino acid sequence using PC GENE revealed some sequence similarity between N-5 and the KH domain of Sam68 (Figure 7B). Since this domain is found in the N-terminal region of Sam68, it is possible that these amino acid residues within the KH domain form an epitope recognizeable by the Sam68 antibody. N-5 also exhibited a high degree of homology to human ribosomal protein S5 with 66.10% identity at the amino acid level (Figure 7A). Furthermore, it was found that other RNA binding proteins, such as human hnRNP K; fragile-X mental retardation gene product, FMR-1, and *Halobacterium halobium* S3 ribosomal protein, Hh S3, shared the same conserved amino acid residues region with N-5 (Figure 7B). Another equally interesting finding was that some other cDNA clones that were partially sequenced also exhibited some degree of homology to RNA binding proteins. N-1.1 exhibited 62% identity at the nucleotide level to the human small nuclear ribo-nucleoprotein gene, SNRP E, in its first 200 nucleotides, whereas, N-1.2 exhibited 51,6% identity to the human X-linked nuclear protein

Figure 6. Sequence analysis of clone N-5. The sequence of N-5 contained 534 nucleotides, which encoded 167 amino acids. A possible start codon at nucleotide position 137-139 was indicated by box. No stop codon was found in the C-terminal region of this sequence, indicating that this cDNA clone did not represent a full length transcript. The conserved residues in KH domain were underlined.

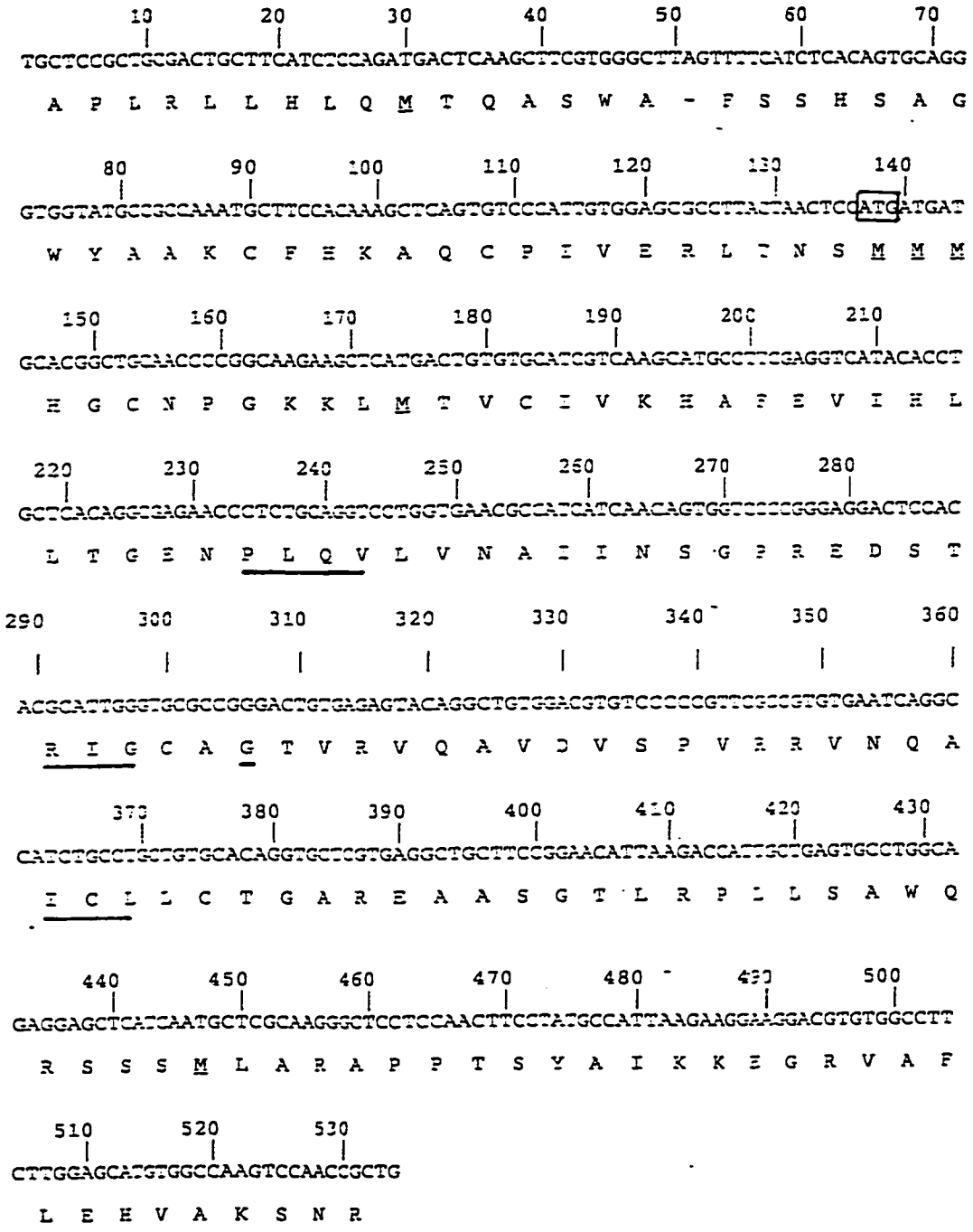
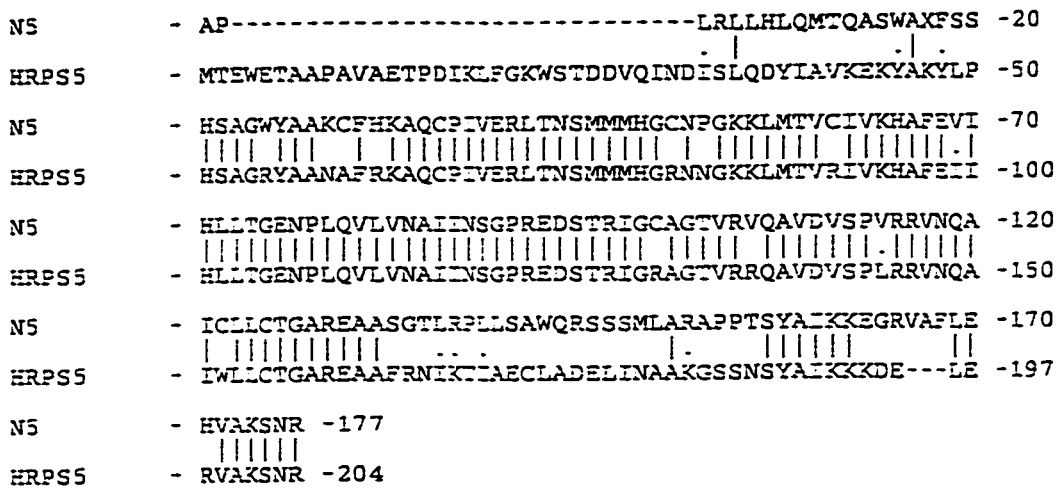
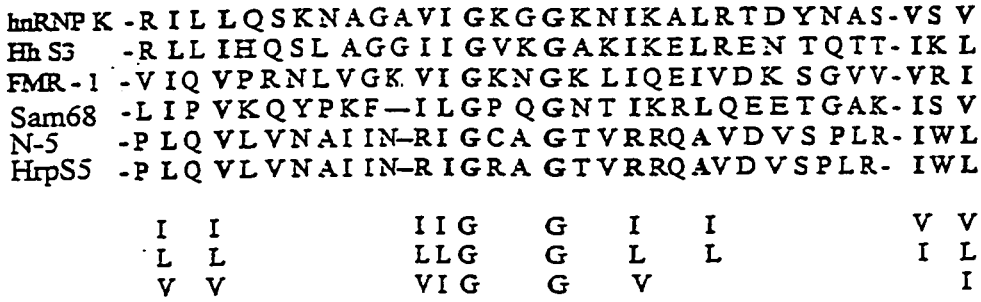

10 20 30 40 50 60 70
TGCTCCGCTGCGACTGCTTCATCTCCAGATGACTCAAGCTTCGTGGGCTTAGTTTTCATCTCACAGTGCAGG
A P L R L L H L Q M T Q A S W A - F S S H S A G
80 90 100 110 120 130 140
GTGGTATGCCGCCAAATGCTTCCACAAAGCTCAGTGTCCCATTGTGGAGCGCCTTACTAACTCCATGATGAT
W Y A A K C F H K A Q C P I V E R L T N S M M M
150 160 170 180 190 200 210
GCACGGCTGCAACCCCGGCAAGAAGCTCATGACTGTGTGCATCGTCAAGCATGCCTTCGAGGTCATACACCT
H G C N P G K K L M T V C I V K H A F E V I H L
220 230 240 250 260 270 280
GCTCACAGGTGAGAACCCTCTGCAGGTCCTGGTGAACGCCATCATCAACAGTGGTCCCCGGGAGGACTCCAC
L T G E N P L Q V L V N A I I N S G P R E D S T
290 300 310 320 330 340 350 360
ACGCATTGGGTGCGCCGGGACTGTGAGAGTACAGGCTGTGGACGTGTCCCCCGTTCGCCGTGTGAATCAGGC
R I G C A G T V R V Q A V D V S P V R R V N Q A
370 380 390 400 410 420 430
CATCTGCCTGCTGTGCACAGGTGCTCGTGAGGCTGCTTCCGGAACATTAAGACCATTGCTGAGTGCCTGGCA
I C L L C T G A R E A A S G T L R P L L S A W Q
440 450 460 470 480 490 500
GAGGAGCTCATCAATGCTCGCAAGGGCTCCTCCAACTTCCTATGCCATTAAGAAGGAAGGACGTGTGGCCTT
R S S S M L A R A P P T S Y A I K K E G R V A F
510 520 530
CTTGGAGCATGTGGCCAAGTCCAACCGCTG
L E H V A K S N R

Figure 7. Alignments between N-5, human ribosomal protein S5 and several RNA binding proteins. N-5 exhibited a high degree of homology to human ribosomal protein S5 with 66.10% identity at the amino acid level (A). Some conserved amino acid residues within the KH region were found in N-5 and several RNA binding proteins, such as hnRNP K, fragile-X mental retardation gene, FMR-1, and *Halobacterium halobium* S3 ribosomal protein (B).

A


N5 - AP----------LRLLHLQMTQASWAXFSS -20
HRPS5 - MTEWETAAPAVAETPDIKLFGKWSTDDVQINDISLQDYIAVKEKYAKYLP -50
N5 - HSAGWYAAKCFHKAQCPIVERLTNSMMMHGCNPGKKLMTVCIVKHAFEVI -70
HRPS5 - HSAGRYAANAFRKAQCPIVERLTNSMMMHGRNNGKKLMTVRIVKHAFEII -100
N5 - HLLTGENPLQVLVNAIINSGPREDSTRIGCAGTVRVQAVDVSPVRRVNQA -120
HRPS5 - HLLTGENPLQVLVNAIINSGPREDSTRIGRAGTVRRQAVDVSPLRRVNQA -150
N5 - ICLLCTGAREAASGTLRPLLSAWQRSSSMLARAPPTSYAIKKEGRVAPLE -170
HRPS5 - IWLLCTGAREAAFRNIKTIAECLADELINAAKGSSNSYAIKKKDE---LE -197
N5 - HVAKSNR -177
HRPS5 - RVAKSNR -204

B


hnRNP K -R I L L Q S K N A G A V I G K G G K N I K A L R T D Y N A S - V S V
Hn S3 -R L L I H Q S L A G G I I G V K G A K I K E L R E N T Q T T - I K L
FMR-1 -V I Q V P R N L V G K V I G K N G K L I Q E I V D K S G V V - V R I
Sam68 -L I P V K Q Y P K F — I L G P Q G N T I K R L Q E E T G A K - I S V
N-5 -P L Q V L V N A I I N — R I G C A G T V R R Q A V D V S P L R - I W L
HrpS5 -P L Q V L V N A I I N — R I G R A G T V R R Q A V D V S P L R - I W L
I I I I G G I I V V
L L L L G G L L I L
V V V I G G V I

(XNP) gene over a region of around 500 nucleotides.

B.2. Further cDNA library screening in order to obtain the full transcript of clone N-5.

In order to obtain the full length cDNA of N-5, Southern blot-DNA hybridization method was utilized to screen the same human T–lymphocytes λgt11 cDNA expression library. The cDNA labelling of N-5 was carried out using random hexamer priming. After performing primary screening on a population of 1.2 X 10^6 plaques and proceeding to subsequent secondary and tertiary screenings, 14 clones of λgt11 phages were isolated. The molecular size of the inserts was revealed by PCR-amplification of the λgt11 phage clones using λgt11 forward and reverse primers. However, the PCR products of all fourteen isolated clones exhibited the same molecular size. Four clones were randomly selected and subcloned into pBluescript plasmid for DNA sequencing. The sequencing results revealed that all four clones contained the identical sequence as N-5. These data indicated that the λgt11 cDNA library did not contain the full N-5 transcript. The cDNA expression library was over ten years old and the phage titre had decreased by three logs. It is possible that the library may have been amplified too many times. As a result, the project was terminated.

C. Library screening using an antiserum raised against a C-terminal region of Sam68 resulted in the isolation of a clone, C-4.3, that exhibits proline-rich regions and a coiled-coil structure.

A polyclonal antibody raised against the C-terminal region of Sam68 corresponding

to amino acids 331-443 (Santa Cruz) was used to screen a population of 1.2 X 10^6 λgt11 plaques from a human T-lymphocytes cDNA expression library. Nineteen positive plaques were isolated. However, after secondary and tertiary screening only eleven positive plaques remained (Figure 8). The eleven phage clones were C-1.3, C-2.4, C-2.5, C-3.1, C-3.2, C-3.3, C-4.1, C-4.2, C-4.3, C-5.1, and C-5.2 (Figure 9A).

C.1. Southern blot hybridization analysis of the isolated clones

To address whether the eleven isolated phage clones were related to each other on a molecular basis, they were subjected to Southern blot hybridization analysis. The cDNA inserts were PCR-amplified using λgt11 forward and reverse primers. After resolution by agarose gel electrophoresis, the DNA bands were transferred onto a nylon membrane. Two cDNA fragments, C-2.4 and C-4.1, were radioactively labelled as described in material and methods and used as probes. In addition, a cDNA fragment generated by PCR-amplification of the C-terminal region of Sam68 corresponding to nucleotides 691-1431 was also radioactively labelled and used as a probe.

The probe synthesized from C-2.4 cDNA hybridized to two clones, C-2.4 itself and C-2.5, suggesting that the C-2.4 and C-2.5 clones were related (Figure 9B). Similarly, the C-4.1 DNA probe hybridized to seven clones including C-4.1 itself, suggesting that the seven clones were related and could be categorized into one group. They were C-1.3, C-3.1, C-3.2, C-3.3, C-4.1, C-4.2, C-4.3, and C-5.2 (Figure 9C). However, no positive hybridization signal other than the positive control was observed when the probe from the C-terminal cDNA of Sam68 was used (Figure 9D). These data indicated that there were no Sam68 cDNAs isolated

Figure 8. Schematic representation of the screening process of a T-lymphocyte cDNA expression library using an anti-C-terminal region specific Sam68 antibody. Approximately 1,250,000 λgt11 pfu's were plated. Twenty one plaques showed a positive signal in the primary screening. However, after tertiary screening, only eleven purified phage clones that reacted positively were isolated.

Screening of a T-lymphocytes cDNA expression library using an anti-C-terminal specific Sam68 antibody

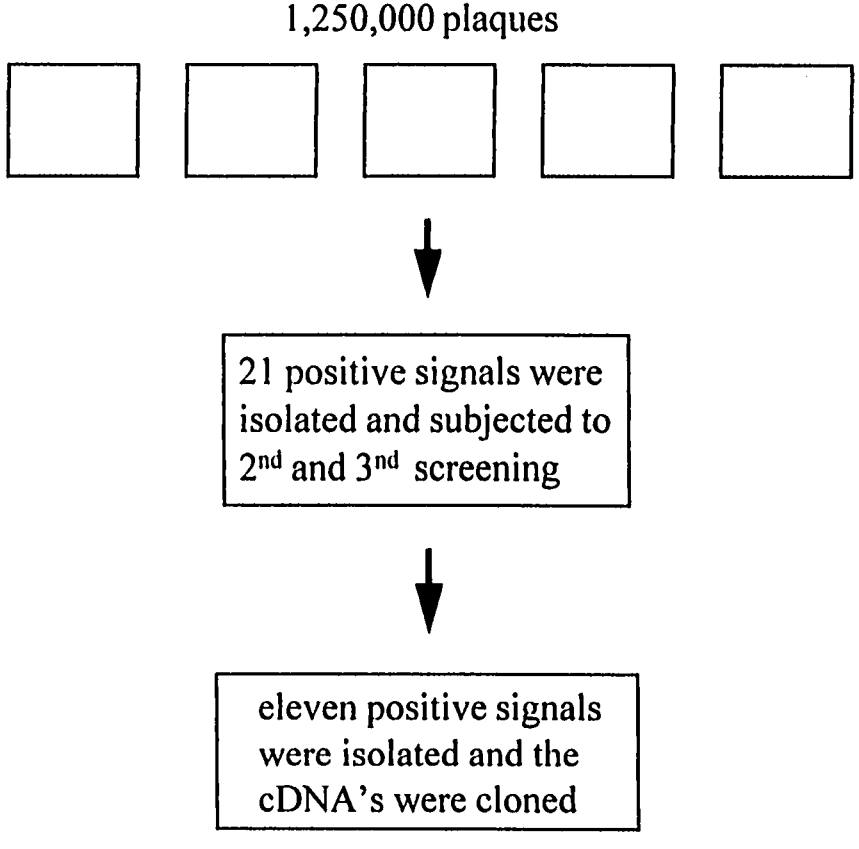
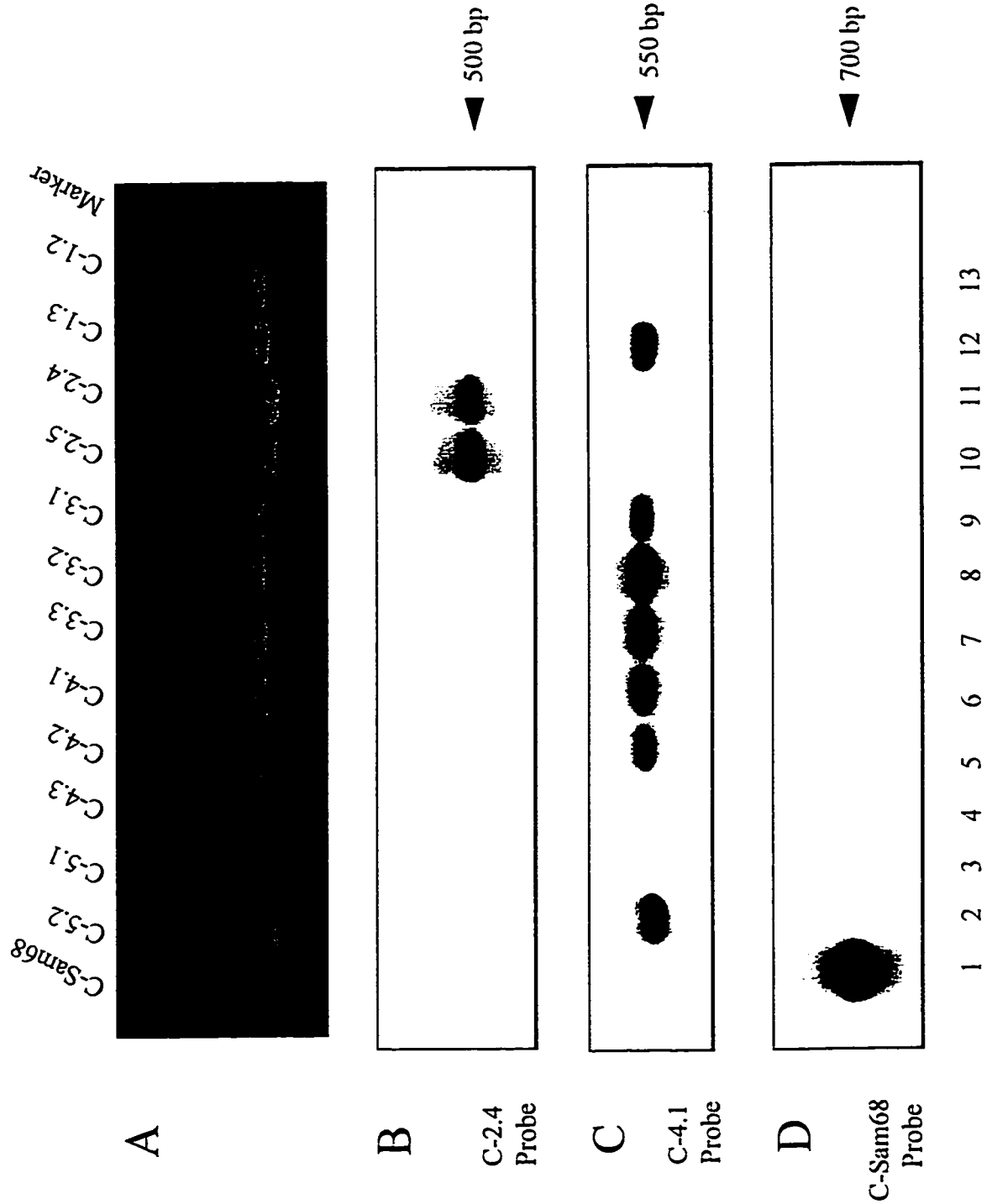


Figure 9. Southern blot analyses of the cDNA clones isolated by screening of a T-lymphocyte cDNA expression library using an anti-C-terminal specific Sam68 antibody. DNAs obtained from PCR amplification of the inserts of the isolated clones were resolved on a 1% agarose gel (A) and transferred onto a nylon membrane. Hybridization with radioactively labelled C-2.4 probes resulted in the detection of C-2.5 and C-2.4 following autoradiography (B). Hybridization using radioactively labelled C-4.1 probes resulted in the detection of C-1.3, C-3.1, C-3.2, C-3.3, C-4.1, C-4.2 and C-5.2 following autoradiography (C). Hybridization using radioactively labelled C-Sam68 probes resulted in the detection of control C-Sam68 only (D). Removal of probes was carried out using moderate stripping (as described in material and methods).


A
Marker
C-1.2
C-1.3
C-2.4
C-2.5
C-3.1
C-3.2
C-3.3
C-4.1
C-4.2
C-4.3
C-5.1
C-5.2
C-Sam68
B
C-2.4 Probe
500 bp
C
C-4.1 Probe
550 bp
D
C-Sam68 Probe
700 bp
1 2 3 4 5 6 7 8 9 10 11 12 13

during the screening process. As a result, an attempt was made to re-screen the original nineteen isolated phages obtained during the primary screening. However, no positive signals were observed.

C.2. Cloning, subcloning and sequencing of C-4.3, C-2.4 and C-4.1

The cDNA inserts were PCR-amplified using λgt11forward and reverse primers and then subcloned into a TA-2.1 cloning vector. They were also subcloned into pBluescript plasmid, which contains more restriction sites, for subsequent manipulation. For the purposes of sequencing, clone C-4.3 which had a molecular size of 1.9 kb was subjected to nested deletion (see materials and methods). This clone was fragmented into twelve pieces with intervals of 150-200 bp (Figure 10) and subcloned back into the TA-2.1 plasmid. Sequencing of the first ~200 nucleotides of each clone indicates that they consist of three different groups of cDNA clones. As a result, only three clones, C-4.3, C-2.4 and C-4.1, which each represents a group of the cDNA clones were selected for complete sequencing. Internal primers of these three clones were then synthesized for complementary sequencing.

C. 3. Structural and sequence analyses of C-4.3

The complete sequence of the cDNA and deduced amino acid sequence of C-4.3 are shown in figure 11A. The nucleotide sequence of C-4.3 was comprised of 1901 bp and contained an open reading frame (ORF) of 1380 bp. The ORF spanned an in-frame ATG initiation codon at nucleotide position 130 and a TAA termination codon at position 1510. This ORF encoded a putative polypeptide of 460 amino acids with a predicted molecular

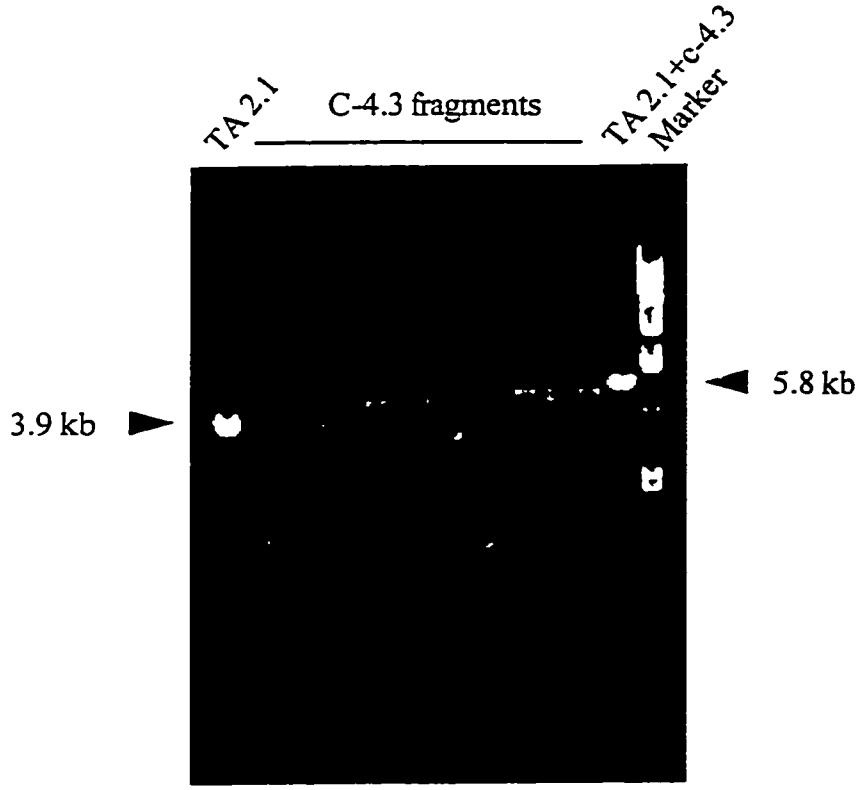


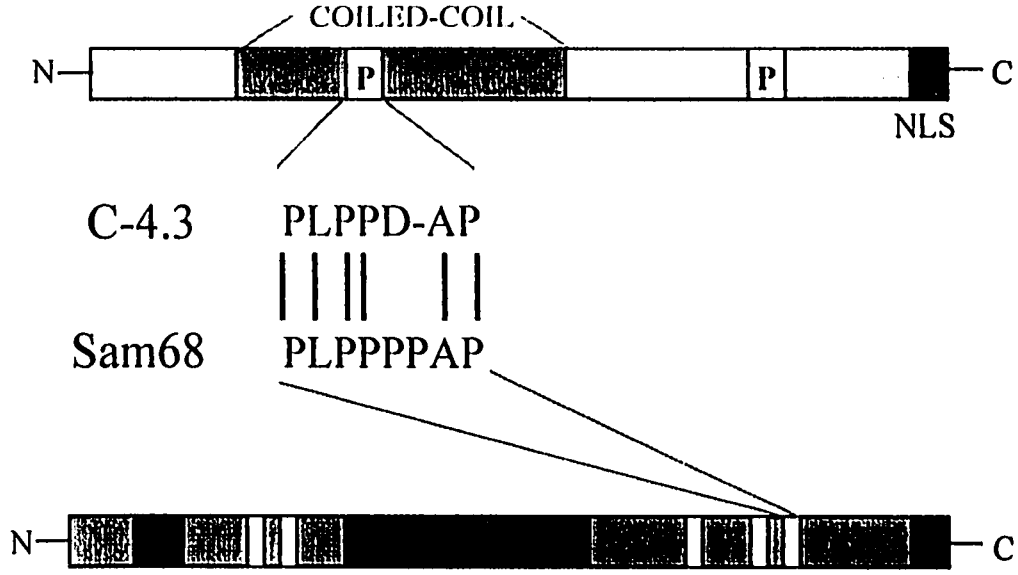
Figure 10. Nested deletion products of the cDNA clone C-4.3. For the purpose of DNA sequencing, clone C-4.3, which had a molecular size of 1.9 kb, was subjected to a nested deletion process. The DNA inserts were fragmented into 12 pieces with respective differences of 150-200 bp, as shown by 1% agarose gel resolution (lane 1-12). Plasmid TA-2.1 had a molecular size of 3.9 kb.

Figure 11. Schematic representation of amino acid sequence and structural analysis of clone C-4.3. The C-4.3 protein consisted of 460 amino acids residues. Amino acids from positions 7-11 contained a possible phosphorylation motif (boxed); amino acids from positions 70-260 contained coiled-coil region (underlined); amino acids from positions 143-149 and 370-392 contained proline-rich regions (bolded) and amino acids from positions 457-460 contained a potential nuclear localization signal (NLS) (boxed) (A). C-4.3 displayed sequence homology to Sam68 in a proline rich region at positions 143-149 (B). Amino acid residues 70 to 260 of C-4.3 had a high probability of forming a coiled-coil structure. This analysis was performed using Macstripe 2.0a1 software (C). The proline rich region at positions 143 to 149 disrupts this structure.

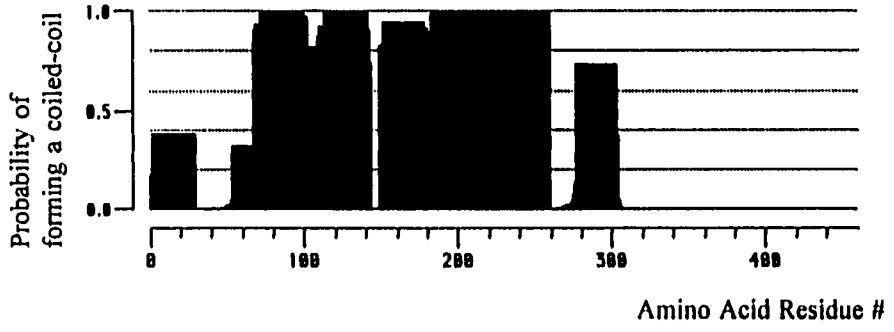
A

1 MKHSLRYFEEESKDLAGRLQRSSQRI
27 GELEWSLCAVAATGKKKPDGFSSRS
52 KALLKRQLEQSIREQILLKGHVTQLK
78 ESLKEVQLERDGYAEQIKGERAQWQ
103 QRMRKMSQEVCTLKEEKKHDTHRV
127 EELERSLSRLKNQMAE**PLPPDAP**AVS
153 SEVELQDLRKELERVAGELQAQVEN
178 NQCISLLNRGQKERLREQEERLQEQQ
204 ERLREQEKRLQQLAEPQSDLEELKHE
230 NKSALGLEQQVKELQEKLGQVTETL
266 TSAEKEPEAAVPASGTGGESSGLMDL
281 LEEKADLREHVEKLELGFIQYRRERC
307 HQNVHRLLTEPGDSAKDAAPGGGHH
332 QAGPGRGGEEGEAVGAAGNGVAAC
356 GSYSEGHGKFLAAARN**PAAEP**SPGA
383 **PAP**QELGAADKHGDLCEASLTNSVEP
408 AQGEAREGSSQDNPTAQPVVQLLGE
432 MQDHQEHPGLGSNCCVPCFCWAWL
456 PRRRRR

B



C



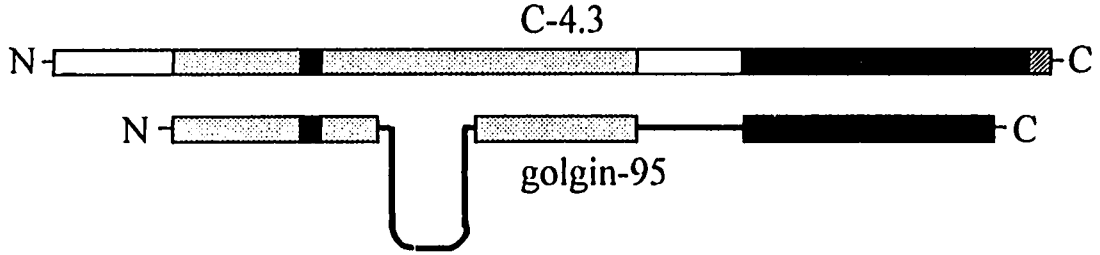
weight of 57,329 Da. The amino acid composition of this putative polypeptide is rich in glutamate (14.8%), leucine (12.3%) and glutamine (9.5%). Using PC GENE alignment analysis between the C-4.3 and Sam68 amino acid sequences, no extensive homology was found. However, a proline-rich region of C-4.3 was found to be homologous to a proline-rich region present in the C-terminal region of Sam68 (75% identity) (Figure 11B). Based on the fact that anti-Sam68 polyclonal antibody was raised against the C-terminal region of Sam68, it is possible that this proline-rich region forms an epitope recognized by the antibody.

A hydrophilic plot of the C-4.3 sequence showed that there were no stretches of hydrophobic and neutral amino acids sufficient in length to span the lipid bilayer. Strikingly, a large part of the protein, spanning from amino acid position 70 to amino acid position 260, had a high probability of assuming a coiled-coil structure (Figure 11C). This structure was stabilized by the heptad repeats of leucine known as a leucine zipper motif (Landschultz *et al.*, 1988). A proline-rich region present at positions 143-149, PLLPDAP, interrupted this structure. Another proline-rich region, PAAPPSPGAPAP, was found within the C-terminal region of this protein at positions 370-392 . Additionally, a nuclear localization signal (NLS) consisted of RRRR (Jans and Hübner, 1996) was found at the C-terminus, at positions 457-460. Finally, a consensus motif, YFEE, resembling the substrate of epidermal growth factor (EGF) (Songyang *et al.*, 1995) and a consensus motif SPG, which is a proline-directed serine phosphorylation site, resembling the substrate of Erk1 (Songyang *et al.*, 1996) were also found in C-4.3.

A BLASTX search for amino acid sequence homology revealed that C-4.3 possesses a high degree of homology to golgin-95 (Figure 12) (Fritzler *et al.*, 1993), and a cis-Golgi

Figure 12. Schematic representation of sequence homology between C-4.3 and golgin-95. Sequences within shaded boxes represent the homologous regions, whereas sequences drawn with thick lines represent the non homologous regions (A). In golgin-95, an insertion of 84 amino acid residues interupts the coiled-coil domain. The amino acid identity and similarity between C-4.3 and golgin -95 are depicted in panel B.

A



B

C-4.3 ESLKEVQL ERDQYAEQI KGERAQWQQ RMRKMSQE VCTLKEEKKHDTHRVEELERSLSR LK
gol-95 ESVRQLQMERDKYAENLKGESAMWRQ RMQQMSE QVHTLREEKECSM SRVQELETSLAELR

C-4.3 NQMAEPLPPDAPAVSSEVELQD LRKELERVAGELQAQVEN NQCI SLLNR
gol-95 NQMAEPPPPEPPAGP SEVEQQLQAEAEHLRKELEGLAGQLQAQVQDNEGLSRLNR

C-4.3 QVTETLT SAEKEPEAAVPASGTGGESSGLMDLLE EKADLREHVEKLELGFIQYRRE
gol-95 RLAHLLASAQKEPEAAAPAPGTGGDSSRFMELMQEKADLKERVEELEHRCIQLSG E

C-4.3 S EG HGKFLAAARNPAAEPSP GAPAPQELGAADKHGDLCEASLTNSVEPAQGEAREGSSQ DNP
gol-95 NEWHGRFLAAAQNPADEPTS GAPAPQELGAANQQGDLCEVSLAGSVEPAQGEAREGSPR DNP

C-4.3 TAQP V VPQLLGEMQDHQEHPGLGSNCCVPCF
gol-95 TAQQ IMQLLR EMQN PRERPGLGSN PCI PFF

Matrix Protein GM130 (Nakamura *et al.*, 1997). Golgin-95 is a Golgi protein identified using an antiserum from a patient with Sjörgren's syndrome (Fritzler *et al.*, 1993), whereas GM130 is a protein reported to be involved in the docking of transport vesicles (Nakamura *et al.*, 1997). The homology between these proteins was found in several regions especially in the coiled-coil motif region, in which C-4.3 exhibited 59% identity and 81% similarity to golgin-95, and 53% identity and 75% similarity to GM130. In comparison to the coiled-coil motif in C-4.3, the coiled-coil motif of golgin-95 contained an insertion of 84 amino acids, whereas that of GM130 contained a long N-terminal addition. The high degree of similarity between these proteins could be indicative of alternative splicing or initiation of translation from different start sites. No other proteins were found to be as highly homologous to C-4.3 as were golgin-95 and GM130. Nevertheless, several proteins with predicted coiled-coil domain known to be structural and motor proteins were found to be weakly similar to C-4.3 (~20% identity and ~40% similarity). They were myosins, kinesin, plectin, dynein and a centromeric motor protein CENP-F.

C. 4. Structural and sequence analyses of clones C-2.4 and C-4.1

Amino acid sequence analysis of clones C-2.4 and C-4.1 revealed no stop codons, indicating these clones did not represent the full transcript cDNA. Clone C-2.4 was comprised of 485 nucleotides which encoded 161 amino acids (Figure 13). Interestingly, the protein sequence contained many NLSs scattered along its sequence. PC GENE analysis using the SOAP program showed that this protein was highly hydrophilic, with a hydrophatic index of -21.14 (Figure 14B). Homology searches using BLASTX revealed that C-2.4 was

Figure 13. Sequence analysis of clone C-2.4. Clone C-2.4 consisted of 485 nucleotides which encoded 161 amino acids. Several nuclear localization signals (NLSs) were found scattered along its sequence (underlined). No stop codon was found in the sequence of clone C-2.4, indicating this clone did not represent the full cDNA transcript.

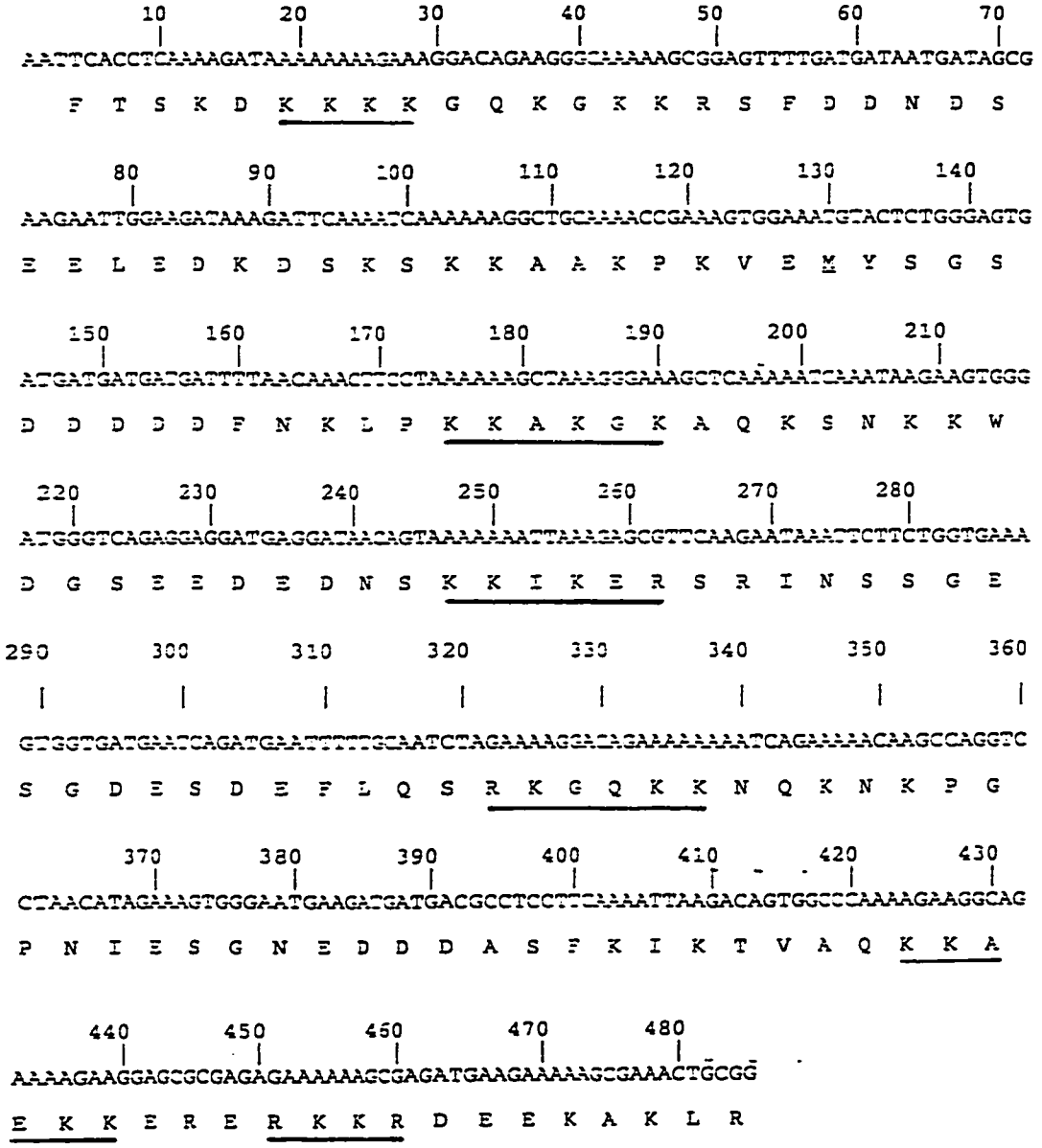
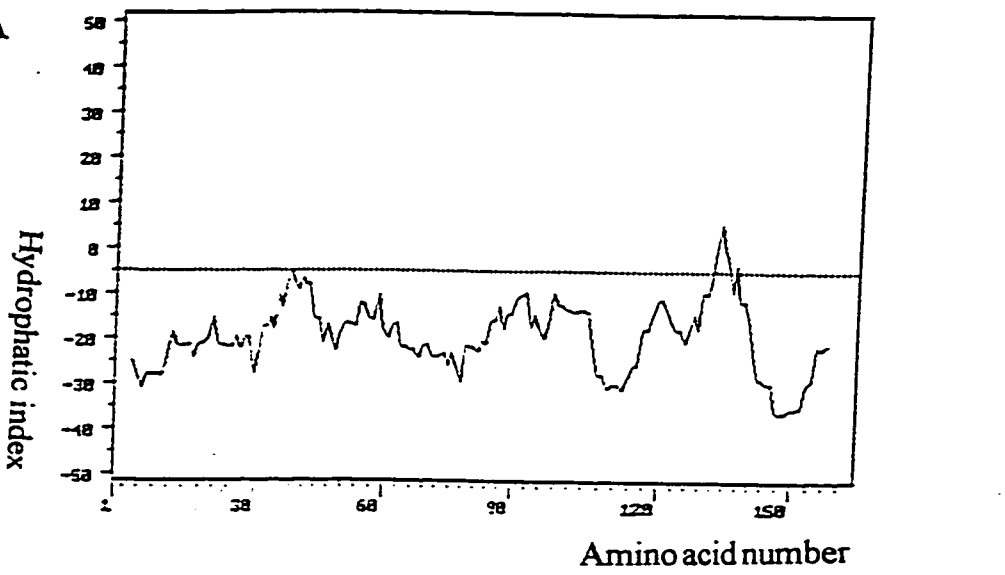

10 20 30 40 50 60 70
AATTCACCTCAAAAGATAAAAAAAAAGAAGGACAGAAGGGCAAAAAGCGGAGTTTTGATGATAATGATAGCG
F T S K D K K K K G Q K G K K R S F D D N D S
80 90 100 110 120 130 140
AAGAATTGGAAGATAAAGATTCAAAATCAAAAAAAGGCTGCAAAACCGAAAGTGGAAATGTACTCTGGGAGTG
E E L E D K D S K S K K A A K P K V E M Y S G S
150 160 170 180 190 200 210
ATGATGATGATGATTTTAACAAACTTCCTAAAAAAAGCTAAAGGGAAAGCTCAAAAATCAAATAAGAAGTGGG
D D D D D F N K L P K K A K G K A Q K S N K K W
220 230 240 250 260 270 280
ATGGGTCAGAGGAGGATGAGGATAACAGTAAAAAAAATTAAAGAGCGTTCAAGAATAAATCTTCTGGTGAAA
D G S E E D E D N S K K I K E R S R I N S S G E
290 300 310 320 330 340 350 360
GTGGTGATGAATCAGATGAATTTTTGCAATCTAGAAAAGGACAGAAAAAAAATCAGAAAACAAGCCAGGTC
S G D E S D E F L Q S R K G Q K K N Q K N K P G
370 380 390 400 410 420 430
CTAACATAGAAAGTGGGAATGAAGATGATGACGCCTCCTTCAAAATTAAGACAGTGGCCCAAAAGAAGGCAG
P N I E S G N E D D D A S F K I K T V A Q K K A
440 450 460 470 480
AAAAGAAGGAGCGCGAGAGAAAAAAGCGAGATGAAGAAAAAGCGAAACTGCGG
E K K E R E R K K R D E E K A K L R

Figure 14. PC GENE alignment between C-2.4 and the sequence of an annexin V binding protein and a hydropathy analysis of C-2.4 sequence. C-2.4 exhibited a very hydrophatic profile (A). C-2.4 exhibited a high degree of homology to annexin V binding protein, ABP7, with 76.40% identity (B). Sequence alignment between C-2.4 and the C-terminal region of Sam68 (S331), from which the anti-C-terminal specific Sam68 antibody was generated, did not show a significant homology (C).

A



B

```
ABP7      -          GQKGKKTSFDENDSEELEDKDSKSKKPAPRPNSEVLLSGSED -42
C24       - TSKDKKKKGQKGKKRSFDDNDSEELEDKDSKSKKAAKPKVEM--YSGSDD -48
ABP7      - ADDPNKLSKKGKK-AQKSTKKRDGSEEDEDNSKRSKERSRVNSSGESGGE -91
C24       - DDDFNKLPKKAKGKAQKSNKKWDGSEEDEDNSKKIKERSRINSSGESGDE -98
ABP7      - SDEFLQSRKGQKKNQKNKSVPTIDSGNEDDDSSFKIKTVAQKKAEKKERE -141
C24       - SDEFLQSRKGQKKNQKNKPGPNIESGNEDDDASFKIKTVAQKKAEKKERE -148
ABP7      - RKKREEEPAKLRKVKEKEELEKGRKEQSKQREPQKRPDEEVLVLRGTPDA -191
C24       - RKKRDEEPAKLR---------------------------------------- -160
ABP7      - GAASEENGDIAATLEDDNEGDKKKKDKKKKKTEKDDKEKEKKKG -235
C24       - ----------------------------------------E -161
```

C

```
C24       -   TSKDKKKKGQKGKKRSFDDNDSEELEDKDSKSKKAAKPKVEMYSGSDD -48
S331      - RGVPPPPTVEGAPAPRARTAGIQRIPL---------PPPPAPETYEEYGY -41
C24       - DDDFNKLPKKAKGKAQKSNKKWDGSEEDEDNSKKIKERSRINSSGESGDE -98
S331      - DDTY----------AEQSYEGYEGYYSQSQGDSEYYDYGH----GEVQDS -77
C24       - SDEFLQSR-KGQKKNQKNKPGPNIESGNEDDDASFKIKTVAQKKAEKKER -147
S331      - YEAYGQDDWNGTRPSLKAPPARPVKGAYREHPYGYR                -113
C24       - ERKKRDEEKAKLRE -161
```

highly homologous to a rat annexin V binding protein, ABP7, with an identity of 76.40 % (Figure 14A). However, in alignments between C-2.4 and Sam68 sequences, no significant homology was found. Clone C-4.1 consisted of 468 nucleotides, that encoded a stretch of 186 amino acids. The protein sequence was very rich in serine amino acid residues (18%). Additionally, several proline-rich motifs were found scattered throughout its sequence (Figure 15). BLASTX searches revealed no significant homologies to the C-4.1 sequence. Therefore, this clone likely represents a novel class of protein.

C. 5. Further library screening in order to obtain the full transcript of clones C-2.4 and C-4.1

In order to find the full length cDNA of C-2.4 and C-4.1, the same λgt11 cDNA expression library was screened using DNA probes synthesized from C-2.4 and C-4.1 (see materials and methods). During the primary, secondary and tertiary screening process of 1.2 X 10^6 λ phages, twenty four purified λ phage clones were finally isolated. Southern blot analysis, following PCR amplification of the clones, showed that 16 clones were specifically hybridized to the C-2.4 probe and eight clones were specifically hybridized to the C-4.1 probe (Figure 16A). However, both groups of clones exhibited the same molecular size as the original cDNA clones (C-2.4=~500 bp, C-4.1= ~550 bp). Sequencing three randomly chosen clones revealed the same nucleotide sequence in C-2.4 and C-4.1. Screening using a PCR amplification method was also carried out using λgt11 forward or reverse primers and internal primers of either C-2.4 or C-4.1 (Figure 16B). Cloning and sequencing of the isolated PCR fragments revealed identical cDNA sequence to that seen in the original clones.

Figure 15. Sequence analysis of clone C-4.1. Clone C-4.1 consisted of 558 nucleotides which encoded 186 amino acids. This protein was very rich in serine amino acid residues (18%). Two proline-rich motifs were found in its sequence. However, no stop codon was found, indicating that this clone did not represent the full cDNA transcript.

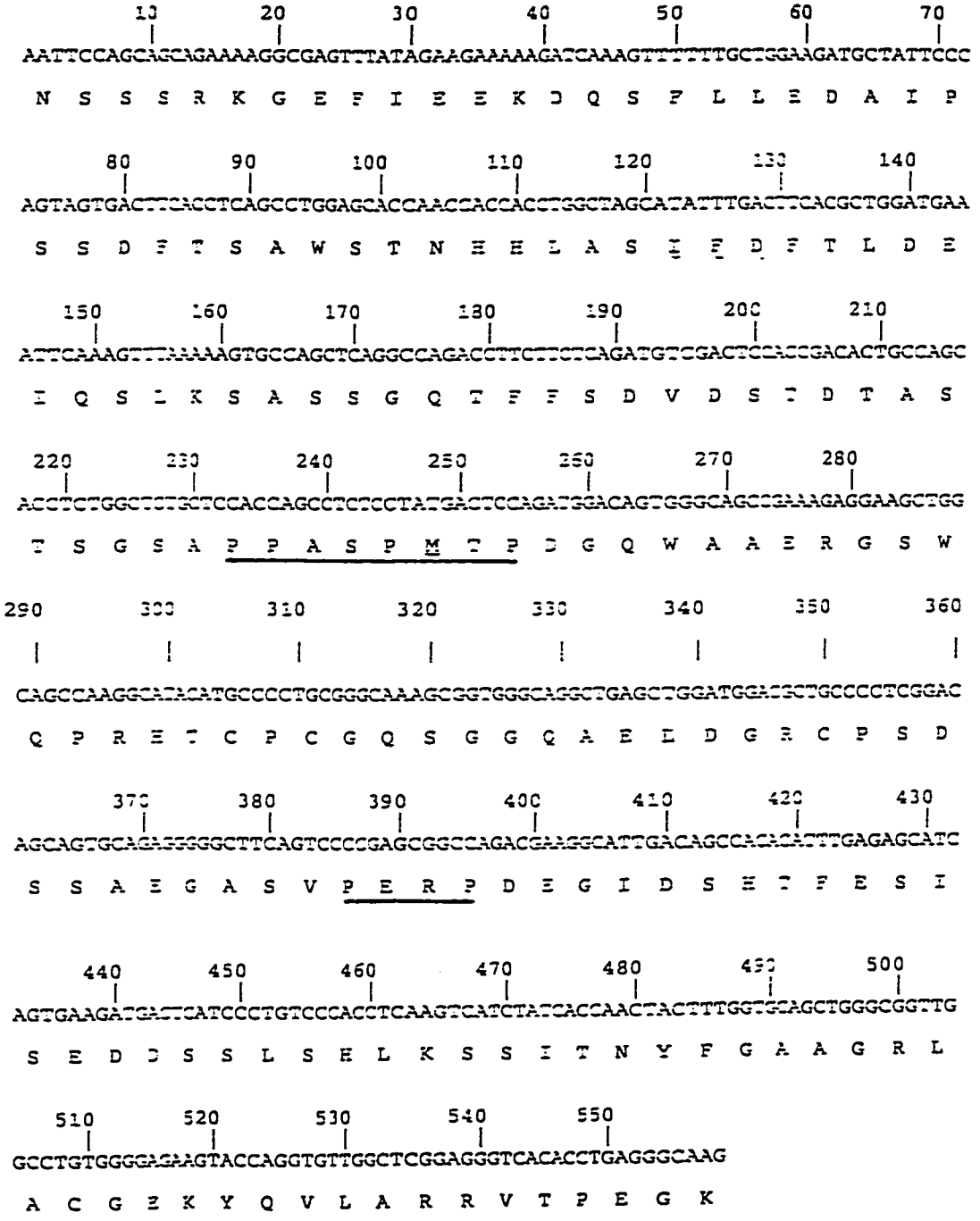
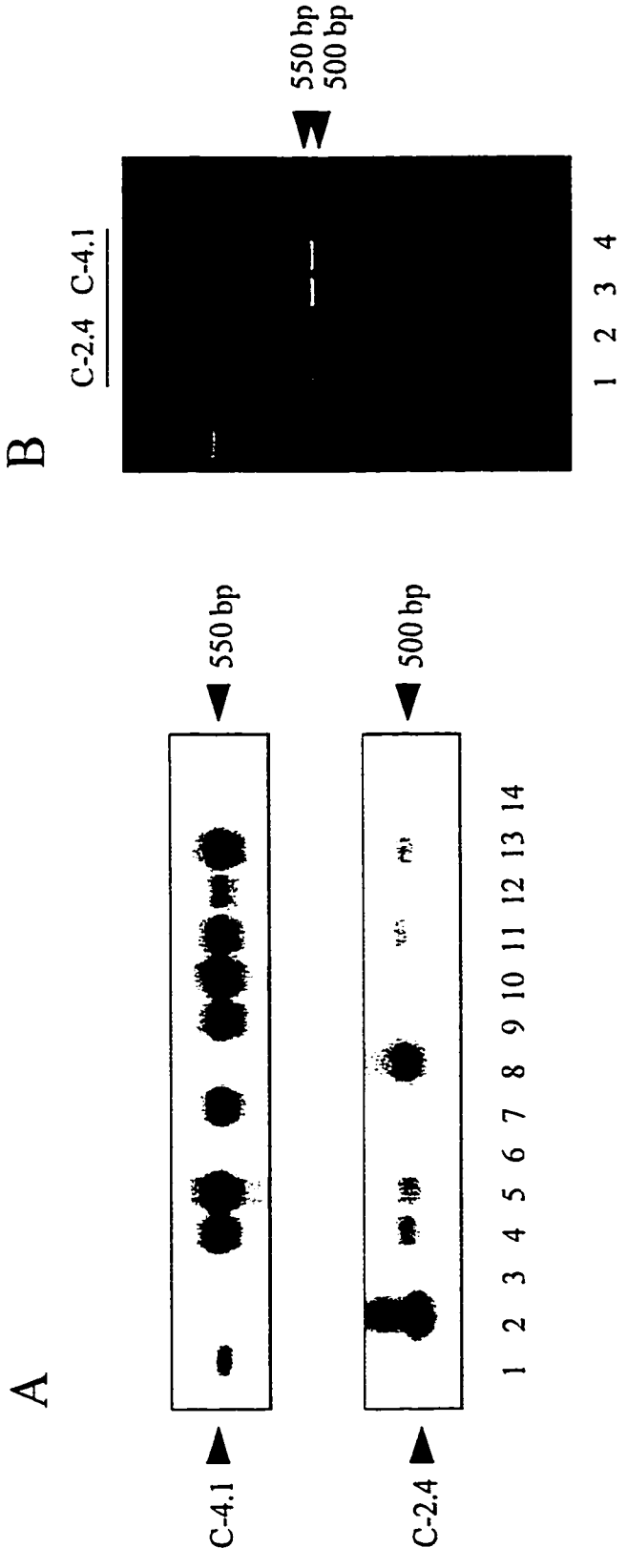

10 20 30 40 50 60 70
AATTCCAGCAGCAGAAAAGGCGAGTTTATAGAAGAAAAAGATCAAAGTTTTTTGCTGGAAGATGCTATTCCC
N S S S R K G E F I E E K D Q S F L L E D A I P
80 90 100 110 120 130 140
AGTAGTGACTTCACCTCAGCCTGGAGCACCAACCACCACCTGGCTAGCATATTTGACTTCACGCTGGATGAA
S S D F T S A W S T N H H L A S I F D F T L D E
150 160 170 180 190 200 210
ATTCAAAGTTTAAAAAGTGCCAGCTCAGGCCAGACCTTCTTCTCAGATGTCGACTCCACCGACACTGCCAGC
I Q S L K S A S S G Q T F F S D V D S T D T A S
220 230 240 250 260 270 280
ACCTCTGGCTCTGCTCCACCAGCCTCTCCTATGACTCCAGATGGACAGTGGGCAGCCGAAAGAGGAAGCTGG
T S G S A P P A S P M T P D G Q W A A E R G S W
290 300 310 320 330 340 350 360
CAGCCAAGGCATACATGCCCCTGCGGGCAAAGCGGTGGGCAGGCTGAGCTGGATGGACGCTGCCCCTCGGAC
Q P R E T C P C G Q S G G Q A E L D G R C P S D
370 380 390 400 410 420 430
AGCAGTGCAGAGGGGCTTCAGTCCCCGAGCGGCCAGACGAAGGCATTGACAGCCACACATTTGAGAGCATC
S S A E G A S V P E R P D E G I D S H T F E S I
440 450 460 470 480 490 500
AGTGAAGATGACTCATCCCTGTCCCACCTCAAGTCATCTATCACCAACTACTTTGGTGCAGCTGGGCGGTTG
S E D D S S L S H L K S S I T N Y F G A A G R L
510 520 530 540 550
GCCTGTGGGGAGAAGTACCAGGTGTTGGCTCGGAGGGTCACACCTGAGGGCAAG
A C G E K Y Q V L A R R V T P E G K

Figure 16. Southern blot analyses of the clones isolated from a DNA hybridization screening using C-4.1 and C-2.4 probes. (A) C-4.1 probes detected 8 positive clones (lane 4, 5, 7, 8, 9, 10, 11 and 12) that exhibited the same molecular sizes as the control C-4.1 (lane 1). Using C-2.4 probes detected a single band (lane 8) which was the same molecular size as the control C-2.4 (lane 2). (B) PCR amplification screening using λgt11 forward or reverse primers and internal primers of either C-2.4 or C-4.1 isolated PCR products (lane 2 and 4) with the sizes identical to the controls, C-2.4 and C-4.1 (lane 1 and 3).


A
550 bp
500 bp
C-4.1
C-2.4
1 2 3 4 5 6 7 8 9 10 11 12 13 14
B
C-2.4 C-4.1
550 bp
500 bp
1 2 3 4

D. Characterization of C-4.3 protein

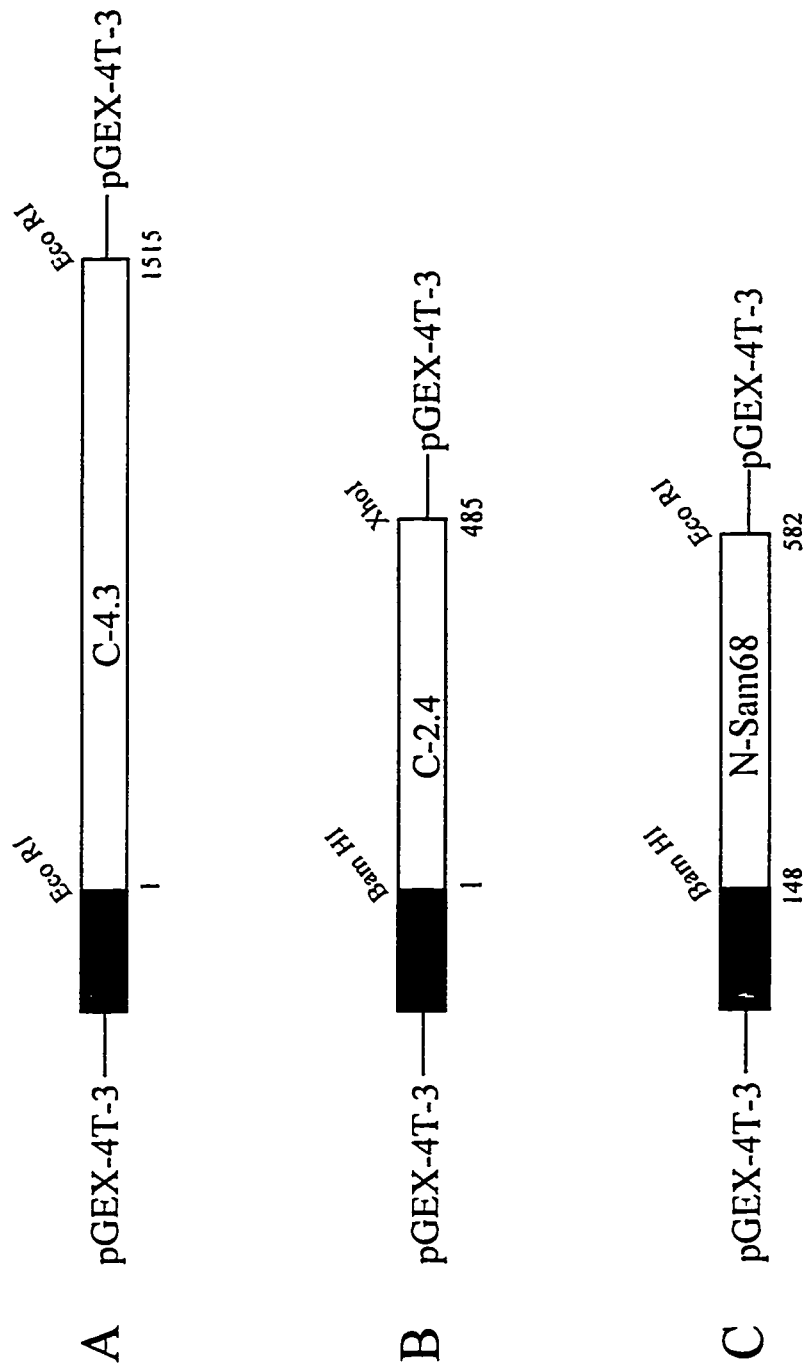
The main focus of this project shifted to the protein characterization of clone C-4.3 for three reasons. First, only C-4.3 possessed a full length cDNA transcript. Second, Western blot analysis of expressed GST-C-4.3 fusion protein showed antigenic relatedness to Sam68. Finally, a rabbit antiserum of C-4.3 protein was succesfully synthesized. An attempt was made to characterize the size of the protein, its localization within cells and its possible role as a putative Src target.

For the purposes of expression, purification and characterization of C-4.3, glutathione S-transferase (GST) constructs were synthesized. The full fragment cDNA of C-4.3 was inserted in-frame into pGEX-4T-3 vector at the C-terminus of the GST gene (Figure 17A). Additionally, a GST construct of the full fragment cDNA of C-2.4 (Figure 17B) and a GST construct of a cDNA fragment of the N-terminal region of Sam68 corresponding to nucleotides 148-582, called GST-N-Sam68 was synthesized (Figure 17C). This latter cDNA was obtained by PCR amplification of a λgt11 cDNA library using internal primers designed from the published sequence of Sam68 (Wong *et al.*, 1992). The proper reading frame of all constructs was confirmed by sequencing on the junction regions.

D. 1. Expression and purification of the GST-C-4.3 protein.

The GST constructs were expressed in *E-coli* strain XL1B or BL21(DE3)pLysS. Unlike GST-C-2.4 and GST-N-Sam68 fusion proteins that were expressed in the soluble fraction of the cell lysates, GST-C-4.3 fusion protein was found in the insoluble fraction. Therefore, certain culture conditions had to be altered in order to increase the solubility of

Figure 17. Glutathione S-transferase (GST) fusion protein constructs. (A) Nucleotide sequences from positions 1-1515 of clone C-4.3 were inserted into Eco RI sites at the multiple cloning sites of plasmid PGEX-4T-3. (B) Nucleotide sequences from positions 1-485 of clone C-2.4 were inserted into Bam HI and Xho I sites at the mutiple cloning sites of plasmid PGEX-4T-3. (C) Sam68 sequences coresponding to nucleotide positions 148-152 (N-Sam68) were inserted into Bam HI and Eco RI sites at the multiple cloning sites of plasmid PGEX-4T-3.

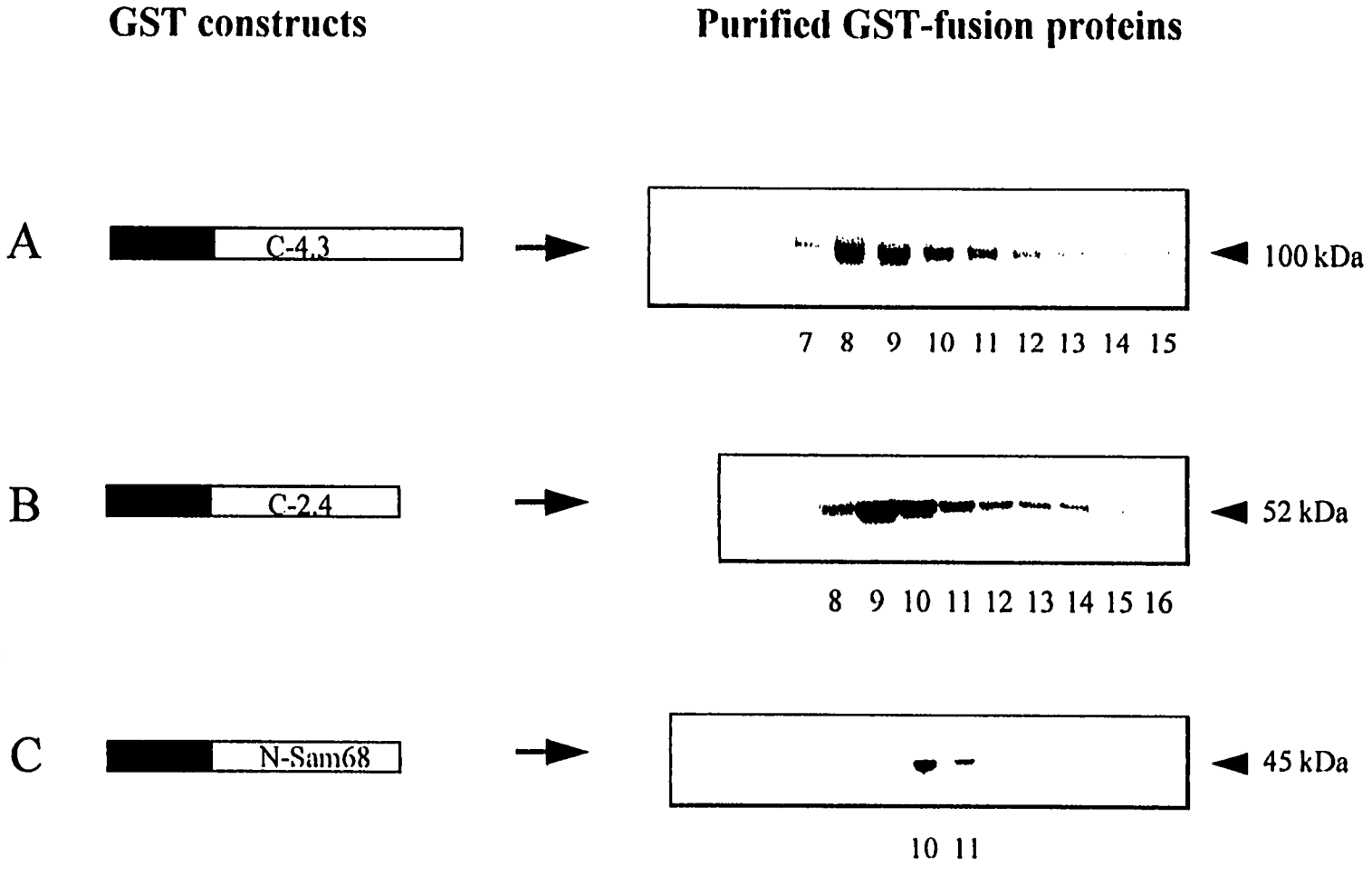

A
pGEX-4T-3
Eco RI
C-4.3
1
1515
pGEX-4T-3
B
pGEX-4T-3
Bam HI
C-2.4
1
485
Xho I
pGEX-4T-3
C
pGEX-4T-3
Bam HI
N-Sam68
148
582
Eco RI
pGEX-4T-3

the expressed GST-C-4.3 protein. GST-C-2.4 and GST-N-Sam68 were expressed inside the *E-coli* strain XL1B, whereas GST-C-4.3 was expressed inside *E-coli* strain BL21(DE3)pLysS. In addition, a purification approach described by Frangione and Neel (1993) was applied to achieve better solubilization and binding of the C-4.3 fusion protein to glutathione agarose beads. On average, a yield of approximately 3 mg of purified the GST-C-4.3 fusion protein was obtained from 500 mL of bacterial culture (Figure 18A). In contrast, a higher yield was obtained from GST-C-2.4 and GST-N-Sam68 constructs; for 500 mL of bacterial culture yielded around 5 mg of the purified GST fusion proteins (Figure 18B and C). This was due to the fact that the latter constructs were able to be induced with higher concentration of IPTG (0.4 mM) and for a longer periods of time (~5 hours).

D. 2. Anti-C-terminal Sam68 polyclonal antibody reacted with GST-C-4.3 and GST-C-2.4 fusion proteins.

In order to confirm that the isolated clones coded for proteins were recognizable by the anti-C-terminal Sam68 polyclonal antibody, Western blot analysis was performed. GST-C-4.3, GST-C-2.4 fusion proteins and a control GST protein were resolved on 10% SDS-PAGE gel and were then transferred onto a nitrocellulose membrane as described in materials and methods. Upon immunoblotting with anti-C-terminal Sam68 polyclonal antibody, a band of approximately 100 kDa (GST-C-4.3 fusion protein) (Figure 19, lane 3) and another band of approximately 52 kDa (GST-C-2.4 fusion protein) were detected (Figure 19, lane 2). In contrast, a 27 kDa GST protein alone was not detected (Figure 19, lane 1). This result indicated that the anti-C-terminal Sam68 polyclonal antibody specifically

Figure 18. Purification of the expressed GST-fusion proteins. The purified GST-fusion proteins were all eluted through GST column, and the fractions were analyzed on 10% SDS-PAGE gels stained with Coomassie blue. (A) The purified GST-C-4.3 fusion protein came up primarily in fractions 8 and 9, and exhibited a molecular mass of approximately 100 kDa. (B) Similarly, the purified GST-C-2.4 fusion protein came up primarily in fractions 8 and 9, and exhibited a molecular mass of approximately 52 kDa. (C) The purified GST-N-Sam68 fusion protein came up primarily in fraction 10, and exhibited a molecular mass of approximately 45 kDa.


GST constructs
Purified GST-fusion proteins
A
C-4.3
100 kDa
7 8 9 10 11 12 13 14 15
B
C-2.4
52 kDa
8 9 10 11 12 13 14 15 16
C
N-Sam68
45 kDa
10 11

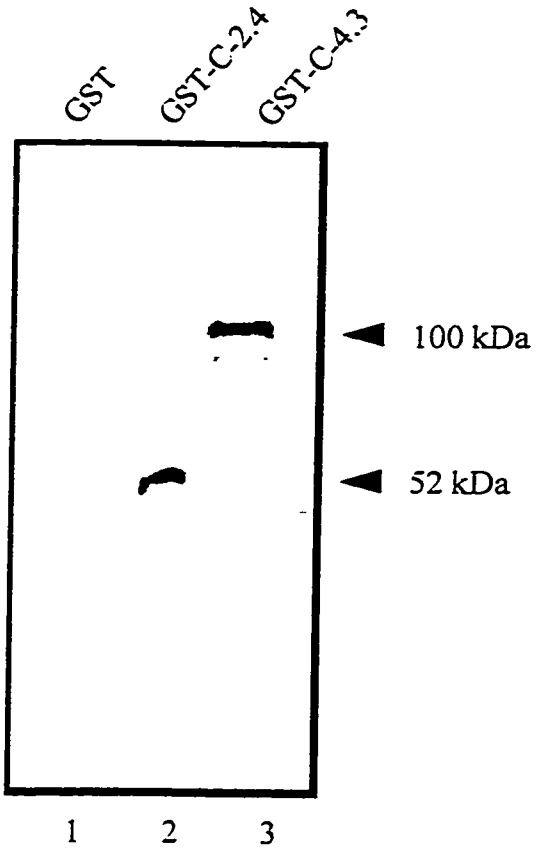


Figure 19. Anti-C-terminal Sam68 polyclonal antibody reacted with GST-C-4.3 and GST-C-2.4 fusion proteins. Two μg of GST-C-4.3 or GST-C-2.4 fusion proteins and a control GST protein were resolved on a 10% SDS-PAGE gel and transferred onto a nitrocellulose membrane. Upon immunoblotting with anti-C-terminal Sam68 polyclonal antibody, a band of approximately 100 kDa (the GST-C-4.3 fusion protein) (lane 3) and another band of approximately 52 kDa (the GST-C-2.4 fusion protein) (lane 2) were detected. A band of 27 kDa (the GST protein alone) was only weakly detected.

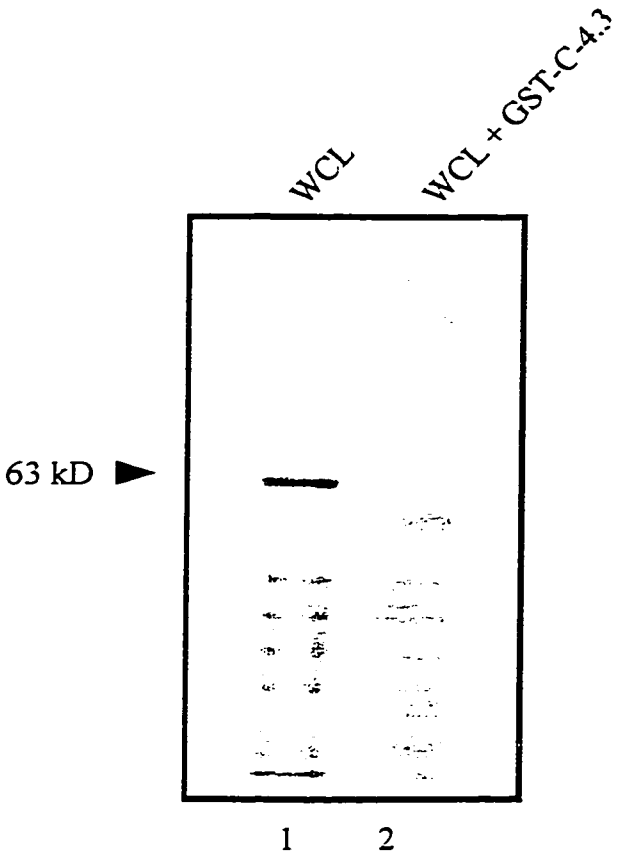
recognized proteins expressed from these isolated clones. Therefore, it was evident that the isolated clones, C-4.3 and C-2.4, exhibited some antigenic identity to Sam68.

D. 3. Production and purification of anti-C-4.3 polyclonal antibody

In order to characterize the biological properties of a protein, such as its molecular mass, cellular localization and possible functions, an antibody is necessary. A polyclonal antibody against C-4.3 was generated by immunizing a New Zealand White (NZW) rabbit with the full length purified GST-C-4.3 fusion protein. Additionally, an attempt was also made to synthesize polyclonal antibodies against the full length purified GST-C-2.4 and the purified GST-N-Sam68 peptide. The rabbit immunized with GST-N-Sam68 peptide died shortly after the first injection. The NZW rabbit antisera bearing anti-C-4.3 polyclonal antibody was affinity purified using purified GST-fusion proteins linked to CNBr-activated sepharose beads. For every mL of rabbit sera, around 125-150 μg purified anti-C-4.3 polyclonal antibody was obtained.

D. 4. The C-4.3 protein exhibited a molecular weight of 63 kDa and was localized in the Golgi complex.

Lysates of NIH 3T3 fibroblasts were resolved on 10% SDS-PAGE gel and transferred onto a nitrocellulose membrane, as described in materials and methods. Upon blotting the membrane with anti-GST-C-4.3 polyclonal antibody, a single band of approximately 63 kDa was identified (Figure 20, lane 1). No bands of 95 kDa (golgin-95) or of 130 kDa (GM130) were observed, indicating that in fibroblasts the antibody did not cross-react with golgin-95,



Blot using anti-C-4.3 antibody

Figure 20. A polyclonal antibody raised against a GST-C-4.3 fusion protein recognized a 63 kDa protein in Western blot analysis of NIH 3T3 fibroblasts. 30 μg of NIH 3T3 whole lysate was separated by 10% gel SDS-PAGE gel, transferred onto a nitrocellulose membrane and immunoblotted with newly synthesized anti-C-4.3 polyclonal antibody. A band of molecular weight of approximately 63 kDa was identified (lane 1). Pre-incubation of 10 μg of purified GST-C-4.3 fusion protein with the anti-C-4.3 polyclonal antibody, competitively removed the 63 kDa band, as shown by Western blot analysis (lane 2).

or GM-130. It is possible that the expression of these proteins is tissue specific. Furthermore, addition of the purified GST-C-4.3 fusion protein into the immunoblot solution resulted in the loss of the 63 kDa band (Figure 20, lane 2), indicating that in fibroblasts the antibody reacted specifically to a protein with molecular weight of 63 kDa.

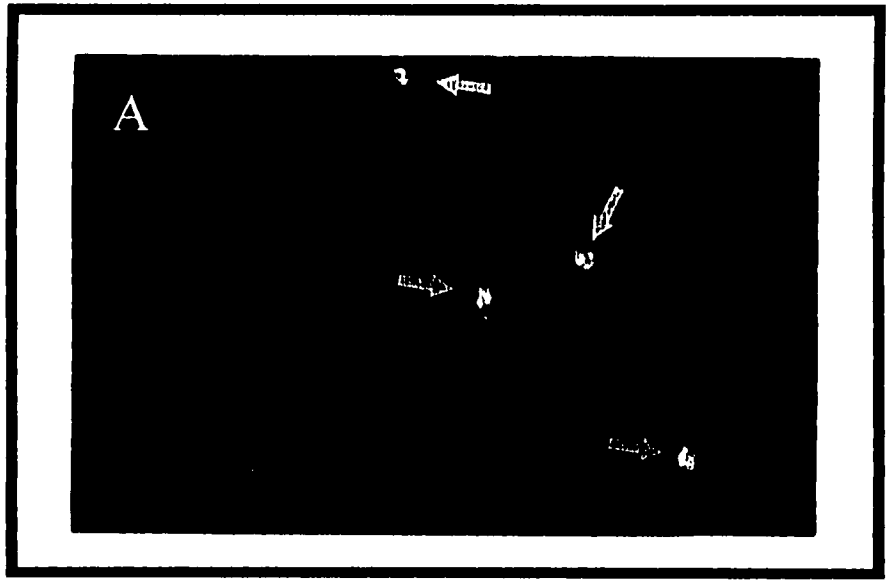
Mapping of the location of the protein inside the cell was performed with indirect immunofluorescence method (IIF) in NIH 3T3 fibroblasts. Detection was achieved by applying a secondary antibody of goat anti-rabbit IgG labelled with rhodamine. A discrete signal was observed in the Golgi complex region in every cell (Figure 21A), indicating that C-4.3 was localized in the Golgi complex. This result confirmed the sequence data that this protein was related to golgin-95 and the cis-Golgi Matrix Protein, GM130. The C-4.3 localization in the Golgi complex was further confirmed using HepG2 cells, prepared at various stages of mitosis. However, to further confirm the precise localization of the protein, a co-localization with antibodies specific for Golgi complex as well as an electron microscope examination might be required. An interesting result observed in the localization experiments with anti-C-4.3 antibody in NRK (newborn rat kidney) cells was that no protein was observed. One possible explanation was that C-4.3 might not be expressed in NRK cells.

D.5. C-4.3 associated with and was phosphorylated by c-Src *in vitro*

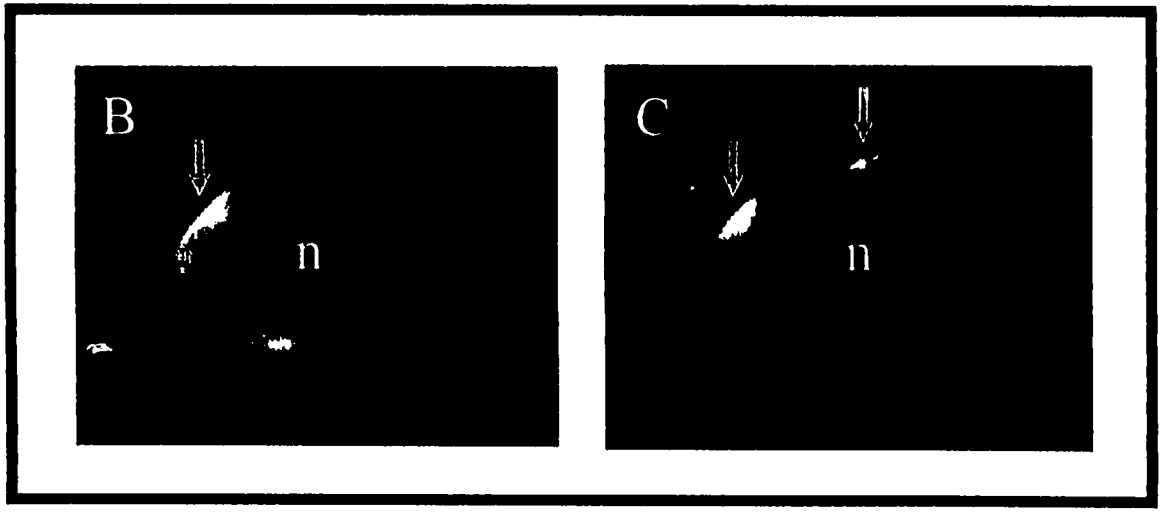
Sam68 was observed to be tyrosine phosphorylated in mitotically arrested Src-transformed cells. Additionally, Sam68 also associated with activated c-Src *in vivo,* and the SH3 domain of c-Src *in vitro* (Fumagalli *et al.*, 1994; Taylor and Shalloway, 1994). Because C-4.3 was isolated using anti-Sam68 antibody, it might be functionally related to Sam68.

Figure 21. C-4.3 localized to the Golgi and co-localized with Src at the perinuclear membrane in NIH 3T3 fibroblasts. (A) Indirect immunofluorescence (IIF) of anti-C-4.3 polyclonal antibody on NIH 3T3 cells mapped the localization of p63 to the Golgi complex (arrows). (B and C) Double immunostaining of anti-p63 polyclonal antibody and anti-Src monoclonal antibody (2-17) on NIH 3T3 cells. Co-localization was observed in some cells. Arrows indicate co-localization; n = nucleus. (These pictures were taken with the assistance of Dr. J.B. Rattner).

Anti-C-4.3



Anti-C-4.3 Anti-Src

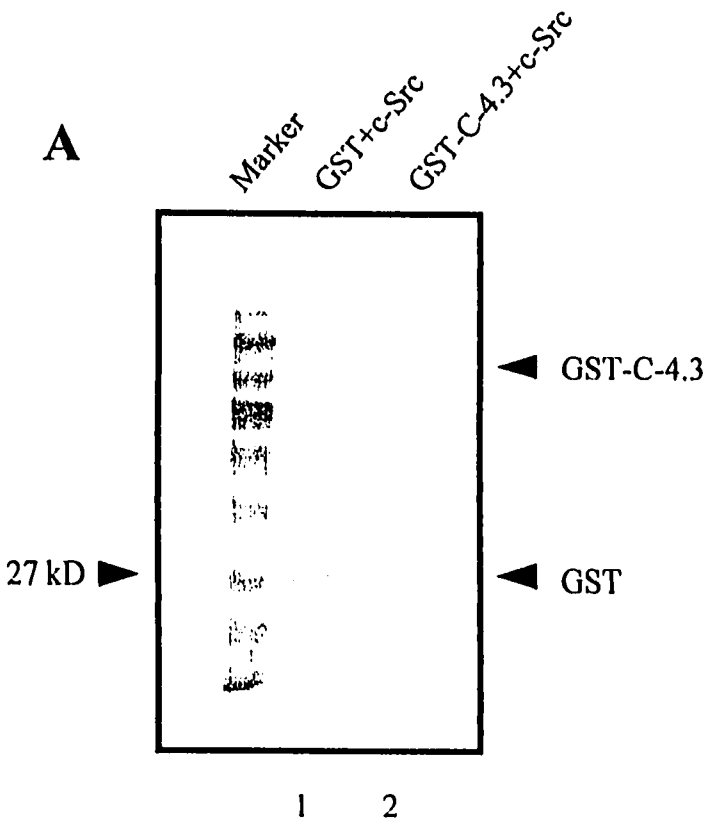


Additionally, C-4.3 exhibited two proline-rich regions that could be potential binding sites for the SH3 domain of Src protein-tyrosine kinase. As a result, it was thought that like Sam68, C-4.3 could also be associated with and phosphorylated by Src protein-tyrosine kinase.

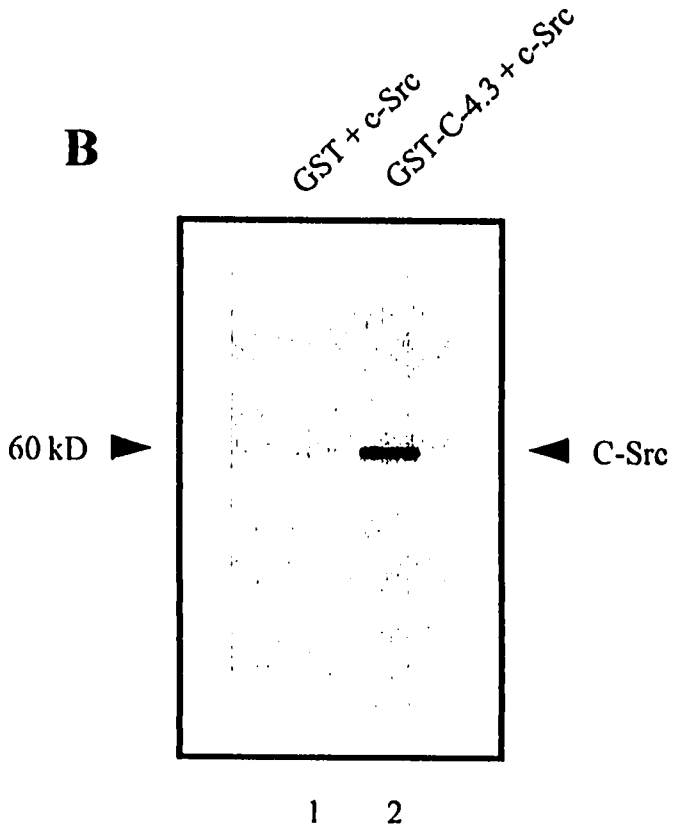
A purified baculovirus-expressed c-Src was used for *in vitro* binding as well as phosphorylation experiments. Purified GST-C-4.3 fusion protein and control GST protein were bound to glutathione agarose beads. After incubation with purified baculovirus-expressed c-Src, the beads were washed 3 times with 0.1% Triton X-100 and the proteins resolved on 10% SDS-PAGE gel, after which they were transferred onto a nitrocellulose membrane as described in materials and methods. Upon immunoblotting with the anti-Src-327 monoclonal antibody, Src was identified as a 60 kDa band in lane two, which contained the GST-C-4.3 fusion protein. On the other hand, no 60 kDa band was seen in lane one, which contained the GST protein alone (Figure 22B). This result indicated that c-Src associated specifically with the C-4.3 portion of the GST-C-4.3 fusion protein.

For the *in vitro* phosphorylation assay, $MgCl_2$ and [γ–^{32}P] ATP were added for proper kinase activation and detection of protein phosphorylation. GST-C-4.3, GST-C-2.4 fusion proteins and GST protein alone (control) were separately incubated with purified baculovirus-expressed c-Src and then resolved on a 10% SDS-PAGE gel and the gel was then dried. Autoradiography revealed a phosphorylation signal in lane three at around 100 kDa, the predicted size of the GST-C-4.3 fusion protein (Figure 23). The GST protein alone (27 kDa) and GST-C-2.4 fusion protein (52 kDa) were not phosphorylated (Figure 23, lane 1 and 2). This indicated that c-Src tyrosine kinase specifically phosphorylated C-4.3 protein

Figure 22. GST-C-4.3 associated with purified baculovirus expressed c-Src *in vitro*. Ten μg of either GST or GST-C-4.3 fusion protein bound to glutathione sepharose beads was incubated with 100 ng of purified baculovirus expressed c-Src in a binding buffer and washed with 0.1% Triton X-100 three times. The precipitated proteins were then separated by 10% SDS-PAGE gel and transferred onto a nitrocellulose membrane. (A). Ponceau S staining of the membrane revealed a 27 kDa band (the GST protein) (lane 1) and a 100 kDa band (GST-C-4.3) (lane 2). (B). Immunoblotting with anti-Src monoclonal antibody (327) showed that the GST protein alone did not associate with Src (lane 1), whereas the GST-C-4.3 fusion protein did (lane 2).



Ponceau S staining



Blot using anti-Src 327

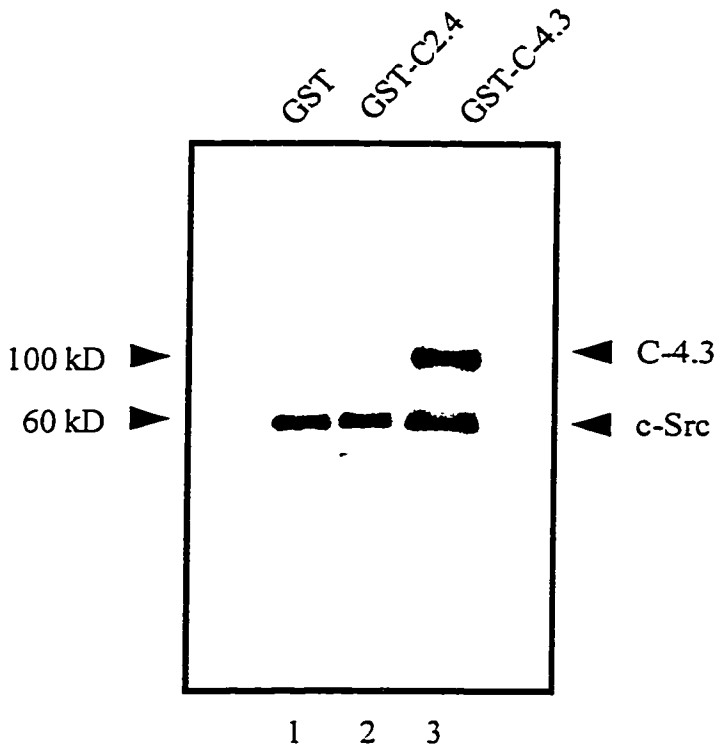


Figure 23. GST-C-4.3 phosphorylation by purified baculovirus expressed c-Src *in vitro*. Ten μg of either purified GST, GST-C-2.4 or GST-C-4.3 fusion proteins were incubated with 30 ng of purified baculovirus expressed c-Src. [γ-^{32}P] ATP and $MgCl_2$ were added to initiate the reaction. Autoradiography showed tyrosine phosphorylation of GST-C-4.3 (100 kDa) by purified baculovirus expressed c-Src *in vitro* (lane 3), whereas neither GST (27 kDa) nor the GST-C-2.4 peptide were phosphorylated (lanes 1 and 2, respectively). The 60 kDa band observed in each lane was autophosphorylated c-Src.

in vitro. The 60 kDa band that was observed in every lane was identified as the auto-phosphorylated form of Src.

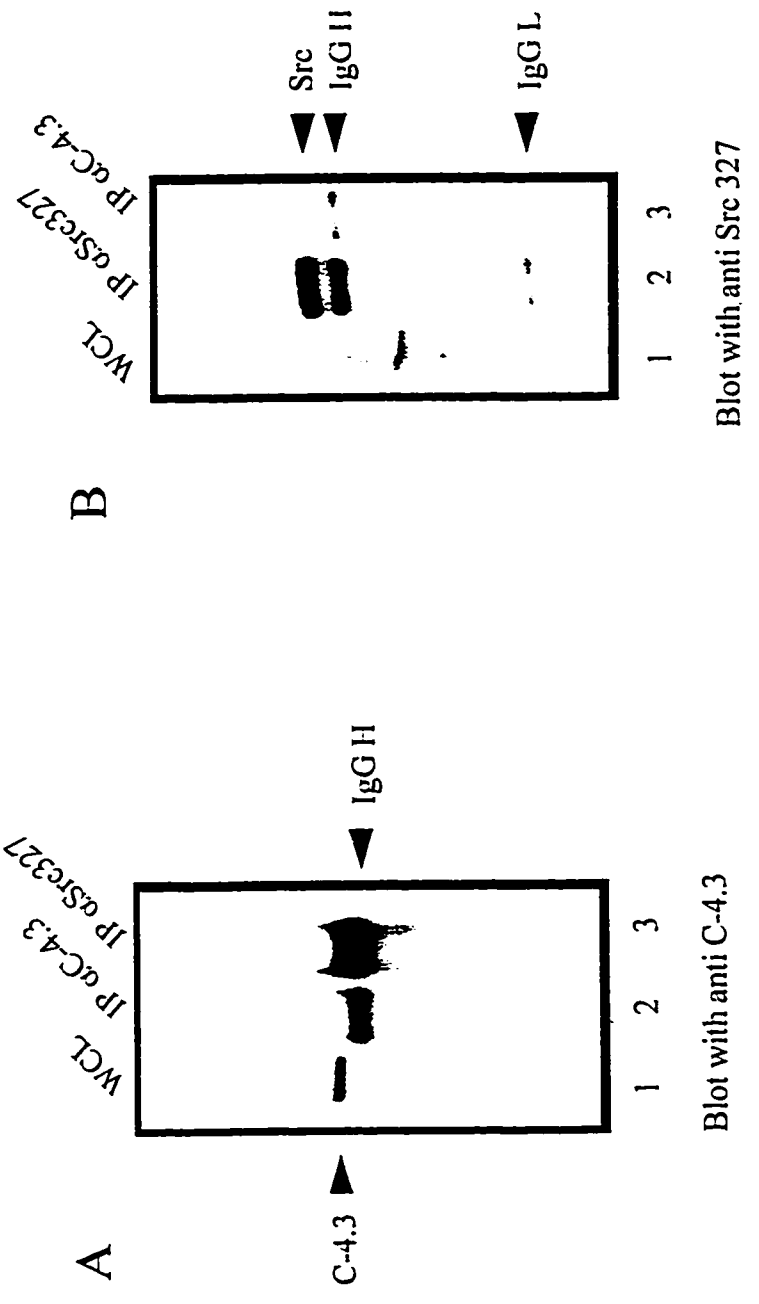
D.6. C-4.3 did not co-immunoprecipitate with Src in NIH 3T3 fibroblasts

In order to determine whether C-4.3 associates with and is phosphorylated by Src tyrosine kinase *in vivo,* co-immunoprecipitation experiments using cell lysates of NIH 3T3 fibroblasts were carried out. For immmunoprecipitating purposes, anti-Src antibodies were used because the anti-C-4.3 polyclonal antibody was not able to immunoprecipitate C-4.3 (Figure 24A, lane 2). However, both anti-Src 327 monoclonal antibody and anti-Src 2-17 monoclonal antibody (the latter was done by another graduate student in our laboratory) failed to immunoprecipitate C-4.3 in association with Src from cell lysates of NIH 3T3 fibroblasts, as determined by Western blot analysis of the lysates with the anti-C-4.3 polyclonal antibody (Figure 24B, lane 3). This result, however, contradicted the *in vitro* data, that C-4.3 associated with Src.

D.7. Src was co-localized with C-4.3 in the perinuclear region of NIH 3T3 fibroblasts

In addition to its localization at the plasma membranes, Src is also found in perinuclear membranes, including endosomes and secretory vesicles of fibroblast cells (Kaplan *et al.*, 1992). There is an increasing body of evidence suggesting that in addition to Src SH4, Src SH2 and SH3 domains play a key role in the cellular localization of this protein (Brown and Cooper, 1996). Based on the *in vitro* data that C-4.3 associated with and was phosphorylated by Src tyrosine kinase, and the idea that the binding might be mediated by

Figure 24. Immunoprecipitations of C-4.3 and Src proteins from NIH 3T3 fibroblast lysates. (A) 500μg of NIH 3T3 lysates were immunoprecipitated with anti-C-4.3 polyclonal antibody or anti-Src 327 monoclonal antibody. The immunoprecipitates were resolved on 10% SDS-PAGE gel and transferred onto a nitrocellulose membrane. Upon immunoblotting with anti-C-4.3 polyclonal antibody, no band with molecular mass of 63 kDa (the C-4.3 protein) was identified (lane 2). (Lane 1) Control Western blotting of NIH 3T3 fibroblast whole cell lysates with anti C-4.3 polyclonal antibody identified a band with a molecular mass of approximately 60-63 kDa. (B) Immunoblotting with anti-Src 327 monoclonal antibody identified a band with a molecular mass of approximately 60 kDa (the Src protein) (lane 2). Ig H = Immunoglobulin heavy chain, Ig L = Immunoglobulin light chain.


A
C-4.3
IgG H
WCL
IP αC-4.3
IP αSrc327
1 2 3
Blot with anti C-4.3
B
Src
IgG H
IgG L
WCL
IP αSrc327
IP αC-4.3
1 2 3
Blot with anti Src 327

SH2 and SH3 domains of Src, it was thought that Src might co-localize in cells with C-4.3.

Indirect immunofluorescence techniques were used to show co-localization of C-4.3 and Src (IIF) as described by Kaplan *et al*, (1992). In NIH 3T3 fibroblasts, it was shown that double immunostaining with anti-C-4.3 polyclonal antibody and anti-Src 2-17 monoclonal antibody produced a similar pattern of staining at the perinuclear region of the cell (arrows) (Figure 21B and 21C), indicating that C-4.3 and Src were most likely co-localized. Since this co-localization between Src and C-4.3 was observed only in a few cells, it was possible that this co-localization event might occur only at a specific stage of the cell cycle.

E. Summary of the results

- Screening a T-lymphocytes λgt11 cDNA library using an antiserum raised against an N-terminal region of Sam68 resulted in the isolation of several clones that had homologies to RNA binding proteins, whereas using an antiserum raised against a C-terminal region of Sam68 resulted in the isolation of a clone, C-4.3, that exhibited some structural and antigenic relationships to Sam68.
- The predicted sequence of C-4.3 possesses a coiled-coil motif and proline-rich regions, which are potential binding sites for SH3-containing proteins. C-4.3 exhibits a high homology to golgin-95 and cis-Golgi Matrix Protein, GM130.
- C-4.3 exhibited a molecular weight of approximately 63 kDa and was localized to the Golgi complex in NIH 3T3 fibroblasts.
- C-4.3 associated with and was phosphorylated by Src protein-tyrosine kinase *in vitro*.
- Finally, C-4.3 was observed to be co-localized with Src in the perinuclear region of

NIH 3T3 fibroblasts.

CHAPTER FOUR

DISCUSSION

A. Structural and antigenic relationship between the isolated clones and Sam68

Based on antigenic cross-reactivity and sequence homology, Sam68 was thought to be related to a GTP-activating protein (GAP)-associated protein, p62. Later, it was proven that Sam68 and GAP-associated p62 were distinct proteins (Lock *et al.*, 1996; Taylor *et al.*, 1995). Yet the existence of Sam68-related proteins cannot be ruled out. Vogel and Fujita (1995) observed that a tyrosine-phosphorylated 70 kDa protein associated with the SH2 domain of $p56^{lck}$ and that it possessed sequence homology to Sam68. In addition, a natural isoform of Sam68, known as Sam68ΔKH, was recently identified (Barlat *et al.*, 1997).

In our attempt to isolate Sam68-related proteins, two rabbit antisera generated against both N- and C-terminal regions of Sam68 (Santa Cruz) were used to screen a human T-lymphocyte cDNA expression library. Interestingly, N-5, a clone isolated using antiserum raised against the N-terminal region of Sam68, exhibits high homology to an RNA binding protein, human ribosomal protein S5. Additionally, N-5 possesses conserved residues within the KH domain that are also found in Sam68 and other RNA binding proteins, such as human hnRNP K; fragile-X mental retardation protein, FMR-1, and *Halobacterium halobium* S3 ribosomal protein, Hh S3. Because the KH domain of Sam68 lies within its N-terminal region, against which the antibody was raised, it is possible that these RNA binding regions share conserved epitopes that are recognized by the antibody. However, since the sequence homology between N-5 and Sam68 is only restricted to few conserved residues within the

RNA binding domain, it cannot be concluded that the two proteins are closely related. Further analyses of the antigenic relationship between N-5 and Sam68 were not possible since the rabbit antiserum was no longer commercially available and an attempt to generate it was unsuccessful.

Screening the same cDNA expression library, using an antiserum raised against the C-terminal region of Sam68, however, resulted in the isolation of eleven cDNA clones. A Southern blot hybridization of these clones revealed that they represented three different sets of clones, C-4.3, C-2.4 and C-4.1. The sequence data showed that one clone, C-4.3, exhibited a complete open reading frame (ORF) consisting of 460 amino acids. This result was confirmed by Dr. Edward Chan and co-workers in The Scripps Research Institute in La Jolla, California, who cloned a similar cDNA, called G95-L1-1 (unpublished data). The six amino acids difference between C-4.3 and G95-L1-1 probably due to the differences in the cDNA libraries used. G95-L1-1 was isolated from a human placenta library, whereas C-4.3 was isolated from a human T-lymphocytes library. Structurally, C-4.3 exhibits some similarities to Sam68, especially with the C-terminal region of Sam68, the region against which the Sam68 antibody was raised. One of the two proline-rich regions present in C-4.3 shows a strong (75%) identity to a proline-rich region found in the C-terminal portion of Sam68. It is possible that the epitope recognized by the antiserum could be proline-rich regions. However, this can only be confirmed with an epitope mapping procedure. Furthermore, similar to Sam68, C-4.3 also possesses a nuclear localization signal (NLS) right at the C-terminus of the protein. In the case of C-4.3, however, this NLS is apparently non-functional, because our immunofluorescence experiment using an antibody raised against a GST-C-4.3

fusion protein showed that the C-4.3 protein was localized in the Golgi complex, not in the nucleus, in NIH 3T3 fibroblasts. Finally, Western blot analysis with anti Sam68 antibody confirmed that the expressed GST-C-4.3 fusion protein was antigenically related to Sam68. Another isolated clone, C-2.4, also demonstrated an antigenic identity to Sam68. C-2.4 possesses many NLSs and exhibits a very hydrophatic sequence, suggesting that like Sam68, C-2.4 may be a nuclear protein.

The data above, collectively, show that screening of a cDNA library using antisera raised against both the N- and C-terminal regions of Sam68 result in the isolation of several clones that exhibit structural and antigenic similarities to Sam68. However, the minimal sequence homology between the isolated clones and Sam68 do not support a close relationship between these proteins. Possibly, these proteins share more similarities in terms of their biological functions in the cell. In the case of N-5, its functional relationship to Sam68 may be in the regulation of RNA metabolism; whereas C-4.3, as suggested by our data, may have similar biological functions as Sam68, such as being a target of Src tyrosine kinase. Therefore, further studies on N-5 and C-4.3 functions are necessary in order to confirm this hyphothesis, which may as well, contribute towards a better understanding of Sam68 function and the role of Src tyrosine kinase.

B. Coiled-coil motif in C-4.3 and its possible functions

An interesting feature of C-4.3 is the presence of a coiled-coil motif in its sequence. This motif, also known as leucine zipper, is characterized by six repeats of leucine residues at every seventh position (Landschultz *et al.*, 1988). Two amphiphatic α-helices form this

motif so that structurally it resembles a coiled-coil (O'Shea *et al.*, 1989). Many auto-antigens that react with auto antibodies present in SLE (systemic lupus erythematosus) and Sjörgren's syndrome have been noted to possess coiled-coil structures. These include 52-kDa SS-A/Ro (Chan *et al.*, 1991) and 80/86-kD Ku (Yaneva *et al.*, 1989). Leucine zipper motifs have also been described to be present in DNA binding proteins, such as c-fos and c-jun (Turner and Tjian, 1989), as well as in structural and motor proteins, such as myosins and kinesin (Mu *et al.*, 1995). The function of coiled-coil motifs, especially in structural proteins is believed to play a role in the multimerization of the protein. However, a coiled-coil structures can also be dynamic and play a central role in generating conformational changes, resulting in dramatic movement of one part of a protein relative to another. This mechanism is implicated in the translocation of the nonclaret disjunctional kinesin-related microtubule motor protein on microtubules toward their minus ends which is required for proper chromosomes segregation in *Drosophila* oocytes (Endow *et al.*, 1994).

While it shares only weak homology with other coiled-coil-containing proteins such as structural and motor proteins, C-4.3 exhibits a high degree of homology to two coiled-coil-containing autoantigens, golgin-95 and cis-Golgi Matrix Protein, GM130. However, compared to C-4.3, the coiled-coil region of golgin-95 has a long insertion consisting of a stretch of acidic amino acids; whereas the coiled-coil region in GM130 has a long addition to its N-terminal site. Additionally, the N-terminal region of GM130 exhibits many phosphorylation motifs for various protein kinases, suggesting that it may have regulatory functions *in vivo*. Amino acids 70-260 of C-4.3 exhibits a classical coiled-coil motif with a high degree of confidence (calculated using a method according to Lupas *et al.*, 1991), with

the exception of a short region of amino acids 143-149 which causes a break in the coiled-coil structure. Judging from the nature of the overall homology, it is likely that these three proteins belong to the same family.

The function of golgin-95 is unknown. However, some possible functions have been proposed based on its sequence analysis and its distribution upon brefeldin A (BFA) treatment. BFA is known to cause microtubule-dependent retrograde transport of Golgi element to the endoplasmic reticulum (ER) (Lippincott-Schwartz *et al.*, 1989). Five and ten minute BFA treatments into cells resulted in a remarkable reduction of staining of golgin-95 in the perinuclear Golgi complex and the appearance of vesicular and elongated microtubular structures in HepG2 cells, as detected with the rabbit antisera raised against the recombinant protein. Furthermore, 30 minute BFA treatment resulted in the complete loss of Golgi staining. Consequently, golgin-95 is thought to have a role in the transport of vesicles from the ER to the Golgi complex or within the Golgi stack. In addition, together with other Golgi proteins, it may form cytoskeletal structure that is the framework for transport of Golgi vesicles (Fritzler *et al.*, 1993). Unlike golgin-95, some functions of GM130 have been characterized. This protein, together with p115, a component needed for intra Golgi transport (Waters *et al.*, 1992), has been reported to be involved in the docking of transport vesicles. Interestingly, it was also shown that the C-terminal region of GM130 which contains a coiled-coil motif was critical for binding to Golgi membranes. Furthermore, a microinjection of a cDNA encoding GM130 truncated at its coiled-coil regions into NRK cells, resulted in cytoplasmic localization of the truncated protein and weaker binding to the Golgi apparatus (Nakamura *et al.*, 1997). Referring to this result, it is interesting to speculate whether the C-

4.3 truncated at its coil-coiled regions will also behave the same. An NLS presents in C-4.3 sequence, yet it is apparently nonfunctional (our data indicated that C-4.3 localized to the Golgi complex). It is possible that C-4.3 lacking the coiled-coil regions might localize to the nucleus instead of the Golgi complex.

C. Proline-rich regions of C-4.3 and its possible functions

It has been characterized that proline-rich sequences provide a relatively hydrophobic region which is suitable for binding Src homology three (SH3) domain (Yu *et al.*, 1994, Mayer and Eck, 1995). Two consensus sequences called class I and class II proline rich regions are known to be able to bind with high affinity to SH3 domains. They are R-x-q-P-x-q-P and q-P-x-q-P-x-R, where x is any amino acid and q is any hydrophobic residue, often it is a proline, valine or leucine. These two classes of ligands have different polarities with respect to their binding to SH3 domains (Mayer and Eck, 1995). The P-x-x-P motif was characterized to be a minimal consensus sequence that is essential for the binding to a SH3 domains. Because the two prolines are separated by two amino acids, it enables them to reside on the same face of the helix, which is critical for binding (Cohen *et al.*, 1995).

Two proline-rich regions are present in C-4.3, in which both have consensus sequences required for the binding to SH3 domains. One proline-rich region, resembles the class I binding site, R-x-P-x-q-P-P-x-P-x-q-P, is present in the C-terminal half of the protein. The other proline rich region which contains a double minimal consensus sequence, P-x-q-P-x-q-P, is present at the N-terminal portion of the protein. Interestingly, as in golgin-95 and GM130, this latter proline-rich motif lies within the coiled-coil domain. The disruption of

this coiled-coil motif makes the proteins structurally more rod-shaped with a joint consisting of a proline-rich stretch (Nakamura *et al.*, 1995). In addition, these proline-rich regions provide the proteins with a docking site for SH3 domain-containing proteins. Since it is known that the Src family of protein-tyrosine kinases uses their SH3 domains to recruit their specific subtrates (Mayer and Eck, 1995), C-4.3 may also be a candidate target of Src protein tyrosine kinase. Our *in vitro* data in which C-4.3 associated with and was phosphorylated by c-Src tyrosine kinase supports that hypothesis. However, to confirm whether the binding of C-4.3 to Src tyrosine kinase is mediated by proline-rich regions, a deletion or peptide competition assay needs to be performed.

D. C-4.3, a putative substrate of Src tyrosine kinase.

A number of Src substrates have been characterized. Some bind to and are phosphorylated by Src directly, and others only interact indirectly with Src, yet show increased phosphorylation level on tyrosine in Src transformed cells (Brown and Cooper, 1996). Shc, PI3 kinase, PLCγ, ras-GAP are examples of Src substrates that are phosphorylated both in Src transformed cells and in growth factor-stimulated cells (Pawson, 1993; Schlessinger and Ullrich, 1992; Cantley *et al.*, 1991). Other Src substrates include actin filament associated protein p110, cortactin, Fak, paxillin, tensin, talin, vinculin, $p130^{cas}$ (Wu *et al.*, 1991; Kanner *et al.*, 1990; Reynolds *et al.*, 1989 a, 1989 b) and also Sam68/p70 (Vogel and Fujita, 1995; Fumagali *et al.*, 1994; Taylor and Shaloway, 1994). These particular substrates have been shown to interact with SH2 or SH3 domains of Src. Such binding interactions make it likely that these proteins are direct substrates of Src in cells

(Brown and Cooper, 1996).

As discussed previously, C-4.3 contains two proline-rich regions that may function as binding sites for SH3 domain containing proteins, such as Src protein-tyrosine kinase. Additionally, there are a few tyrosine residues that may serve as phosphorylation sites for Src. Although no ideal Src phosphotyrosine binding motif (YEEI) as described by Songyang *et al.* (1995) presents in C-4.3 sequence, a tyrosine residue at position 6 has adjacent acidic residues which is the preferred environment for the kinase domain of Src to associate with (Songyang *et al.*, 1995). Once a tyrosine residue in C-4.3 is phosphorylated, it may probably serve as a strong binding site for the SH2 domain of Src. *In vitro* data showed that C-4.3 associated with and was phosphorylated by c-Src protein-tyrosine kinase. Furthermore, Src was also observed to be co-localized with C-4.3 in the perinuclear region of NIH 3T3 fibroblasts. These data, collectively, indicate that C-4.3 may be a true substrate of Src kinase in the cell. It may also support the notion that SH3 or SH2 domains are involved in protein-protein interactions that assist Src localization within the cell.

Immuno-precipitation (IP) experiments were unsuccesfull in showing *in vivo* interactions between C-4.3 and Src. Our IP assays used asynchronous lysates of NIH 3T3 fibroblasts, and mainly relied on anti-Src monoclonal antibodies because the antiserum generated against GST-C-4.3 fusion protein was unable to be used for IPs. It is possible that the interaction of C-4.3 and Src within the cell only occurs at certain stages of the cell cycle as is the case for Sam68. Due to its localization in the nucleus, Sam68 only associates with and is phosphorylated by Src protein-tyrosine kinase when the nuclear envelope breaks down at M phase (Fumagalli *et al.*, 1994, Taylor and Shalloway, 1994). Our results indicate that

C-4.3 is most likely localized to the Golgi. As a result, one can speculate that a similar type of association occurs between C-4.3 and Src at some stage of the cell cycle, possibly upon Golgi breakdown at mitosis. Interestingly, Nakamura *et al.*, (1997) found that incubation of GM130 with mitotic cytosol had a marked effect on the mobility of the protein, which could be prevented by prior treatment of mitotic cytosol with a general kinase inhibitor, staurosporine. This indicates that GM130 may be regulated by some protein kinases in cell cycle dependent manner. Similarly, C-4.3 might also be regulated in cell cycle dependent manner by Src tyrosine kinase or possibly other kinases.

E. Future perspective on C-4.3

The finding that C-4.3, a putative golgi protein, associates with and is phosphorylated by Src protein-tyrosine kinase *in vitro*; as well as co-localizes with Src in the perinuclear region, deserves further attention. First of all, an effort should be made to prove *in vivo* association and phosphorylation by Src tyrosine kinase. However, there is the possibility that the association may occur at certain stages of the cell cycle. Consequently, co-immunoprecipitation experiments should be applied on various stages of the cell cycle. Another approach that may facilitate the co-immunoprecipitation of the two proteins is by extracting the Golgi complex out of cells. It will minimize any competing substrates of Src not found at the Golgi, such as Sam68. Peptide competition or deletion assay will be required to determine the binding site of Src as well as the epitope sites recognized by the Sam68 antibody. In addition, cyanogen bromide or chymotrypsin cleavage of the protein may also be required to map the phosphorylation site of C-4.3. Furthermore, a co-localization assay

using anti-Golgi antibodies or an electron microscopy will help map precisely the localization of the protein.

Although it may be a little too early to discuss the function of this protein, some speculations can be made based on our data, as well as data reported by other research groups. One can speculate that C-4.3 phosphorylation by Src may regulate the disassembly-reassembly of the Golgi complex during mitosis. A proline-rich region in C-4.3 is positioned such that it interrupts the coiled-coil motif. As a result, one may speculate that the interaction of SH3 containing proteins with this region may disrupt the coiled-coil domain-mediated protein multimerization (Endow *et al.*, 1994). There is also a possibility that C-4.3 may involve in certain physiological events related to Golgi functions in which Src tyrosine kinase plays a role. Finally, Liebl and Martin (1992) have reported that directing Src to perinuclear membranes (endoplasmic reticulum) prevented transformation. As a result, another possible role of C-4.3 may be to regulate the transforming activity of the Src proto-oncogene by sequestering it at the perinuclear region, away from the plasma membrane where Src is likely to be activated by activated receptor tyrosine kinases.

BIBLIOGRAPHY

Alonso, G., Koegl, M., Mazurenko, N. and Courtneidge, S.A. 1995. Sequence requirements for binding of Src family tyrosine kinases to activated growth factor receptors. J. Biol. Chem. **270**, 9840-9848.

Anderson, D., Koch, C., 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.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidmen, J.G., Smith, J.A. and Struhl, K. (eds). 1990. Current protocol in molecular biology. John Wiley & Sons, New York.

Barlat, I., Maurier, F., Duchesne, M., Guitard, E., Tocque, B. and Schweighoffer, F. 1997. A role for Sam68 in cell cycle progression antagonized by a spliced variant within the KH domain. J. Biol. Chem. **272**, 3129-3132.

Bjorge, J.D., O'Connor, T.J. and Fujita, D.J. 1996. Activation of human pp60c-src (review). Biochem. Cell Biol. **74**, 477-484.

Brown, M.T. and Cooper, J.A. 1996. Regulation, substrates and functions of Src. Biochim. Biophys. Acta **1287**, 121-149.

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

Chackalaparampil, I. and Shalloway, D. 1988. Altered phosphorylation and activation of pp60c-Src during fibroblast mitosis. Cell **52**, 801-810.

Chan, E.K.L., Hamel, J.C., Buyon, J.P. and Tan, E.M. 1991. Molecular definition and sequence motifs of the 52 kD component of human SS-A/Ro autoantigen. J. Clin. Invest. **87**, 68-76.

Cohen, G.B., Ren, R. and Baltimore, D. 1995. Modular binding domains in signal transduction proteins (review). Cell **80**, 237-248.

Cooper, G.M. 1990. Oncogenes. Jones and Barlett publisher, Boston.

Cooper, J.A., Gould, K.L., Cartwright, C.A. and Hunter, T. 1986. Tyr 527 is phosphorylated in pp60c-src: implications for regulation. Science **231**, 1431-1434.

Courtneidge, S.A., Levinson, A.D. and Bishop, J.M. 1980. The protein encoded by the transforming gene of avian sarcoma virus (pp60src) and a homologous protein in normal cells (pp60 proto-src) are associated with the plasma membrane. Proc. Natl. Acad. Sci. USA **77**, 3783-3787.

Courtneidge, S.A., Dhand, R., Pilat, D., Twamley, G.M., Waterfield, M.D. and Roussel, M.F. 1993. Activation of Src family kinases by colony stimulating factor-1, and their association with its receptor. EMBO J. **12**, 943-950.

Courtneidge, S.A. and Fumagalli, S. 1994. A mitotic function for Src. Trends in Cell Biol. **4**, 345-347.

David-Pfeuty, T. and Nouvian-Dooghe, Y. 1990. Immunolocalization of the cellular Src protein in interphase and mitotic NIH c-src overexpresser cells. J. Cell. Biol. **111**, 3097-3116.

David-Pfeuty, T. and Novian-Dooghe, Y. 1995. Highly specific antibody to Rous sarcoma

virus src gene product recognizes nuclear and nucleolar antigens in human cells. J. Virology **69**, 1699-1713.

De Boulle, K., Verkerk, A.J., Reyniers, E., Vits, L., Hendrickx, J., Van Roy, B., Van den Bos, F., De Graaff, E., Oostra, B.A. and Willems, P.J. 1993. A point mutation in the FMR-1 gene associated with fragile X mental retardation. Nature Genet **3**, 31-35.

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

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**, 377-381.

Endow, S.A., Chandra, R., Komma, D.J., Yamamoto, A.H. and Salmon, E.D. 1994. Mutants of the Drosophila ncd microtubule motor protein cause centrosomal and spindle pole defects in mitosis. J. Cell Sci. **107**, 859-867.

Erlich, R., Gleeson, P.A., Campbell, P., Dietzsch, E. and Toh, B.-H. 1996. Molecular characterization of trans-Golgi p230. J. Biol. Chem. **271**, 8328-8337.

Feng, S., Chen, J.K., Yu, H., Simon, J.A. and Schreiber, S.L. 1994. Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. Science **266**, 1241-1247.

Flynn, P.C., Leu, T.H., Reynolds, A.B. and Parsons, T. 1993. Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated pp60src substrate. Mol. Cell. Biol. **13**, 7892-7900.

Frangioni, J.V. and Neel, B.G. 1993. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. Anal. Biochem. **210**, 179-187.

Fritzler, M.J., Hamel, J.C., Ochs, R.L. and Chan, E.K.L. 1993. Molecular characterization of two human autoantigens: unique cDNAs encoding 95- and 160-kD proteins of a putative family in the Golgi complex. J. Exp. Med. **178**, 49-62.

Fumagalli, S., Totty, N.F., Hsuan, J.J. and Courtneidge, S.A. 1994. A target for Src in mitosis. Nature **368**, 871-874.

Gibson, T.J., Thompson, J.D. and Heringa, J. 1993. The KH domain occurs in a diverse set of RNA-binding proteins that include the anti terminator NusA and is probably involved in binding to nucleic acid. FEBS Lett. **324**, 361-366.

Goding, J.W. 1996. Monoclonal antibodies: principles and practice. 3rd ed. Academic Press, London.

Gould, K. and Hunter, T. 1988. Platelet-derived growth factor induces multisite phosphorylation of pp60c-src and increases its protein-tyrosine kinase activity. Mol. Cell. Biol. **8**, 3345-3356.

Hunter, T. and Cooper, J.A. 1986. Protein control by phosphorylation in The Enzymes (Boyer, P.D. and Krebs, E.G. eds). vol XVII, pp 191-246. Academic Press, New York.

Jans, D.A. and Hübner, S. 1996. Regulation of protein transport to the nucleus: central role of phosphorylation (review). Physiol. Rev. **76**, 651-685.

Kanner, S.B., Reynolds, A.B., Vines, R.R. and Parsons, J.T. 1990. Monoclonal antibodies to individual tyrosine-phosphorylated protein substrates of oncogene-encoded tyrosine

kinases. Proc. Natl. Acad. Sci. USA. **87**, 3328-3332.

Kanner, S.B., Reynolds, A.B., Wang, H.C., Vines, R.R. and Parsons, J.T. 1991. The SH2 and SH3 domains of pp60src direct stable association with tyrosine phosphorylated proteins p130 and p110. EMBO J. **10**, 1689-1698.

Kaplan, K.B., Swedlow, J.R., Varmus, H.E. and Morgan, D.O. 1992. Association of p60c-src with endosomal membranes in mammalian fibroblasts. J. Cell. Biol. **118**, 321-333.

Kaplan, K.B., Bibbins, K.B., Swedlow, J.R., Arnaud, M., Morgan, D.O. and Varmus, H.E. 1994. Association of the amino-terminal half c-Src with focal adhesions alter their properties and is regulated by phosphorylation of tyrosine 527. EMBO J. **13**, 4745-4756.

Kaplan, J.M., Varmus, H.E. and Bishop, J.M. 1990. The Src protein contains multiple domains for specific attachment to membranes. Mol. Cell. Biol. **10**, 1000-1009.

Kappeller, R., Prasad, K.V.S., Janssen, O., Hou, W., Schaffhausen, B.S., Rudd, C.E. and Cantley, L.C. 1994. Identification of two SH3-binding motifs in the regulatory subunit of phosphatidylinositol 3-kinase. J. Biol. Chem. **269**, 1927-1933.

Kmiecik, T.E. and Shalloway, D. 1987. Activation and suppression of pp60c-src transforming ability by mutation of its primary sites of tyrosine phosphorylation. Cell **49**, 65-73.

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

Landschultz, W.H., Johnson, P.F. and McKnight, S.L. 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240**, 1759-

1764.

Lawe, D.C., Hahn, C. and Wong, A.J. 1997. The Nck SH2/SH3 adaptor protein is present in the nucleus and associates with the nuclear protein SAM68. Oncogene. **14**, 223-231.

Levine, T.P., Rabouille, C., Kieckbusch, R.H. and Warren, G. 1996. Binding of the vesicle docking protein p115 to Golgi membranes is inhibited under mitotic conditions. J. Biol. Chem. **271**, 17304-17311.

Liebl, E.C. and Martin, G.S. 1992. Intracellular targeting of pp60src expression: localization of v-src to adhesion plaques is sufficient to transform chicken embryo fibroblasts. Oncogene **7**, 2417-2428.

Lin, Q., Taylor, S.J. and Shalloway, D. 1997. Specificity and determination of Sam68 RNA binding. J. Biol. Chem. **272**, 27274-27280.

Lippincott-Schwartz, J., Yuan, L.C., Bonifacino, J.S. and Klausner, R.D. 1989. Rapid redistribution of Golgi proteins into the ER in cells treated with brefeldin A: evidence for membrane cycling from Golgi to ER. Cell **56**, 801-813.

Liu, X., Marengere, L.E., Koch, C.A. and Pawson, T. 1993. The v-Src SH3 domain binds phosphatidylinositol 3' kinase. Mol. Cell. Biol. **13**, 5225-5232.

Lock, P., Fumagalli, S., Polakis, P., McCormick, F. and Courtneidge, S.A. 1996. The human p62 cDNA encodes Sam68 and not the Ras GAP-associated p62 protein. Cell **84**, 23-24.

Lowell, C.A., Soriano, P. and Varmus, H.E. 1994. Functional overlap in the Src gene family: inactivation of hck and fgr impairs natural immunity. Genes Dev. **8**, 387-398.

Lupas, A., Van Dyke, M. and Stock, J. 1991. Predicting coiled-coils from protein sequences. Science **252**, 1162-1164.

Mayer, B.J. and Eck, M.J. 1995. SH3 domains. Minding up p's and q's (review). Curr. Biol. **5**, 364-367.

McGlade, J., Cheng, A., Pelicci, G., Pelicci, P.G. and Pawson, T. 1992. Shc proteins are phosphorylated and regulated by the v-Src and v-Fps protein-tyrosine kinases. Proc. Natl. Acad. Sci. USA **89**, 8869-8873.

Moran, M.F., Koch, C.A., Anderson, D., Ellis, C., England, L., Martin, G.S. and Pawson, T. 1990. Src homology region 2 domains direct protein-protein interactions in signal transduction. Proc. Natl. Acad. Sci. USA **87**, 8622-8626.

Moran, M.F., Polakis, P., McCormick, F., Pawson, T. and Ellis, C. 1991. Protein tyrosine kinase regulate the phosphorylation protein interactions, subcellular distribution, and activity of p21ras GTPase-activating protein. Mol. Cell. Biol. **11**, 1804-1812.

Morgan, D.O., Kaplan, J.M., Bishop, J.M. and Varmus, H.E. 1989. Mitosis-specific phosphorylation of pp60c-src by p34cdc2-associated protein kinase. Cell **57**, 775-786.

Mori, S., Ronnstrand, L., Yokote, K., Engstrom, A., Courtneidge, S.A., Claesson-Welsh, L. and Heldin, C.H. 1993. Identification of two juxtamembrane autophosphorylation sites in the PDGF beta receptor; involvement in the interaction with Src family tyrosine kinases. EMBO J. **12**, 2257-2264.

Mu, F.T., Callaghan, J.M., Steele-Mortimer, O., Stenmark, H., Parton, R.G., Campbell, P.L., McCluskey, J., Yeo, J.P., Tock, E.P.C. and Toh, B.H. 1995. EEA1, an early endosome-associated protein. EEA1 is a conserved alpha-helical peripheral membrane protein flanked

by cysteine "fingers" and contains a calmodulin-binding IQ motif. J. Biol. Chem. **270**, 13503-13511.

Nakamura, N., Rabouille, C., Watson, R., Nilsson, T., Hui, N., Slusarewicz, P., Kreis, T.E. and Warren, G. 1995. Characterization of a cis-Golgi matrix protein, GM130. J. Cell Biol. **131**, 1715-1726.

Nakamura, N., Lowe, M., Levine, T.P., Rabouille, C. and Warren, G. 1997. The vesicle docking protein p115 binds GM130, a cis-Golgi matrix protein, in a mitotically regulated manner. Cell **89**, 445-455.

O'Connor, T.J., Bjorge, J.D., Cheng, H.C., Wang, J.H. and Fujita, D.J. 1995. Mechanism of c-src activation in human melanocytes: elevated level of protein tyrosine phosphatase activity directed against the carboxyl-terminal regulatory tyrosine. Cell Growth Differ. **6**, 123-130.

O'Shea, E.K., Rutkowski, R. and Kim, P.S. 1989. Evidence that leucine zipper is a coiled-coil. Science **243**, 538-542.

Pawson, T. 1993. Signal transduction- a conserved pathway from the membrane to the nucleus (review). Devel. Genet. **14**, 333-338.

Pawson, T. and Schlessinger, J. 1993. SH2 and SH3 domains. Curr. Biol. **3**, 434-442.

Pharmacia Biotech. 1997. GST gene fusion system manual.

Pillay, I., Nakano, H. and Sharma, S.V. 1996. Radicicol inhibits tyrosine phosphorylation of the mitotic Src substrate Sam68 and retards subsequent exit from mitosis of Src-transformed cells. Cell Growth and Differ. **7**, 1487-1499.

Prasad, K.V.S., Janssen, O., Kapeller, R., Raab, M., Cantley, L.C. and Rudd, C.E. 1993. Src-homology 3 domain of protein kinase p59fyn mediates binding to phosphatidylinositol 3-kinase in T cells. Proc. Natl. Acad. Sci. USA **90**, 7366-7370.

Ralston, R. and Bishop, J.M. 1985. The product of the protooncogene c-src is modified during the cellular response to platelet-derived growth factor. Proc. Natl. Acad. Sci. USA. **82**, 7845-7849.

Rameh, L.E., Chen, C.-S. and Cantley, L.C. 1995. Phosphatidylinositol (3,4,5) P3 interacts with SH2 domains and modulates PI3-kinase association with tyrosine-phosphorylated proteins. Cell **83**, 821-830.

Resnick, R.V., Taylor, S.J., Lin, Q. and Shalloway, D. 1997. Phosphorylation of the Src substrate Sam68 by Cdc2 during mitosis. Oncogene **15**, 1247-1253.

Reynolds, A.B., Kanner, S.B., Wang, H.-C. and Parsons, J.T. 1989. Stable association of activated pp60src with two tyrosine-phosphorylated cellular proteins. Mol. Cell. Biol. **9**, 3951-3958.

Reynolds, A.B., Roesel, D.J., Kanner, S.B. and Parsons, J.T. 1989. Transformation-specific tyrosine phosphorylation of a novel cellular protein in chicken cells expressing oncogenic variants of the avian cellular Src gene. Mol. Cell. Biol. **9**, 629-638.

Richard, S., Yu, D., Blumer, K.J., Hausladen, D., Olszowy, M.W., Connelly, P.A. and Shaw, A.S. 1995. Association of p62, a multifunctional SH2- and SH3-domain-binding protein, with src family tyrosine kinases, Grb2, and phospholipase C gamma-1. Mol. Cell. Biol. **15**, 186-197.

Rodriguez, J.L., Gelpi, C., Thomson, T.M., Real, F.J. and Fernandez, J. 1982. Anti-Golgi

complex autoantibodies in a patient with Sjörgren's syndrome and lymphoma. Clin. Exp. Immunol. **49**, 579-586.

Rohrschneider, L.R. 1979. Immunofluorescence on avian sarcoma virus-transformed cells: localization of the Src gene product. Cell **16**, 11-24.

Rozakis-Adcock, M., McGlade, J., Mbamalu, G., Pelicci, G., Daly, R., Li, W., Batzer, A., Thomas, S., Brugge, J., Pelicci, P.G., Schlessinger, J. and Pawson, T. 1992. Association of the Shc and Grb2/Sem5 SH2-containing proteins is implicated in activation of the Ras pathway by tyrosine kinases. Nature **360**, 689-692.

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

Schaller, M.D., Bouton, A.H., Flynn, D.C. and Parsons, J.T. 1993. Identification and characterization of novel substrates for protein tyrosine kinases (review). Prog. Nucl. Acid Res. and Mol. Biol. **44**, 205-227.

Schaller, M.D., Hildebrand, J.D., Shannon, J.D., Fox, J.W., Vines, R.R. and Parsons, J.T. 1994. Autophosphorylation of the focal adhesion kinase, pp125FAK, directs SH2-dependent binding of pp60src. Mol. Cell. Biol. **14**, 1680-1688.

Schlessinger, J. and Ullrich, A. 1992. Growth factor signaling by receptor tyrosine kinases (review). Neuron **9**, 383-391.

Siomi, H., Siomi, M.C., Nussbaum, R.L. and Dreyfuss, G. 1993. The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. Cell **74**, 291-298.

Siomi, H., Choi, M., Siomi, M.C., Nussbaum, R.L. and Dreyfuss, G. 1994. Essential role for

KH domains in RNA binding : impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. Cell **77**, 33-39.

Songyang, Z., Shoelson, S.E., Chauduri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., Neel. B.G., Birge, R.B., Fajardo, J.E., Chou, M.M., Hanafusa, H., Schaffhausen, B. and Cantley, L.C. 1993. SH2 domains recognize specific phosphopeptide sequences. Cell **72**, 767-778.

Songyang, Z., Carraway, K.L, Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schlessinger, J., Hubbard, S.R., Smith, D.P., Eng, C., Lorenzo, M.J., Ponder, B.A.J., Mayer, B.J. and Cantley, L.C. 1995. Catalytic specificity of protein-tyrosine kinases is critical for selective signaling. Nature **373**, 536-539.

Songyang, Z., Lu, K.P., Kwon, Y.T., Tsai, L.-H., Filhol, O., Cochet, C., Brickey, D.A., Soderling, T.R., Bartleson, C., Graves, D.J., DeMaggio, A.J., Hoekstra, M.F., Blenis, J., Hunter, T. and Cantley, L.C. 1996. A structural basis for substrate spesificities of protein Ser/Thr kinases: primary sequence preference of casein kinases I and II, NIMA, phosphorylase kinase, calmodulin-dependent kinase II, CDK5, and Erk1. Mol. Cell. Biol. **16**, 6486-6493.

Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S.A. and Draetta, G. 1993. Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. EMBO J. **12**, 2625-2634.

Tanaka, A. and Fujita, D.J. 1986. Expression of a molecularly cloned human c-src oncogene by using a replication-competent retroviral vector. Mol. Cell. Biol. **6**, 3900-3909.

Taylor, S.J. and Shalloway, D. 1994. An RNA-binding protein associated with Src through its SH2 and SH3 domains in mitosis. Nature **368**, 867-871.

Taylor, S.J, Anafi, M., Pawson, T. and Shalloway, D. 1995. Functional interaction between c-Src and its mitotic target, Sam68. J. Biol. Chem. **270**, 10120-10124.

Turner, R. and Tjian, R. 1989. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cfos-cjun heterodimers. Science **243**, 1689-1694.

Verkerk, A.J., de Graaff, E., De Boulle, K., Eichler, E.E., Konecki, D.S., Reyniers, E., Manca, A., Poustka, A., Willems, P.J., Nelson, D.L. 1993. Alternative splicing in the fragile X gene FMR1. Hum. Mol. Genet **2**, 399-404.

Vogel, L.B. and Fujita, D.J. 1993. The SH3 domain of p56lck is involved in binding to phosphatidylinositol 3'-kinase from T lymphocytes. Mol. Cell. Biol. **13**, 7408-7417.

Vogel, L.B. and Fujita, D.J. 1995. p70 phosphorylation and binding to p56lck is an early event in interleukin-2-induced onset of all cycle progression in T-lymphocytes. J. Biol. Chem. **270**, 2506-2511.

Wang, L.L., Richard, S. and Shaw, A.S. 1995. p62 association with RNA is regulated by tyrosine phosphorylation. J. Biol. Chem. **270**, 2010-2013.

Waters, M.G., Clary, D.O. and Rothman, J.E. 1992. A novel 115-kD peripheral membrane protein is required for intercisternal transport in the Golgi stack. J. Cell Biol. **118**, 1015-1026.

Wong, G., Müller, O., Clark, R., Conroy, L., Moran, M.F., Polakis, P. and McCormick, F. 1992. Molecular cloning and nucleic acid binding properties of the GAP-associated tyrosine phosphoprotein p62. Cell **69**, 551-558.

Wu, H., Reynolds, A.B., Kanner, S.B., Vines, R.R. and Parsons, J.T. 1991. Identification and

characterization of a novel cytoskeleton-associated pp60src substrate. Mol. Cell. Biol. **11**, 5113-5124.

Yaneva, M., Wen, J., Ayala, A. and Cook, R. 1989. cDNA-derived amino acid sequence of the 86 kDa subunit of the Ku antigen. J. Biol. Chem. **264**, 13407-13411.

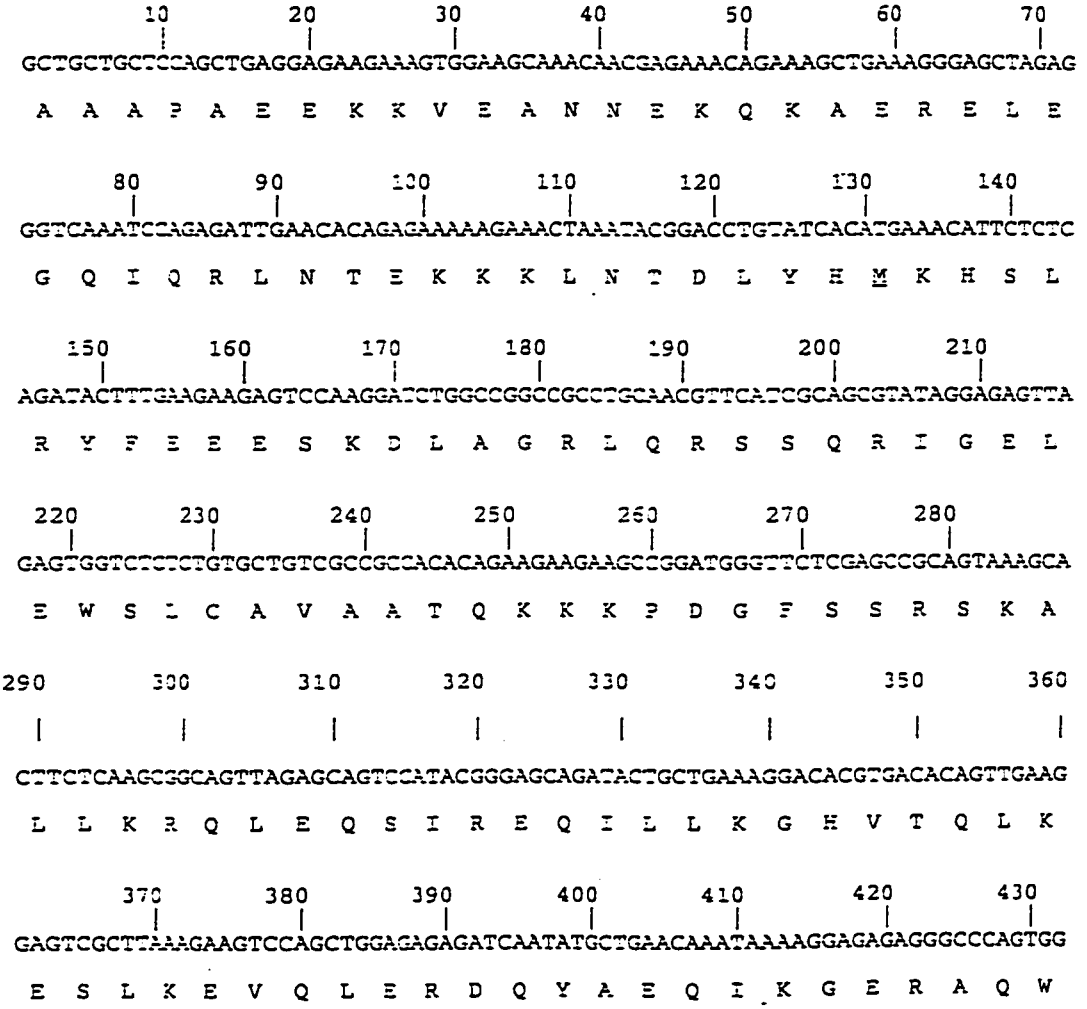
Yu, H., Rosen, M.K., Shin, T.B., Seidel-Dugan, C., Brugge, J.S. and Schreiber, S.L. 1992. Solution structure of the SH3 domain of Src and identification of its ligand-binding site. Science **258**, 1665-1668.

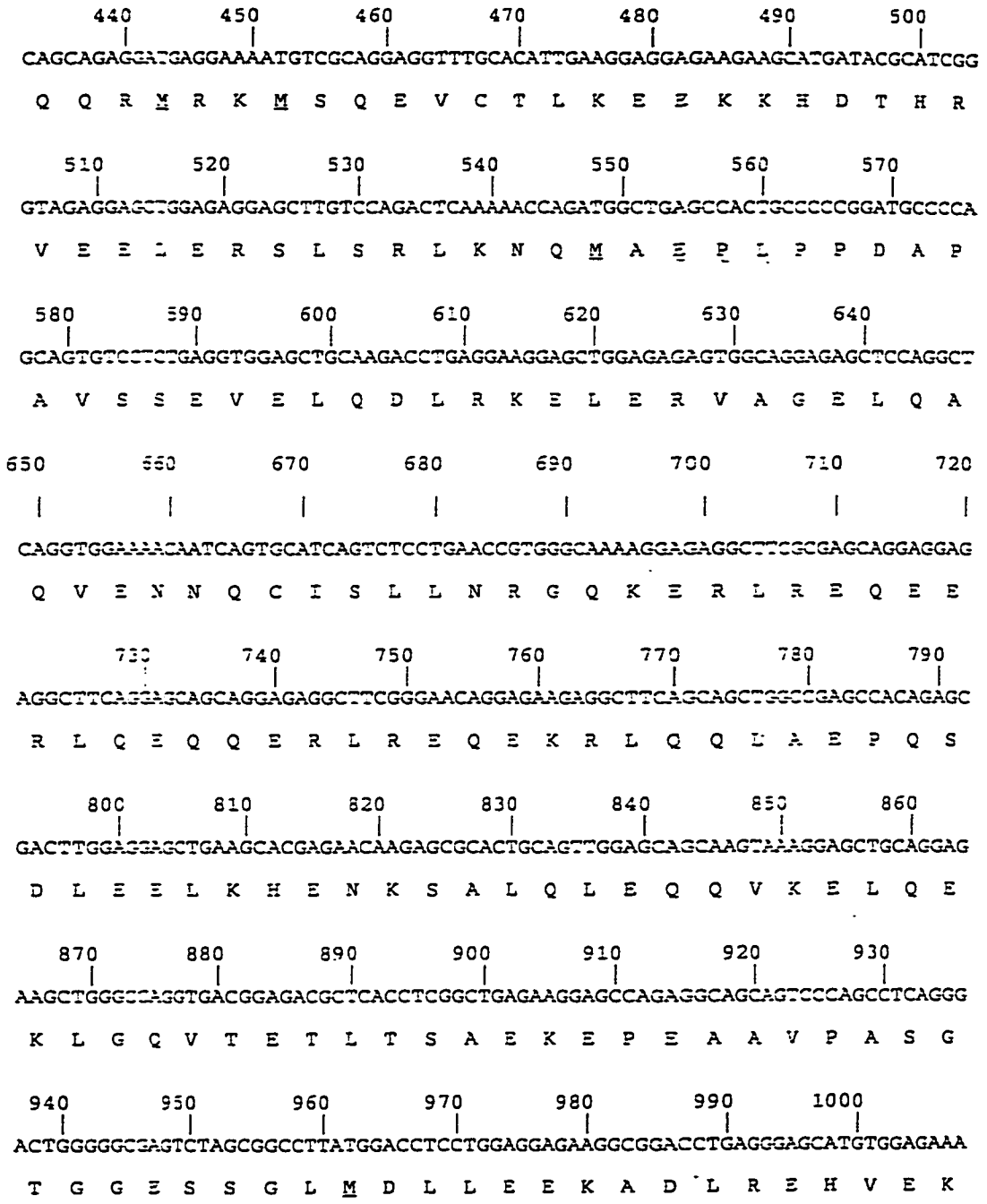
Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W. and Schreiber, S.L. 1994. Structural basis for the binding of proline-rich peptides to SH3 domains. Cell **76**, 933-945.

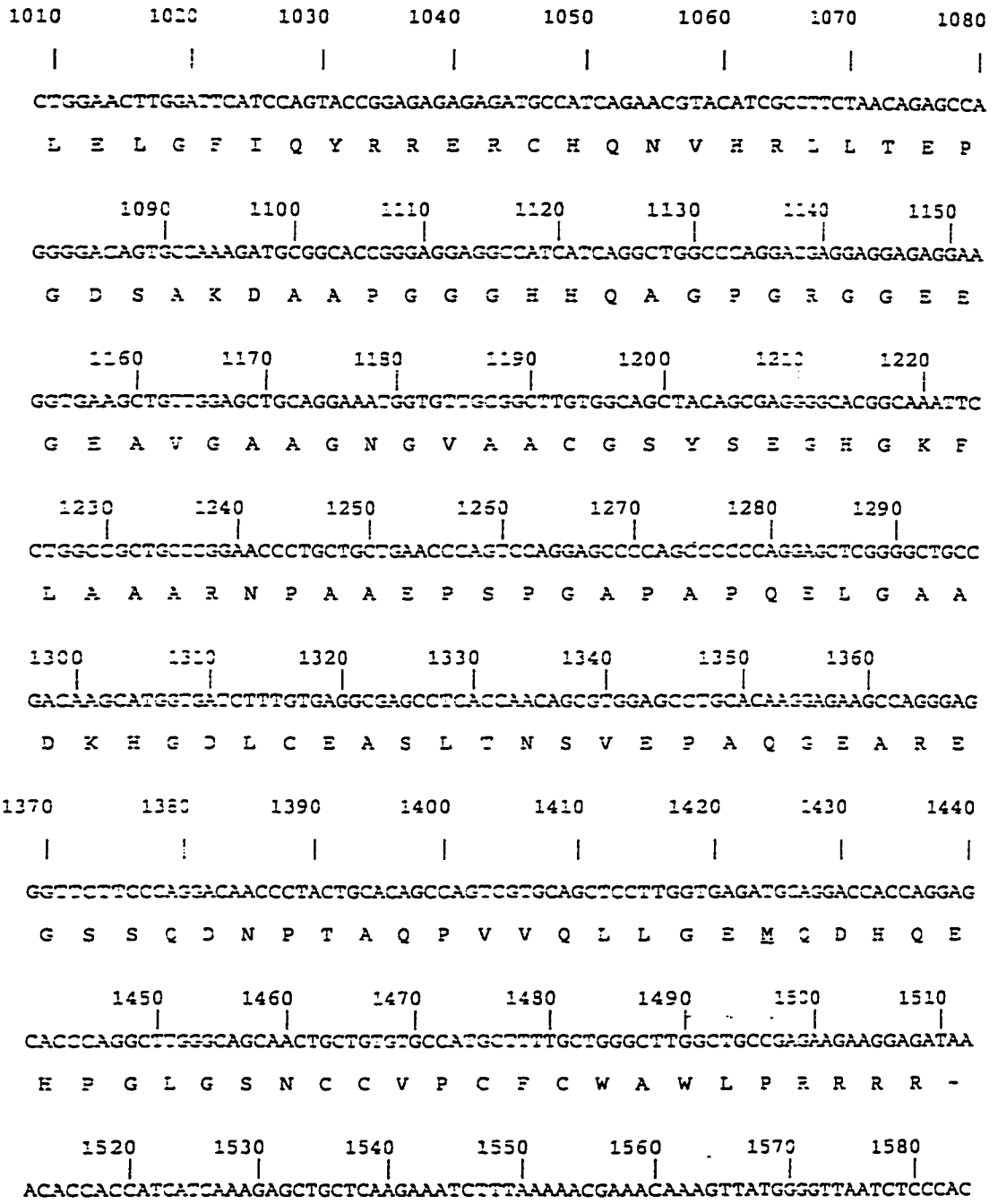
Zheng, X.M., Wang, Y. and Pallen, C.J. 1992. Cell transformation and activation of pp60src by overexpression of a protein tyrosine phosphatase. Nature **359**, 336-339.

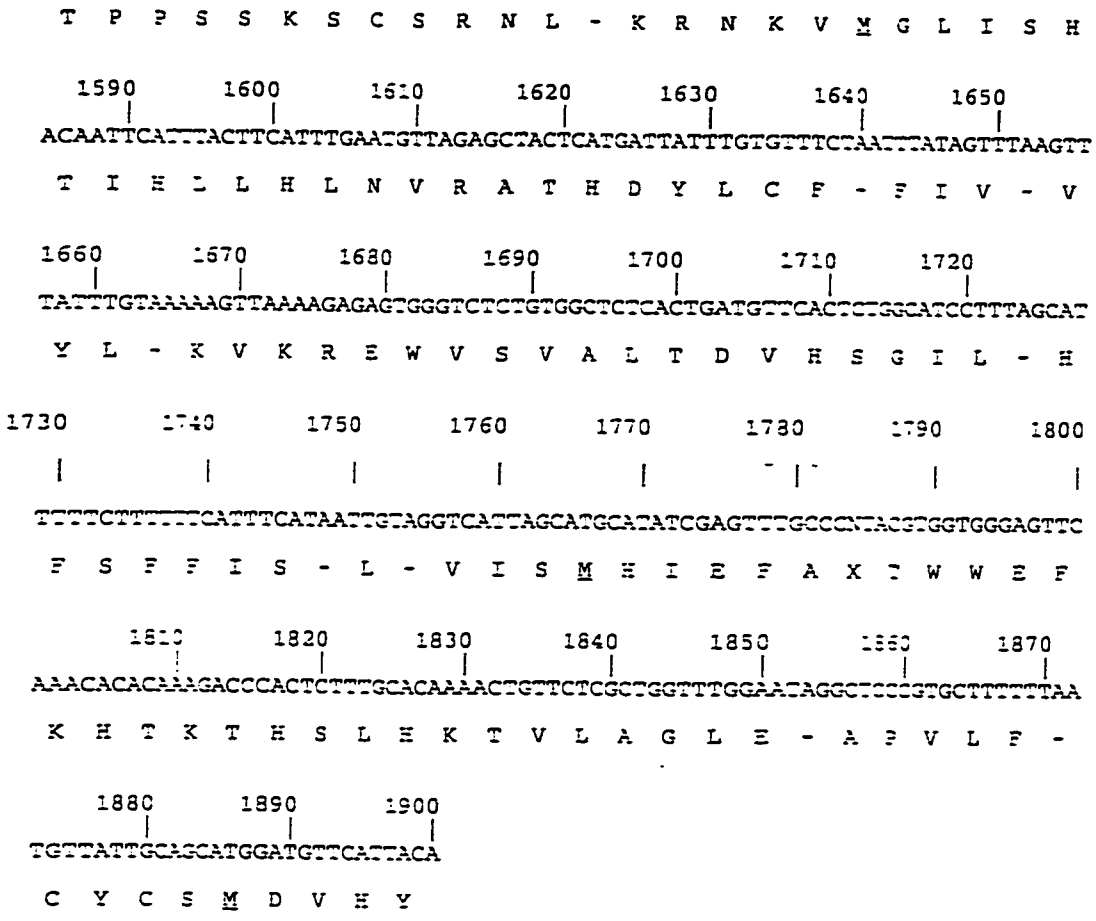
APPENDIX

The complete sequence of C-4.3.

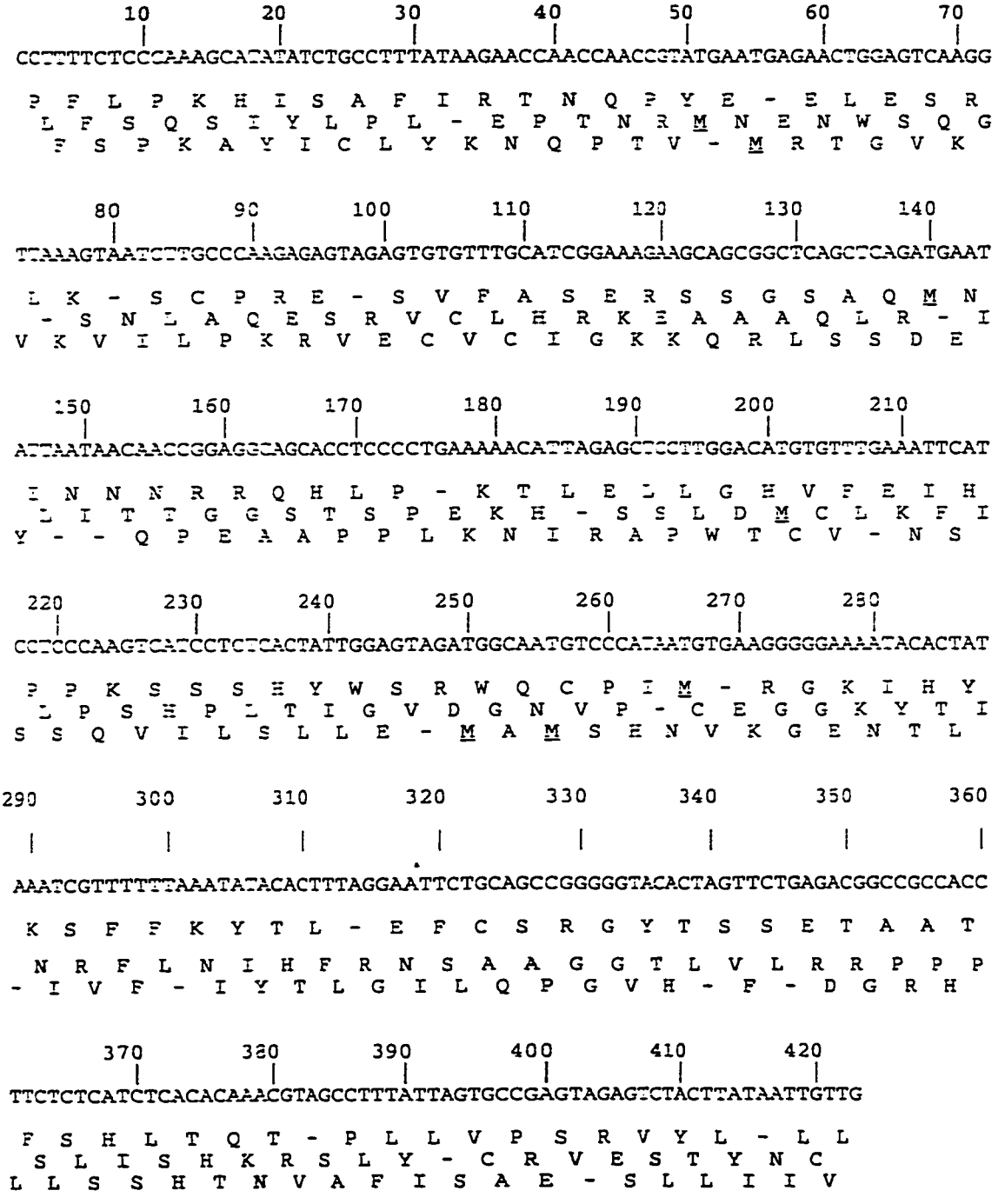



440 450 460 470 480 490 500
CAGCAGAGGATGAGGAAAATGTCGCAGGAGGTTTGCACATTGAAGGAGGAGAAGAAGCATGATACGCATCGG
Q Q R M R K M S Q E V C T L K E E K K H D T H R
510 520 530 540 550 560 570
GTAGAGGAGCTGGAGAGGAGCTTGTCCAGACTCAAAAACCAGATGGCTGAGCCACTGCCCCCGGATGCCCCA
V E E L E R S L S R L K N Q M A E P L P P D A P
580 590 600 610 620 630 640
GCAGTGTCTCTGAGGTGGAGCTGCAAGACCTGAGGAAGGAGCTGGAGAGAGTGGCAGGAGAGCTCCAGGCT
A V S S E V E L Q D L R K E L E R V A G E L Q A
650 660 670 680 690 700 710 720
CAGGTGGAAAACAATCAGTGCATCAGTCTCCTGAACCGTGGGCAAAAGGAGAGGCTTCGCGAGCAGGAGGAG
Q V E N N Q C I S L L N R G Q K E R L R E Q E E
730 740 750 760 770 780 790
AGGCTTCAGGAGCAGCAGGAGAGGCTTCGGGAACAGGAGAAGAGGCTTCAGCAGCTGGCCGAGCCACAGAGC
R L Q E Q Q E R L R E Q E K R L Q Q L A E P Q S
800 810 820 830 840 850 860
GACTTGGAGGAGCTGAAGCACGAGAACAAGAGCGCACTGCAGTTGGAGCAGCAAGTAAAGGAGCTGCAGGAG
D L E E L K H E N K S A L Q L E Q Q V K E L Q E
870 880 890 900 910 920 930
AAGCTGGGGCAGGTGACGGAGACGCTCACCTCGGCTGAGAAGGAGCCAGAGGCAGCAGTCCCAGCCTCAGGG
K L G Q V T E T L T S A E K E P E A A V P A S G
940 950 960 970 980 990 1000
ACTGGGGGGCGAGTCTAGCGGCCTTATGGACCTCCTGGAGGAGAAGGCGGACCTGAGGGAGCATGTGGAGAAA
T G G E S S G L M D L L E E K A D L R E H V E K

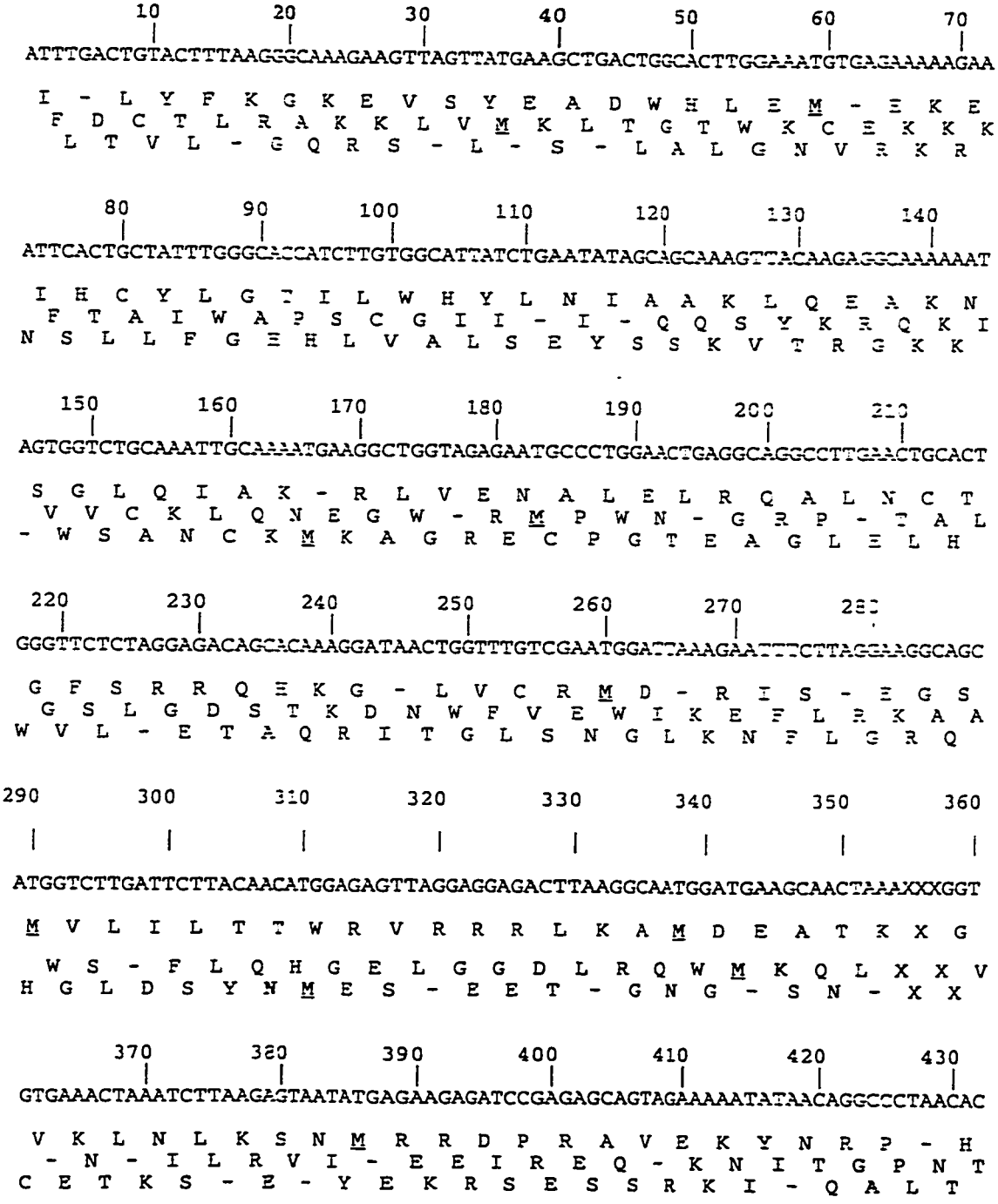

1010 1020 1030 1040 1050 1060 1070 1080
CTGGAACTTGGATTCATCCAGTACCGGAGAGAGAGATGCCATCAGAACGTACATCGCTTCTAACAGAGCCA
L E L G F I Q Y R R E R C H Q N V H R L L T E P
1090 1100 1110 1120 1130 1140 1150
GGGGACAGTGCCAAAGATGCGGCACCGGGAGGAGGCCATCATCAGGCTGGCCCAGGACGAGGAGGAGAGGAA
G D S A K D A A P G G G H H Q A G P G R G G E E
1160 1170 1180 1190 1200 1210 1220
GGTGAAGCTGTTGGAGCTGCAGGAAATGGTGTTGCGGCTTGTGGCAGCTACAGCGAGGGGCACGGCAAATTC
G E A V G A A G N G V A A C G S Y S E G H G K F
1230 1240 1250 1260 1270 1280 1290
CTGGCCGCTGCCCGGAACCCTGCTGCTGAACCCAGTCCAGGAGCCCCAGCCCCCCAGGAGCTCGGGGCTGCC
L A A A R N P A A E P S P G A P A P Q E L G A A
1300 1310 1320 1330 1340 1350 1360
GACAAGCATGGTGATCTTTGTGAGGCGAGCCTCACCAACAGCGTGGAGCCTGCACAAGGAGAAGCCAGGGAG
D K H G D L C E A S L T N S V E P A Q G E A R E
1370 1380 1390 1400 1410 1420 1430 1440
GGTTCTTCCCAGGACAACCCTACTGCACAGCCAGTCGTGCAGCTCCTTGGTGAGATGCAGGACCACCAGGAG
G S S Q D N P T A Q P V V Q L L G E M Q D H Q E
1450 1460 1470 1480 1490 1500 1510
CACCCAGGCTTGGGCAGCAACTGCTGTGTGCCATGCTTTTGCTGGGCTTGGCTGCCGAGAAGAAGGAGATAA
H P G L G S N C C V P C F C W A W L P R R R R -
1520 1530 1540 1550 1560 1570 1580
ACACCACCATCATCAAAGAGCTGCTCAAGAAATCTTTAAAAACGAAACAAAGTTATGGGGTTAATCTCCCAC


T P P S S K S C S R N L - K R N K V M G L I S H
1590 1600 1610 1620 1630 1640 1650
ACAATTCATTTACTTCATTTGAATGTTAGAGCTACTCATGATTATTTGTGTTTCTAATTCATAGTTTAAGTT
T I H L L H L N V R A T H D Y L C F - F I V - V
1660 1670 1680 1690 1700 1710 1720
TATTTGTAAAAAGTTAAAAGAGAGTGGGTCTCTGTGGCTCTCACTGATGTTCACTCTGGCATCCTTTAGCAT
Y L - K V K R E W V S V A L T D V H S G I L - H
1730 1740 1750 1760 1770 1780 1790 1800
TTTTCTTTTTTCATTTCATAATTGTAGGTCATTAGCATGCATATCGAGTTTGCCCNTACGTGGTGGGAGTTC
F S F F I S - L - V I S M H I E F A X T W W E F
1810 1820 1830 1840 1850 1860 1870
AAACACACAAAGACCCACTCTTTGCACAAAACTGTTCTCGCTGGTTTGGAATAGGCTCCCGTGCTTTTTTAA
K H T K T H S L H K T V L A G L E - A P V L F -
1880 1890 1900
TGTTATTGCAGCATGGATGTTCATTACA
C Y C S M D V H Y

Sequence of N-1.1



Sequence of N-1.2



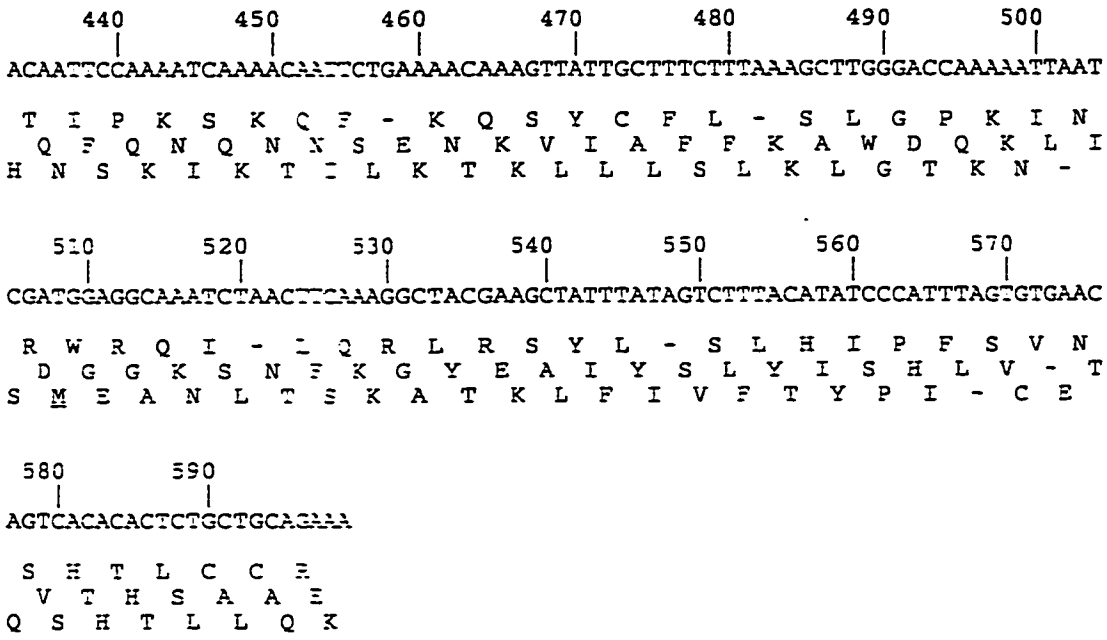
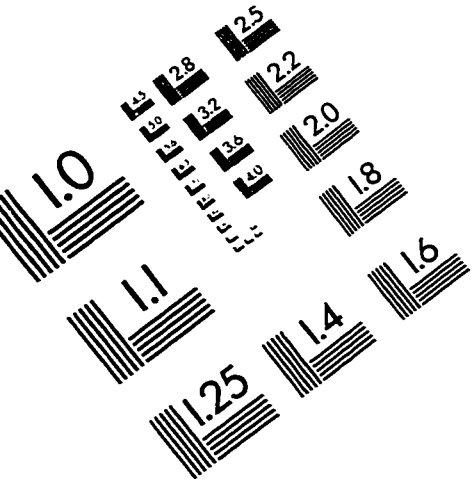
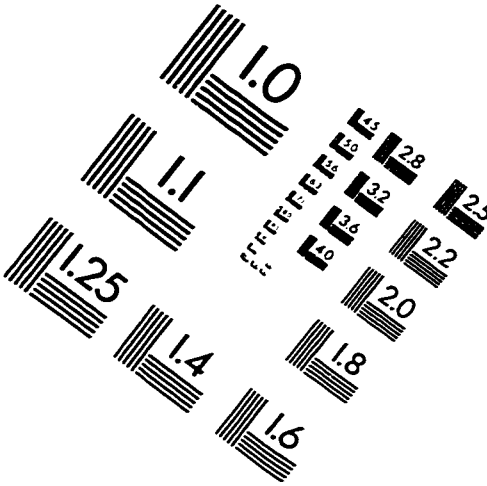
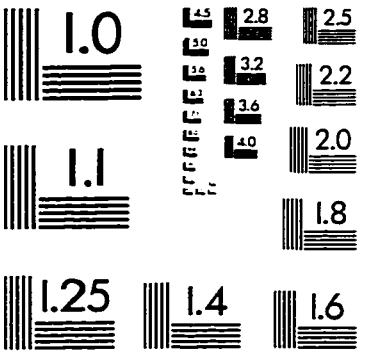
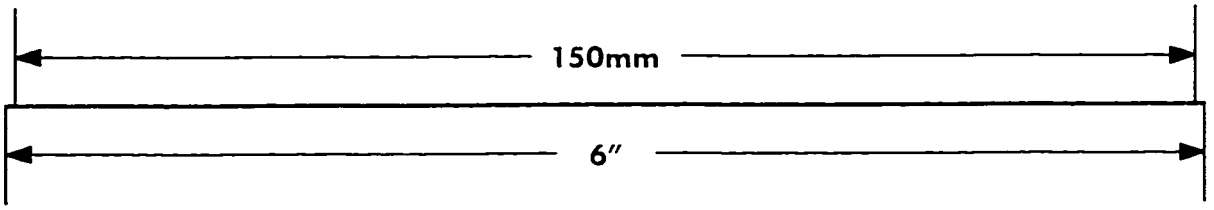

440 450 460 470 480 490 500
ACAATTCCAAAATCAAAACAATTCTGAAAACAAAGTTATTGCTTTCTTTAAAGCTTGGGACCAAAAATTAAT
T I P K S K Q F - K Q S Y C F L - S L G P K I N
Q F Q N Q N X S E N K V I A F F K A W D Q K L I
H N S K I K T I L K T K L L L S L K L G T K N -
510 520 530 540 550 560 570
CGATGGAGGCAAATCTAACTTCAAAGGCTACGAAGCTATTTATAGTCTTTACATATCCCATTTAGTGTGAAC
R W R Q I - I Q R L R S Y L - S L H I P F S V N
D G G K S N F K G Y E A I Y S L Y I S H L V - T
S M E A N L T S K A T K L F I V F T Y P I - C E
580 590
AGTCACACACTCTGCTGCAGAAA
S H T L C C E
V T H S A A E
Q S H T L L Q K

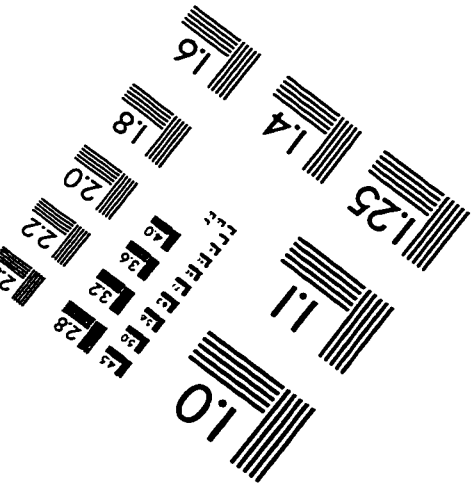
IMAGE EVALUATION TEST TARGET (QA-3)

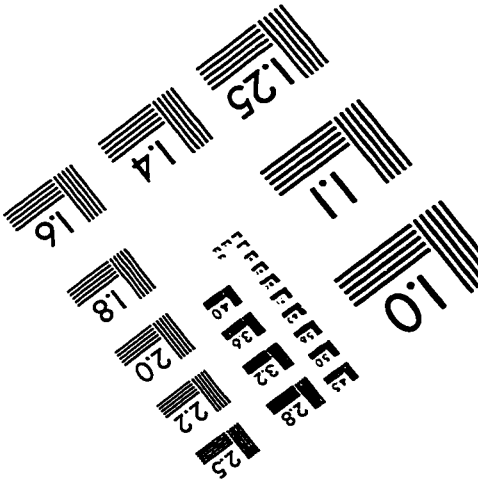












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989