University of Alberta

Structural Aspects of the Interaction of the Cytoplasmic Domain of Mucin-1 (MUC1) with the SH3 Domain of Src Kinase

by

Bodhi Nirosha Marasinghe Arachchige

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Master of Science

in

Medical Sciences - Laboratory Medicine and Pathology

Fall 2011

Edmonton, Alberta

Dedication

This thesis is dedicated to the loving memory of my beloved father,

M.A. Gunasekara and to my beloved mother Leela Gunasekara

Abstract

Breast cancer is the second most frequent cause of cancer deaths in Canadian women with death resulting from the spread of cancer cells or metastasis to distal organs. Our laboratory was the first to show that MUC1, a type-1 transmembrane glycoprotein highly overexpressed in breast tumors, may contribute to migration of breast cancer cells by binding to the Intercellular adhesion molecule-1 (ICAM-1), which triggers the recruitment of non-receptor tyrosine kinase, Src that initiates the downstream signaling. However, the structural aspects of the interaction of cytoplasmic domain of MUC1 (MUC1-CD) and the Src-SH3 domain are still unknown. This thesis, aims to determine the affinity and specificity of this interaction using multinuclear, multidimensional nuclear magnetic resonance (NMR) spectroscopy/titration studies using ^{15}N labeled Src-SH3 domain and the synthetic peptides of MUC1-CD. The results revealed that the dissociation constant (K_D) for the interaction of 69-residue full-length MUC1-CD and Src-SH3 domain is 1.85 mM, based on the residues that show the highest chemical shift changes (> 0.04 ppm). Although the residue-shifts were very small (< 0.1 ppm) different-length MUC1-peptides produced the same results. The most perturbed residues were, Arg98, Glu100, Leu103, His125, Thr132 and Gly130 located outside the canonical binding site, suggesting that MUC1-CD binds with a high specificity but a low affinity to a non-canonical site. The results form a foundation for further structural studies exploring the molecular recognition mechanisms of the MUC1/Src-SH3 interaction.

Acknowledgements

First I would like to thank my supervisor, Dr. Judith Hugh for her guidance, encouragement and understanding that allowed me to design and launch a completely novel study, independently. I am indebted to Dr. Brian Sykes (Department of Biochemistry), a member of my supervisory committee, for teaching me about NMR as well as about life. His cheerful attitude and invaluable advice have tremendously helped me to overcome obstacles. Many thanks go to the other members of my supervisory committee, Dr. David Wishart and Dr. Andrew Shaw, for their guidance.

I would like to offer my sincerest thanks to the current and former members of Dr. Brian Sykes's lab; David Corson for his guidance on protein expression and purification, Dr. Ryan Hoffman for providing guidance on assigning NMR spectra, Dr. Monica Li, Dr. Marta Oleszczuk, Dr. Olivier Julian, Ian Robertson, Robert Boyko and Mellisa Crane for educating me about conducting NMR experiments. My special thanks goes to Dr. Peter Hwang for providing invaluable guidance on interpreting NMR data, Dr. Tony Pawson (University of Toronto) and Gary Ritzel (University of Alberta) for donating Src plasmids and PGEX vectors. Many thanks to Dr. Jing Zhang, Maxine Farr-Jones (Dr. Hugh's Lab - former), Rayan Anderson (Department of Biochemistry), Dr. Jodi Carter (University of Alberta Hospital) Dr. Darrin Lindhout (Dr. Sykes's lab-former) for their help and advice on molecular cloning and protein expression.

Last but not the least, I am forever grateful to my beloved mother and late fater, Leela Gunasekara and M.A. Gunasekara, my husband, Harris and my little son, Harrin, who have been the inspiration of my life.

Table of Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

ADAM17	–	ADAM metallopeptidase domain 17
APC	–	Adenomatous Polyposis Coli
AKT/PKB	–	Protein Kinase B
bFGF	–	basic Fibroblast Growth Factor
Bcl-2	–	B-cell lymphoma-2
BME	–	Beta Mercaptoethanol
BMRB	–	Biological Magnetic Resonance data Bank
BRCA 1/2	–	BReast CAncer type 1 (or 2) susceptibility genes
CAS	–	Crk Associated Substrate
CCR7	–	C-C chemokine Receptor type 7
CCL21	–	chemokine (C-C motif) Ligand
CCND1	–	Cyclin-D1
CDC42	–	Cell Division Control protein 42 homolog
CD8	–	Cluster of Differentiation 8
CHEK2	–	CHK2 Checkpoint Homolog
CrkL	–	CT10 regulator of kinase like
Csk	–	C-terminal src kinase
CTC	–	Circulating Tumor Cell
CXCR4	–	C-X-C chemokine Receptor type 4
CXCL12	–	Chemokine (C-X-C motif) Ligand 12
cDNA	–	Complementary DNA
DCIS	–	Ductal Carcinoma In Situ
Dock180	–	Dedicator of cytokinesis -180 kDa
DNA	–	DeoxyriboNucleic Acid
DTT	–	Dithiothreitol
E-cadherin	–	Epithelial cadherin
ECM	–	Extra Cellular Matrix
EGFR	–	Epidermal Growth Factor Receptor
EMA	–	Epithelial Membrane Antigen
EMT	–	Epithelial Mesenchymal Transition

ErbB2/3/4/	–	Erythroblastic leukemia viral oncogene homologs
ER	–	Endoplasmic Reticulum
ER-Positive	–	Estrogen Receptor Positive
ER- Negative	–	Estrogen Receptor Negative
ERK-1/2	–	Extracellular signal-Regulated Kinase
E.coli	–	*Escherichia coli*
ETA	–	Epithelial Tumor Antigen
FAK	–	Focal Adhesion Kinase
FS	–	Fluorescence Spectroscopy
FGFR3	–	Fibroblast Growth Factor-3
GEF	–	Guanine nucleotide Exchange Factor
GSK-3 β	–	Glycogen Synthase Kinase-3-beta
GST	–	Glutathione S-Transferase
GS4B	–	Glutathione Sepharose 4B
Grb2	–	Growth factor receptor-bound protein 2
GTP	–	Guanosine-5'-TriPhosphate
GTPase	–	Guanosine-5'-TriPhosphate hydrolase
HCK	–	Hemopoietic Cell Kinase
HER2	–	Human Epidermal growth factor Receptor 2
HIF	–	Hypoxia Inducible Factor
HGF	–	Hepatocyte Growth Factor
HIV-nef	–	Human Immunodeficiency Virus-negative regulatory factor
hmT	–	hamster middle T antigen
HRG	–	Heregulin
HSQC	–	Heteronuclear Single Quantum Coherence
ICAM-1	–	Intercellular Adhesion Molecule 1
IDC	–	Invasive Ductal Carcinoma
IL-8	–	Interleukin-8
ILC	–	Invasive Lobular Carcinoma
IFN-γ	–	Interferon-gamma
IKK	–	IκB kinase

IPTG	–	IsoPropyl-β-D-ThioGalactopyrenoside
ITC	–	Isothermal Titration Calorimetry
JNK	–	c-Jun N-terminal kinases
K_d	–	Dissociation Constant
LCIS	–	Lobular Carcinoma In Situ
Lck	–	Lymphocyte-specific protein tyrosine kinase
LFA-1	–	Leukocyte Function-associated Antigen-1
Mac-1	–	Macrophage adhesion ligand-1
MALDI-TOF	–	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MAPK	–	Mitogen Activated Protein Kinase
MCA	–	Mammary Carcinoma Antigen
MCF-7	–	Human breast adenocarcinoma cell line
MEK-1	–	Mitogen-activated protein kinase kinase 1
MLCK	–	Myosin Light-Chain Kinase
MMP	–	Matrix Metalloproteinase
MMTV-PyV-MT	–	Mouse Mammary Tumor Virus-driven PolyomaVirus Middle T-antigen transgenic mice
MT1-MMP	–	Membrane Type 1 Matrix Metalloproteinase
MUC1	–	Mucin-1
MUC1-CD	–	MUC1-Cytoplasmic Domain
MUC1-ECD	–	MUC1-ExtraCellular Domain
MUC1-TMD	–	MUC1-TransMembrane Domain
NF-κB	–	Nuclear Factor kappa-light-chain-enhancer of activated B cells
N-cadherin	–	Neural cadherin
NMR	–	Nuclear Magnetic Resonance
NOESY	–	Nuclear Overhauser Effect Spectroscopy
NSCLC	–	Non Small Cell Lung Cancer
p21	–	cyclin-dependent kinase inhibitor 1
p53	–	tumor protein 53
pp60c-Src	–	proto-oncogene protein-60KDa-cellular Src

PDGFRβ	–	Platelet Derived Growth Factor Receptor- beta
PEM	–	Polymorphic Epithelial Mucin
PI3K	–	Phosphoinositide 3-Kinase
PKCδ	–	Protein Kinase C-delta
PLCγ	–	PhosphoLipase-C gamma
PMSF	–	Phenyl-methane-sulfonyl-fluoride
PPII	–	PolyProline type II
ppm	–	parts per million
PTEN	–	Phosphatase and Tensin homolog
PX Domain	–	Phosphoinositide-binding Domain
Rac	–	Ras-related C3 botulinum toxin substrate
Rb	–	Retinoblastoma protein
RhoGTPases	–	Ras homology (Rho) family of GTPases
RTK	–	Receptor Tyrosine Kinase
SDSPAGE	–	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
SEA	–	Sea Urchin Sperm Protein Enterokinase and Agrin
SEC	–	Size Exclusion Chromatography
SFK	–	Src Family Kinase
SH2	–	Src Homology 2
SH3	–	Src Homology 3
SOS	–	Son of Sevenless
SPR	–	Surface Plasmon Resonance
STAT3	–	Signal Transducer and Activator of Transcription 3
T47D	–	Human ductal breast epithelial tumor cell line
TACE	–	Tumor necrosis factor-alpha-Converting Enzyme
TDLU	–	Terminal Ductal Lobular Units
TGF-β	–	Transforming Growth Factor-beta
TNF-α	–	Tumor Necrosis Factor-alpha
TOCSY	–	Total Correlation Spectroscopy
uPA	–	urokinase Plasminogen Activator
uPAR	–	uPA Receptor

VEGF	–	Vascular Endothelial Growth Factor
VNMRJ	–	Varian-NMR-Java-based data acquisition and processing tool
VNTR	–	Variable Number of Tandem Repeats
Wnt-1	–	Wingless-type-1

Chapter 1

INTRODUCTION

1.1. Background

Cancer arises from the uncontrolled growth of abnormal or genetically mutated cells in a particular organ or a tissue in the body. Breast cancer is the most frequently diagnosed cancer in women, with over 23,000 new cases diagnosed in 2010 in Canada and continues to be the second leading cause of cancer deaths in Canadian women (1). The spread of cancer cells, or metastasis, to distal organs such as the brain and bones is the major cause of death due to breast cancer. Development of metastasis depends on the ability of cancer cells to detach from the primary tumor, travel through the circulatory system and establish secondary colonies in distant sites, a process which involves multiple molecular mechanisms (2).

Although advances in diagnosis and treatment, over the recent years, have contributed to the reduced rates of mortality caused by breast cancer, metastasis still results in treatment failures and deaths, suggesting that there is a shortage of pertinent knowledge about the molecular mechanisms of breast cancer metastasis. Due these gaps in knowledge, current prognostic criteria are unable to predict the exact metastasis-risk for an individual breast cancer patient (3). As a result, some women may receive systemic adjuvant therapies such as cytotoxic chemotherapy unnecessarily while other women may be inappropriately classified as low-risk individuals, and excluded from therapy.

Therefore, it is absolutely necessary to launch in-depth investigations of the molecular recognition events involved in breast cancer metastasis, to completely understand the underlying mechanisms, which could provide the pragmatic bridge between the bench and bedside.

1.2. The Normal Breast

The formation of the breast begins in the embryo as a thickening or a ridge of tissue that gives rise to rudimentary milk-carrying (lactiferous) ducts by the end of gestation (4). The human breast remains relatively undeveloped until puberty but undergoes morphological and functional changes during puberty and then during pregnancy and lactation (5). The elongation of milk ducts is stimulated by estrogen, growth hormone, insulin-like growth factor-1 and epidermal growth factor while prolactin, progesterone and thyroid hormones are involved in further ductal branching and formation of alveoli or acini (secretory terminal end buds) (4).

The mature human breast is composed of fatty (adipose) tissue, connective tissue (stromal matrix) and glandular tissue (ducts and lobules) (Fig. 1-A). Each breast consists of 15 to 25 lobes, each connected to a lactiferous duct that carries milk to the nipple. Every lobe is formed by 20 – 40 lobules that drain milk into the intralobular ducts and each lobule consists of a variable number (usually 10-100) of alveoli (6). The alveoli and the intralobular ducts together form the major structural and functional unit of the breast, known as the terminal ductal lobular unit (TDLU) (Fig. 1-iii), which has somatic stem cell activity for further development and differentiation during pregnancy and lactation (7).

The epithelial cells form a continuous lining of the lactiferous ducts, interlobular ducts, intralobular ducts and alveoli. Each TDLU consists of three layers, a luminal epithelium (epithelial cells that cover the surface of lumina or cavities), a basal myoepithelium (contractile epithelial cells located within a glandular epithelium) and a basement membrane that separates the TDLUs from

the intralobular matrix (5). The stromal matrix is comprised of fibroblasts, adipocytes, endothelial cells and nerve cells. The cross-talk between epithelial cells and stromal cells is crucial for the proper function of the breast, failure of which may lead to breast cancer (8).

1.3. The Cancerous Breast

1.3.1. Types of Breast Cancer

Breast cancer is a heterogeneous and phenotypically diverse disease (9). More than 95% of breast cancers are carcinomas, which are malignant tumors arising from the epithelial stem cells of the ducts or lobules (10) and histologically subdivided as ductal carcinomas (DC) or lobular carcinomas (LC), respectively. The *in situ* (IS) carcinomas that are limited to the ducts or lobules and have not penetrated the basement membrane are known as ductal carcinoma *in situ* (DCIS) (Fig. 1-B) or lobular carcinoma *in situ* (LCIS) respectively. If the malignant cells have penetrated the basement membrane into the surrounding stroma, the tumors are called invasive ductal (IDC) (Fig. 1-C) or invasive lobular (ILC) carcinoma, which may be capable of spreading to the distal organs or tissues via the blood vessels and the lymphatic system. Invasive ductal carcinoma is the most common type that comprises 80 – 95 % of all breast cancers (11).

Based on the gene expression profiling, breast cancers can be divided into a limited number of molecular subtypes (12), (13), These subtypes have distinct biological characteristics and response to therapy (14). There are 2 major genetic variants of breast cancer, estrogen receptor (ER)-positive and ER-negative, which are classified based on the level of expression of Estrogen receptor alpha (15).

The ER positive tumors can be further subdivided into Luminal A type tumors and Luminal B type tumors, where the former is associated with a better prognosis than the latter (16), (13). The ER negative breast tumors include three subgroups; i) HER2 positive, ii) normal breast-like and iii) basal-like tumors (17). Overexpression of HER2 (ErBb2 or human epidermal growth factor receptor-2) defines the HER2 positive subtype. The normal breast-like phenotype has an expression profile that is similar to non-cancerous breast tissue while the basal-like group consists of tumors that show a more aggressive clinical behavior and is associated with a poor prognosis (18). A percentage of HER2 positive tumors, however, may also be ER positive, while some Luminal B tumors can be ER negative (17).

Recurrence patterns vary between the ER negative and positive tumors suggesting different metastatic behaviors. The ER negative tumors that recur after surgery do so within the first three years and can be rapidly fatal (19), (20). Approximately 30-40% of all breast cancers are ER positive, Luminal B tumors and patients with these tumors experience a slow but steady increase in recurrence over 10 years or more (21).

The ER positive luminal B tumors show increased expression of the transmembrane mucin-1, commonly known as MUC1 (22), which has been widely studied and recognized as an oncogenic molecule (recently reviewed in (23)). It has been demonstrated that MUC1 is directly involved in the *in vivo* transformation of the mammary gland; the prolonged expression of MUC1 is shown to be associated with the mammary gland tumorigenesis and alveolar differentiation (24).

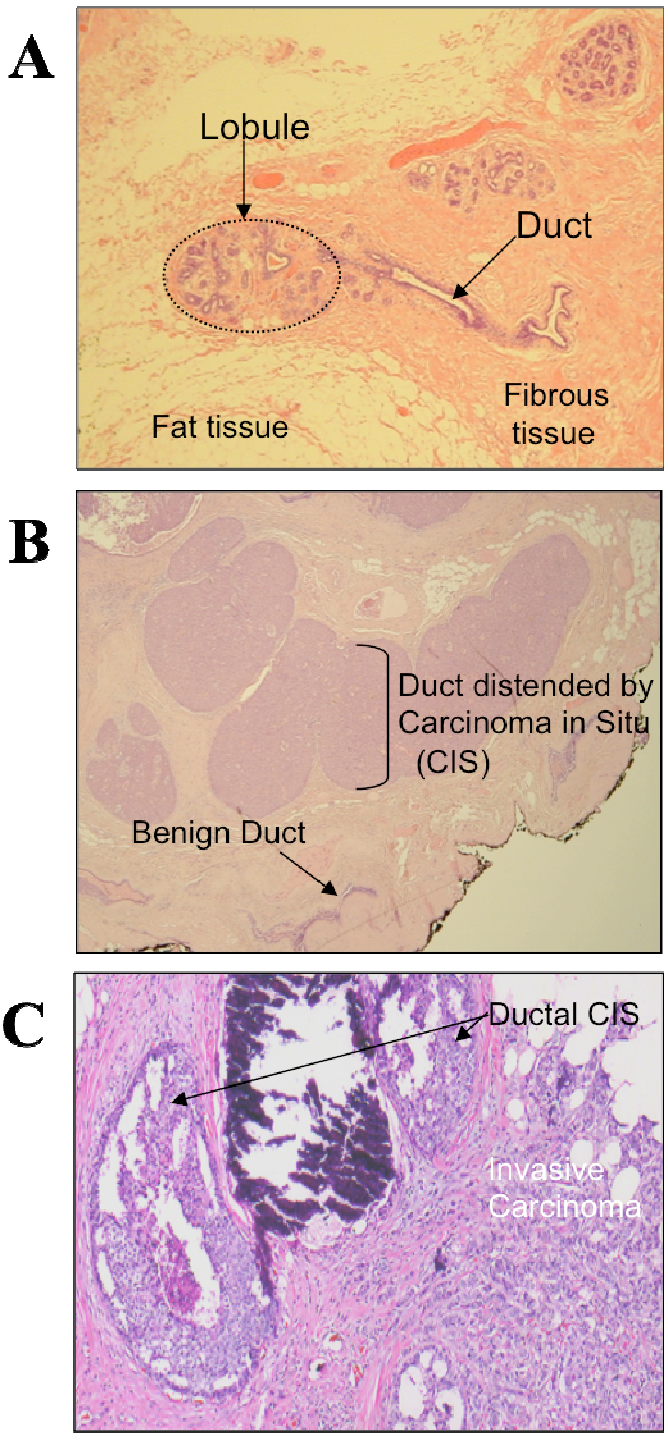


Figure 1.1. Histology of breast cancer subtypes. Photomicrographs of (A) Normal breast (B) Ductal carcinoma *in situ*. (C) Invasive ductal carcinoma. Samples were stained with Hematoxylin and Eosin. Pictures courtesy of J. Hugh.

1.3.2. The Molecular basis of Breast Carcinogenesis

The molecular etiology of breast cancer is extremely complex (13). Genomic analyses of breast cancers indicate that there are only a few genes, which are frequently mutated but many are infrequently mutated, providing an explanation for the observed cancer heterogeneity (25), (26). Like other solid tumors, breast cancer also follows the multistep evolution of a cancer or "multi-hit hypothesis", which postulates that cancer originates from gene mutations occurring in a single cell or a few cells that eventually accumulate and lead to uncontrolled proliferation of a population of cells (10). A healthy breast maintains the balance of normal growth of cells and programmed cell death (apoptosis) by genetically controlled cell-cycle mechanisms and apoptotic pathways. Activated tumor suppressor genes guard these mechanisms while the proto-oncogenes are tightly deregulated. The genetic alterations in the cells however, lead to the inactivation of tumor-suppressor genes and/or activation of proto-oncogenes that promote the arrest of cell cycle control and facilitate the cells to escape from apoptosis (27). As a result, the malignant cells may become self-sufficient in growth signaling and may gain a limitless potential to replicate and invasive abilities to metastasize (28).

Of all breast cancers, 5 - 10% are hereditary, which are characterized by an inherited susceptibility to breast cancer on the basis of a germline mutation in one allele of a high penetrance susceptibility genes (tumor suppressor genes), such as BRCA-1, BRCA-2, p53, PTEN, CHEK2 (29), (30), (31), (32), (33). The hereditary susceptibility for breast cancer has been estimated as; 30-40 % of BRCA1, 10-30 % of BRCA2, 5% of CHEK2, less than 1 % of p53, less than 1 %

of PTEN and one third of unknown mutations (34), (33). The normal BRCA1 protein is involved in repairing the DNA double strand breaks and cell cycle control via dephosphorylation of the retinoblastoma (Rb) protein as well as in the transcriptional regulation of the tumor suppressor proteins, p53 and p21 (35). The BRCA2 protein binds to and regulates the protein produced by the RAD51 gene to fix the breaks in DNA (36). The CHEK2 is one of the most recently identified breast cancer susceptibility gene, which encodes a cell-cycle checkpoint kinase and is implicated in DNA repair processes involving BRCA1 and p53 (37). The mutations in these genes may thus usurp cell cycle control and eventually lead to oncogenic transformation of cells.

More than 90% of all breast cancers are "sporadic" that are thought to occur from mutations that accumulate in the somatic cells (non-germ cells) (33). Activation or amplification of mutated genes or oncogenes such as HER2 (ErbB-2), c-Myc, CCND1 (Cyclin D1) and EGFR (Epidermal growth factor receptor) accounts for early sporadic tumorigenesis in the breast (33), (38), (39), (40), (41).

Deregulation of growth factor signaling pathways and hormones play a major role in breast carcinogenesis (42). The epidermal growth factor (EGF) (43), transforming growth factor beta (TGF-β) (44), insulin-like growth factor (IGF) (45) and hepatocyte growth factor (HGF) (46) are commonly known to contribute to the normal development of the ducts and lobules of the mammary gland and thus are frequently implicated in tumor cell proliferation, epithelial-mesenchymal transition (EMT) and/or anti-apoptotic signaling. Inside the cell, the Src family non-receptor tyrosine kinases (SFKs), particularly the cellular Src (c-Src) contributes to growth, proliferation, invasion, migration and angiogenesis of

breast cancer via several oncogenic signaling pathways (discussed in section 1.7).

The female hormone, estrogen also plays a critical role in the development of breast cancer, since it stimulates the growth of normal breast tissue that eventually convert to the cancerous tissue (47). Therefore, the breast cancer risk among women is associated with the lifetime exposure to estrogen, which depends on factors such as early menarche, late menopause, late first full-term pregnancy or no full-term pregnancies, breast-feeding deficiency and hormonal therapy. Estrogen regulates gene expression via the Estrogen Receptors (ER), that activate oncogenic signaling molecules and contribute to oncogenic cell signaling pathways (15). As introduced in the previous section, the MUC1 molecule is overexpressed in ER positive tumors (22), due to transcriptional upregulation of the MUC1 gene by the transcription factor, ERα (48). Numerous studies report the involvement of MUC1 in adhesion, proliferation, invasion and metastasis of breast cancer cells (23), thus characterizing MUC1 as an oncoprotein.

1.4. Metastasis of Breast Cancer Cells

Spread of tumor cells to other organs or metastasis is the leading cause of mortality of breast cancer patients. The process of metastasis involves a series of steps (Fig. 1.2) that tumor cells must complete to exit the primary tumor and develop a new tumor at a distant site (49), each of which can be a rate-limiting step. The major steps of a metastatic cascade are, i) detachment from the primary tumor mass, penetrating the basement membrane and invading the stroma, ii) intravasation of tumor cells into lymphatic and/or blood circulation systems, iii) adhesion to the endothelial cells and formation of tumor cell emboli and iv)

extravasation into a new tissue, and establishment of new tumor growth by formation of new blood vessels (2).

To invade the stroma, the tumor cells must first detach from the cell - cell and cell – extracellular matrix (ECM) junctions and degrade the basement membrane and extracellular matrix (ECM) (50). The degradation of ECM is mediated by matrix metalloproteinases (MMPs) (51) and the urokinase plasminogen activator (uPA) system (52), in which uPA acts as a substrate to uPAR (uPA receptor) and initiates a proteolytic cascade that aid the degradation.

In epithelial cancers, normal cell-cell junctions involve the E-Cadherin-β catenin complex. Downregulation of E-cadherin and β catenin as well as upregulation of the mesenchymal marker, N-cadherin are closely associated with the process of invasion (49). This leads to a loss of the epithelial cell-cell adhesion and gain of mesenchymal characteristics that convert the tumor cells into migratory and invasive cells, commonly known as epithelial mesenchymal transition (EMT). Loss of E-cadherin and reduced cell-cell adherent junctions have been observed in invasive lobular carcinomas in the breast (53). Unlike lobular carcinomas, which do not frequently express other EMT markers, basal-like breast tumors show a coordinated expression of EMT markers (e.g. vimentin, N-cadherin) in addition to the reduction of epithelial markers (e.g. E-cadherin, luminal cytokeratins) (54). The stromal fibroblasts and myofibroblasts are shown to modulate invasion and migration of transformed epithelial cells (55).

Several molecular pathways contribute to tumor cell invasion and migration (56). Among the molecules that are involved in these pathways, the Src family non-receptor tyrosine kinases (SFK) stand out since their activation

initiates critical downstream signaling pathways (57), (58). For instance, in cancer cells, Src kinase is activated by Focal adhesion kinase (FAK) and forms a Src/FAK signaling complex that initiate tumor cell migration and invasion, involving oncogenic signaling molecules such as, PI3K, p130CAS, Crk, Dock180, RhoGTPases (e.g. Rac, CDC42) (59). Cell migration is a multistep process initiated by the protrusions of actin cytoskeleton of the membrane of invading cells, commonly known as filopodia, lamellipodia or invadopodia (podosomes) (60). These membrane protrusions are dynamic structures, of which polarized extension-contraction cycles coupled with adhesion and de-adhesion, facilitate migration of the cell (56). Both EMT and the activity of RhoGTPases lead to activation of MMPs with further degradation of ECM (49).

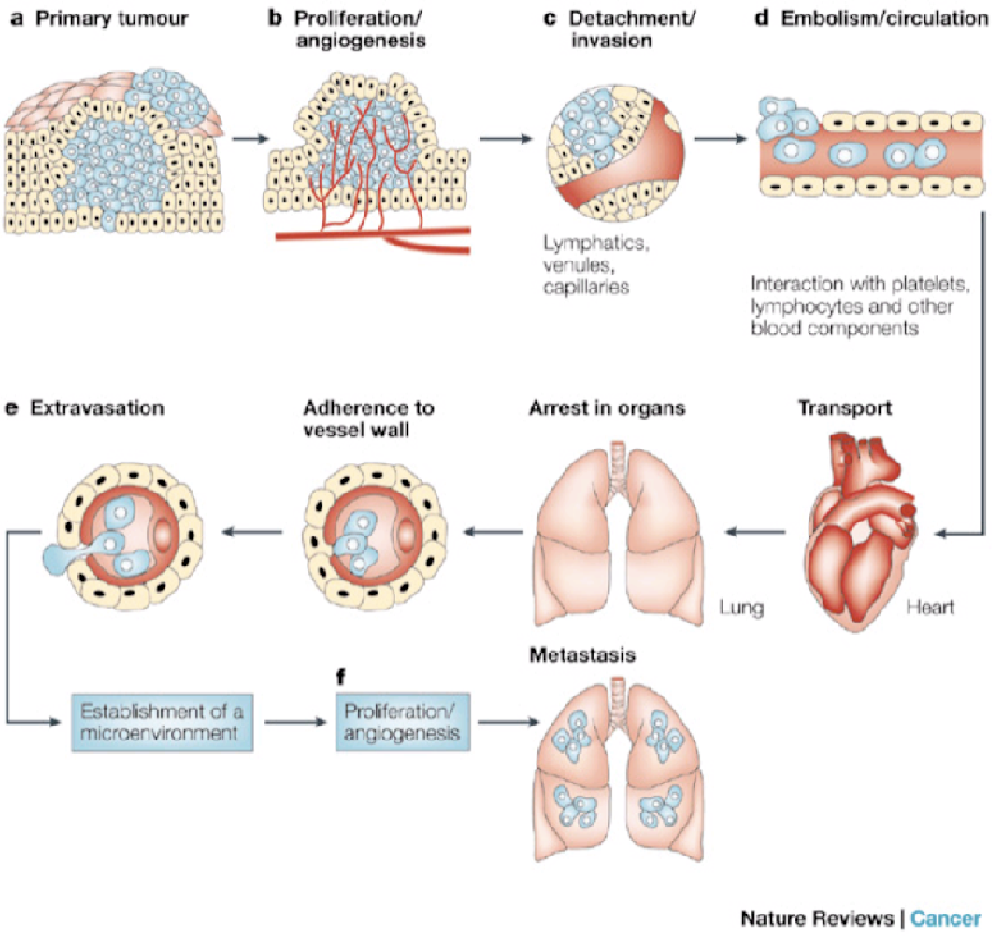


Figure 1.2. The metastatic cascade of a cancer cell showing the different stages of metastasis (adapted from (2)). The major steps are, i) detachment from the primary tumor mass, penetration of the basement membrane and invasion of the stroma (*a, b, c*), ii) intravasation of tumor cells into the blood circulation system (*c, d*), iii) adhesion to the endothelial cells and formation of a tumor cell embolus (*d*) and iv) extravasation into a new tissue, and establishment of new tumor growth by formation of new blood vessels (*e, f*).

Those cells that have gained the invasive/migratory properties can then detach from the primary tumor and intravasate into the lymphatic and/or vascular system and become circulating tumor cells (CTCs) that act as seeds to form tumor emboli. Certain CTCs then extravasate or depart from the blood/lymphatic system, through the vascular endothelium and develop subsequent growth in distal tissues.

The fates of CTCs can vary according to their molecular profile and the micro-environment of the primary tumor and the host organ (61). These cells must escape immune surveillance, avoid programmed cell death, anchorage-dependent cell death or anoikis (cell death that occurs when the attachment between the cell and the ECM is lost), be highly efficient at embolizing to survive in the circulation of lymphatic or blood vessels, be able to extravasate into a new tissue and form new blood vessels (angiogenesis). Thus, CTCs must express several signaling molecules in order to survive (e.g. survivin, telomerase, EGFR and Bcl-2), invade (e.g. MMP and uPA), migrate (e.g. RhoGTPases), colonize [adhesion molecules such as integrins, focal-adhesion-kinases (FAK), cadherins and laminins], and form new blood vessels [e.g. vascular endothelial growth factor (VEGF), basic fibroblast growth factor (bFGF) and hypoxia inducible factor (HIF)] (61).

Several theories have attempted to explain the reasons for metastatic cells to colonize certain organs and tissues. The ‘seed and soil’ theory, (62) suggests that some tumor cells (seeds) selectively colonize in distant organs (soil) where there is a favorable environment for localization and growth. Accordingly, the spread of breast cancer cells to the bones is thought to have a selective advantage

because the matrix of bones contains high concentrations of growth factors that would accelerate the proliferation of tumor cells (63). Alternatively, the chemoattraction theory postulates that organ-specific attractant molecules help direct migrating tumor cells to specific sites (64). In several breast cancer specimens, the chemokine receptors, CXCR4 and CCR7 are upregulated while their ligands (CXCL12 and CCL21 respectively) have been localized to the potential sites of metastasis such as lung liver and bone marrow (65).

Another theory is that vascular endothelial cells of certain organs express adhesion molecules that specifically trap circulating tumor cells and facilitate extravasation or the transit of tumor cells from the blood stream into a new tissue (66). There are several functional similarities in the process of extravasation of white blood cells (leukocytes) and tumor cells that involve, rolling, adhesion and transmigration (67). In response to inflammation, the endothelial cells express several adhesion molecules (e.g. LFA1, Mac-1, E-selectin) to support leukocyte extravasation (68). During the rolling step, integrins such as the leukocyte β-2 ($\beta 2$) integrins, LFA-1 (leukocyte function-associated antigen) and Mac-1 (macrophage adhesion ligand-1) bind to the intercellular adhesion molecule 1 (ICAM-1) to establish firm adhesion of leukocytes to endothelium (69), (67).

Similarly, the circulating breast tumor cells may exploit the ability of adhesion molecules such as ICAM-1, which is expressed on peri-tumoral stromal cells and endothelial cells (70), for extravasation by binding to other molecules that are overexpressed on the cell surface. In support of this, previous work in our laboratory established that MUC1 may contribute to breast cancer metastasis by binding to ICAM-1, (71) and demonstrated that this interaction mediates

downstream signaling events leading to the trans-endothelial migration of breast cancer cells (72). Collectively, these findings suggest that ER positive, Luminal B type tumors may use a mechanism based on receptor-ligand recognition to mediate cell movement showing a slow but cumulative increase in mortality over time. The structure-function relationships and the role of MUC1 in breast cancer are reviewed in the next section.

1.5. The Mucin, MUC1

1.5.1. The Mucins

Mucins are large, heavily glycosylated proteins involved in regulating diverse cellular functions both in normal and pathological conditions (73) and are normally expressed in the respiratory and gastrointestinal tracts and the ductal surfaces of organs such as the breast, pancreas, liver and kidney (74). The amount of glycosylation depends on the type of mucin, the site of mucin expression and the physiological or pathological conditions (75). In turn, the extent and nature of the mucin glycosylation determines the biochemical and biophysical properties of mucins (76).

The mucins are classified into two main classes; *viz*. i) the secreted mucins (MUC2, MUC5AC, MUC5B, MUC6, MUC7, MUC8 and MUC19) which lack transmembrane and cytoplasmic domains, and ii) the membrane bound mucins (MUC1, MUC3A, MUC3B, MUC4, MUC12, MUC13, MUC15, MUC16, MUC17, MUC20 and MUC21) (23), (77), that are composed of a large extracellular domain, a single-pass transmembrane domain and a relatively short cytoplasmic domain (78).

The secreted mucins form a thick mucous gel that act as a mechanical barrier for lubrication and protection of epithelial cells from pathogens as well as chemical, enzymatic, and mechanical damage (79). In addition to these functions, the transmembrane mucins play diverse roles in the progression of cancer, some of which are thought to increase the metastatic capability of tumor cells (reviewed in (23), (73), (76), (74), (80), (81), (78)). The amino acid sequence variation among different transmembrane mucins is very high, which may be attributed to their unique functions (76).

Of the transmembrane mucins that are associated with pathological conditions, MUC1 is the most studied mucin and has been reported to function as a cell surface receptor (82), which is involved in tumor cell adhesion, survival and signal transduction in the cancerous breast. The cytoplasmic domain of MUC1 forms molecular complexes with various signaling molecules including, oncoproteins, growth factor receptors and adaptor proteins that are involved in cancer cell proliferation, invasion and migration related oncogenic pathways (73), (23).

1.5.2. Expression of MUC1

The human MUC1 (also known as Episialin, DF3, CA15-3, CD227, PEM, EMA, ETA, MEA, MCA) is an integral membrane glycoprotein, which is encoded on chromosome 1q21. It was first isolated from breast milk and subsequently cloned from breast and pancreatic carcinomas as a tumor antigen (83), (84), (85). In the normal breast, MUC1 is expressed on the apical surface of glandular epithelial cells, but in breast cancer cells, it is overexpressed and shows

an abnormal distribution (86). The overexpression of MUC1 in breast cancer cells is due to gene amplification and/or elevated transcription (87). A recent study claims that the gene copy number of MUC1 increases from normal breast tissue to primary invasive breast carcinomas, and that this correlates with MUC1 protein overexpression (88). In estrogen receptor positive breast cancer, MUC1 expression is significantly upregulated by the transcription factors, Estrogen receptor alpha (ERα) (89) and GATA3 (90). Transcription of MUC1 can also be elevated in breast tumor cells due to constant stimulation by pro-inflammatory cytokines, such as tumor necrosis factor-alpha (TNF-alpha) and interferon-gamma (IFN-gamma) (91).

MUC1 expression has been correlated with prognosis in human breast cancers. For instance, the circumferential membrane expression of MUC1 in breast cancer cells is associated with increased lymph node metastases while increasing amounts of cytoplasmic MUC1 is correlated with poor survival of breast cancer patients (92). The level of expression of MUC1 and its aberrant glycosylation patterns are positively correlated with the aggressiveness and metastatic potential of the breast cancer cells (93) and thus, MUC1 is implicated as a prognostically significant breast tumor marker (94). Increased expression of MUC1 has also been reported in other malignancies such as cancers of the ovary, colon, stomach, lung, and bladder (95), (93), (96).

1.5.3. The Structure-Function Relationships of MUC1 Molecule

MUC1, is a type I transmembrane glycoprotein that consists of three major domains (83), extracellular (ECD), transmembrane (TM) and cytoplasmic (CD) (Fig. 1.3). The newly synthesized MUC1 polypeptide undergoes an autoproteolytic cleavage in the endoplasmic reticulum due to the mechanical stress, induced by the folding of the polypeptide chain (97), (98), (99). The resulting subunits remain associated during the maturation of MUC1 molecule by multiple cycles of clathrin-coated endocytosis followed by sialylation in the Golgi complex before being finally tethered to the cell surface as a relatively large glycoprotein (300 -500 KDa) (100), (101), (78). The large N terminal (a) subunit (> 200 KDa) is non-covalently associated with the short extracellular stub of the C terminal or b subunit (17 KDa) that contains transmembrane and cytoplasmic domains. The correlation of the structural features and functional uniqueness of each domain is discussed in the next subsections.

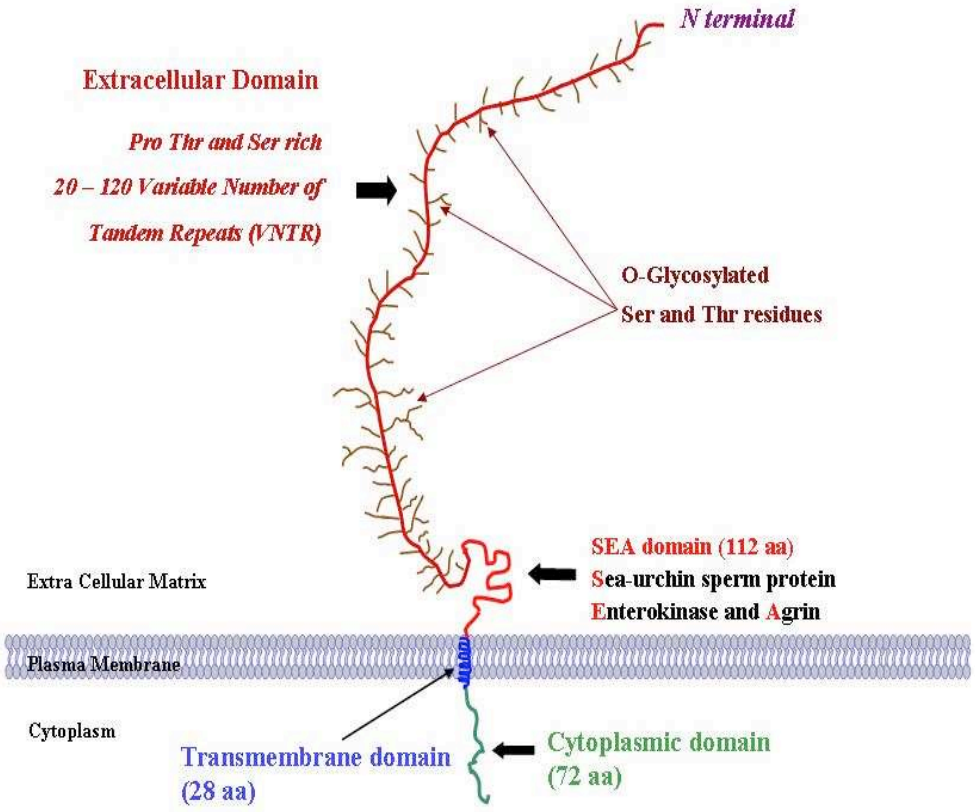


Figure 1.3. A schematic of the domain organization of MUC1 molecule, adapted from (102).

1.5.3.1. The Extracellular Domain

The extracellular domain (MUC1-ECD hereafter) consists of the N-terminal subunit and the extracellular stub of the C terminal subunit of the MUC1 molecule. The N terminal region of MUC1-ECD has a signal peptide that directs the insertion of MUC1 polypeptide into the endoplasmic reticulum for cell-surface delivery (78). The majority of the ECD consists of a variable number (20 – 120) of tandem repeats (VNTR) of a 20 amino acid sequence (VTSAPDTRPAPGSTAPPAHG) that contain a relatively higher number of serine (S), threonine (T) and proline (P) residues, flanked by unique non-repetitive sequences (103).

MUC1-ECD was first believed to be in a random coil conformation (104) however, the structural data of 1 – 3 tandem repeats, experimentally obtained by NMR spectroscopy, confirmed that the VNTR region forms an extended polyproline type II (PPII) like backbone conformation with large repeating loops crested by a beta turn (105). The tandemly repeated beta turn PPII conformation does not fold into a higher ordered structure with increasing numbers of repeats. The MUC1 molecule, therefore, protrudes more than 200 nm above the apical surface of breast epithelia (106). Structural analyses of O-glycosylated (at Threonine residue of the PDTRP part of the tandem repeat) and non-glycosylated peptides (tandem repeats) revealed that the O-glycosylated peptide was more extended and rigid compared to the non-glycosylated peptide. The non-glycosylated PDTRP motif formed knob like protrusions (107). The exposed repeats of the PDTRP motif facilitate effective antibody binding and are responsible for the immunodominant properties of MUC1-ECD that characterizes

MUC1 as a tumor antigen (108). Thus, the underglycosylated MUC1 molecule, which is typical to malignant cells, contains slightly more protruded and exposed PDTRP motifs in the backbone that may increase the chances of ligand binding. Further, the extended tandemly repeated structure of the ECD increases the multivalency of the MUC1 molecule, which enhances the antibody binding affinity (106). Thus, the MUC1 molecule has been used effectively for the development of tumor markers, tumor vaccines and targeted antibodies that are commonly used to screen breast cancer patients (109).

MUC1-ECD is O-glycosylated at the serine and threonine residues of the tandem repeat region and N glycosylated at 5 sites on asparagine residues (85) located at the C terminal segment of the ECD that flanks the tandem repeats (110). There are five potential O-glycosylation sites in each tandem repeat. The O-glycosylated residues support the rigid extended structure of its backbone due to the highly negatively charged side chains that avoid "close packing" with other molecules (111), (112) while the N-Glycosylation is essential for the stability, folding, transport, and secretion of MUC1 (110). The degree of glycosylation of MUC1 molecule decreases from normal to cancerous breast (113) and facilitates several biological functions such as acting as a barrier to pathogens and cytotoxic lymphocytes (114), modulation of clathrin mediated endocytosis (115) and altering adhesive and anti-adhesive functions that regulate its binding properties (116). The anti-adhesive properties of the highly glycosylated MUC1-ECD are thought to interfere with the E-cadherin and integrin mediated cell-cell and cell-matrix adhesion thereby enhancing the invasiveness of malignant cells (117), (118), (119), (120) while the adhesive properties of underglycosylated MUC1-

ECD allows binding to ligands such as ICAM-1, which initiates oncogenic signaling leading to cell migration (71).

At the C terminal end of the MUC1-ECD, a 120-residue SEA (Sea Urchin Sperm Protein Enterokinase and Agrin) domain is located close to the membrane (121) (Fig.). The MUC1 molecule is cleaved posttranslationally in the endoplasmic reticulum due to conformational stress imposed on the site, G^SVVV, within the SEA domain (99), (98) and exists on the cell surface as a non-covalently associated heterodimer (100). Although the exact function still remains elusive it is postulated that the SEA domain, i) acts like a molecular mechanical fracture device (98) for autoproteolytic cleavage and ii) permits the shedding of the bulky N terminal tandem repeat region of MUC1-ECD, as necessary. Mutation of the Glycine and Serine at the cleavage site to Valine and Proline prevents cleavage and shedding of the MUC1-ECD (122). The extracellular domain is also cleaved and shed proteolytically by TACE/ADAM17 (123), or MT1-MMP (124). The processing of the MUC1 molecule is consistent with the fact that ectodomain shedding regulates most cellular functions of type 1 transmembrane proteins, in which the released intracellular domains interact with cytoplasmic signaling intermediaries and transcription factors in the nucleus [74].

In addition to the full-length MUC1 molecule, alternative splice isoforms of MUC1 with incomplete extracellular domains are expressed on the surface of normal and/or malignant cells (125), (126), (127). The isoform, MUC1/Y (42 – 45 KDa), lacks the bulky N terminal tandem repeat region of the MUC1-ECD as well as a part of the SEA domain and is not cleaved in the ER, posttranslationally (99). It binds to another isoform, secreted MUC1, MUC1/SEC (that lacks both TMD

and CD), which results in tyrosine phosphorylation of the MUC1/Y cytoplasmic domain (128). Another alternative splice isoform, MUC1/X, lacks the tandem repeats but contains the complete SEA domain segment of MUC1-ECD and thus can undergo autoproteolytic cleavage similar to the full-length MUC1 molecule (99). The full length MUC1 (MUC1/TM) and the splice isoforms, MUC1/X and MUC1/Y are expressed in tumor cells, whereas MUC1/SEC is mostly expressed in normal cells (48).

1.5.3.2. The Transmembrane Domain

MUC1 has a highly hydrophobic, 28-residue single pass transmembrane domain (MUC1-TMD hereafter) (101). As a type-1 transmembrane glycoprotein, MUC1 is anchored to the plasma membrane with a membrane anchor sequence, which orients its N-terminal end to the lumen of endoplasmic reticulum during synthesis and then to the extracellular space on the cell surface (129). It is postulated that MUC1-TMD plays a major role in the membrane localization of MUC1 as well as distribution of MUC1 molecule in lipid rafts (130). The cysteine residues in the cysteine-glutamine-cysteine (CQC) motif at the junction of the cytoplasmic and transmembrane domains are involved in cysteine-mediated palmitoylation that aids the membrane localization of MUC1 molecule (131). Site directed mutagenesis of these cysteine residues to alanine (CQC to AQA) resulted in failure of trafficking of the MUC1 molecule to the membrane from endosome recycling. Since these AQA mutants could not be extracted from the membranes of transfected pancreatic cell lines (132) it has been postulated that the CQC motif is essential in the lipid raft localization of MUC1. Recent studies have shown that

the dimers and oligomers of MUC1, formed via disulphide linkages of the CQC motif, are necessary for the downstream signaling events (133) and the nuclear localization of the cytoplasmic domain of MUC1 (134).

1.5.3.3. The Cytoplasmic Domain

The cytoplasmic domain of MUC1 (MUC1-CD hereafter) is relatively short with 72 amino acids (7.8 KDa) (Fig. 1.4) and has been identified as a hub of oncogenic signaling since it interacts with various signaling molecules, most of which are involved in adhesion, proliferation, invasion and metastasis of tumor cells (135). The MUC1-CD has no intrinsic kinase activity but phosphorylation is proven to be the key mechanism that allows it to regulate downstream signaling, functionally similar to the cytokine receptor-like signaling molecules, which also show some sequence similarity to MUC1-CD (136), (137). To initiate phosphorylation dependent signaling cascades in tumor cells, MUC1-CD contain several functional motifs that include tyrosine, serine or threonine residues, some of which are highly conserved across mammalian species (138).

In breast cancer cells, MUC1-CD is phosphorylated by non-receptor tyrosine kinase, Src (139), serine/threonine protein kinases, glycogen synthase kinase 3 beta (GSK 3 β) (140) and the delta isoform of protein kinase C (PKC δ) (141). Activated cell-surface growth factor receptors also phosphorylate MUC1-CD in breast cancer cells; e.g.(s) epidermal growth factor receptor (EGFR) (142), (143), human epidermal growth factor receptor-2 (Her2/ErbB2), ErbB3 and ErbB4 (143), fibroblast growth factor receptor 3 (FGFR3) (144).

There are other signaling molecules that associate with MUC1-CD either

directly or in a phosphorylation dependent manner and contribute to oncogenic signaling pathways in the breast; e.g.(s) Grb2 (145), beta (β) catenin (139), (142), (140), (146), p120-catenin (147), gamma (γ) catenin (148), Adenomatous Polyposis Coli (APC) (149) and CT10 regulator of kinase like (CrkL) (150). (Fig. 4)

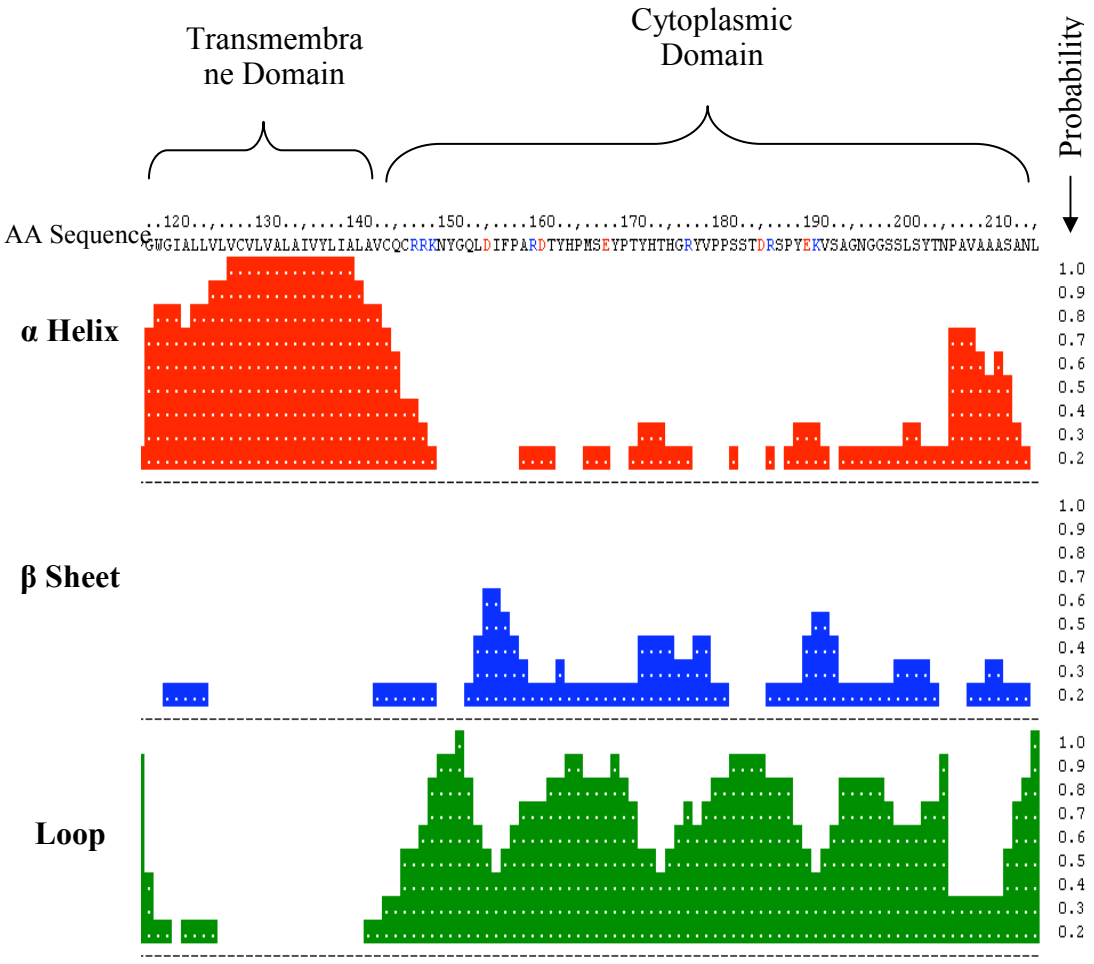
MUC1-CD is reported to be associated with many other molecules in different cell types and signaling contexts (151): e.g.(s) i) the SH2 domains of the non receptor tyrosine kinases, Lyn in multiple myeloma cells (152) and Lck in Jurkat lymphoma cells and normal T cells (153), ii) the SH2 domain of the P85 subunit of phosphatidylinositol 3-kinase (PI3K) (136) in Rat 3Y1 Fibroblasts (154) and Non-small cell lung cancer (NSCLC) cells (155) (See Fig. 1.4 for binding sequences).

Figure 1.4. The amino acid sequence of MUC1 cytoplasmic domain (MUC1-CD) showing the currently identified binding motifs and putative binding sites of the following signaling molecules; **PI3K** –Y^{20}HPM; **Src-SH3** – RYVPPSSTDR ; **PKCδ**– ST^{41}DRSP ; **GSK3β** - $TDRS^{44}$PY ; **Src SH2** - Y^{46}EKV; **β Catenin** – SAGNGGSSLS ; **Grb2** – Y^{60}TNP and putative motifs for **CrkL** binding – Y^{35}VPP or Y^{60}TNP

Since MUC1-CD interacts with a wide variety of signaling molecules of different signaling pathways that integrate cellular signaling networks, it is thought to serve as a scaffold protein (156). The cellular signaling pathways involved in tumor progression are coordinated and/or mediated by several scaffold proteins that assemble signaling complexes, activate enzymes and/or facilitate multiple molecular interactions (157). Increasing evidence suggests that many scaffolding proteins are disordered or intrinsically unstructured in order to accommodate multiple interactions, since the disorder provides flexibility for these molecules to interact with many binding partners with maximized interaction-surface per residue (158). Many disordered proteins maintain flexibility to switch binding specificity via posttranslational modifications and contain structural elements that aid in disorder-to-order transitions needed to bind with different partners (159).

The structure of MUC1-CD has not been experimentally determined. The secondary structure predictions (Fig. 1.5) suggest a largely unstructured or intrinsically disordered nature of MUC1-CD that is consistent with a scaffolding function. The predictions also show that it may have short beta sheet and alpha helix elements interspersed with a largely random coil-like structure (Fig. 1.5), indicating that it may exist as a partially stable, interconverting ensemble of conformations with residual alpha helices and beta sheets. There are multiple interaction motifs of MUC1-CD, some of which partially overlap or are located immediately adjacent to each other, providing further evidence for its scaffolding

function. As a result, some molecular interactions could be sterically hindered by one another and thus be mutually exclusive, causing a form of competitive binding. Alternatively, the flexibility of MUC1-CD may also allow binding of more than one molecule, cooperatively, to non-overlapping sites, which may lead to the assembly of signaling complexes.



MUC1 Domains	Secondary Structural Elements (%)			Solvent accessibility composition (Core/Surface ratio)	
	Alpha (α) helix	Beta (β) Sheet	Random coil (loop)	Surface exposed (>16%) Residues	All other residues
Transmembrane	79.41	0.00	20.59	35	65
Cytoplasmic	4.17	9.72	86.11	92	8

Figure 1.5. The secondary structure predictions of the cytoplasmic domain of MUC1 compared to the transmembrane domain, which indicates that MUC1-CD may be mainly (>80%) unstructured with some residual alpha helical and beta sheet structure. *PredictProtein server (295)*

Collectively, the flexibility and selective phosphorylations of tyrosine, serine and/or threonine residues may allow both cooperative as well as competitive binding of MUC1-CD to other molecules, depending on the nature of the interaction and the structure-function relationships of the binding partners. For instance, the phosphorylation of Y^{46} and binding of Src (to the SP**Y**^{46}EKV motif) (139) as well as the phosphorylation of T^{41} by PKC δ (binds to S**T**^{41}DRS motif) (141) upregulate the direct binding of MUC1-CD to β-catenin (binds to S^{50}AGNGGSSLS motif), a protein overexpressed in breast cancer. However, the phosphorylation of S^{44} by GSK3β (binds to DR**S**^{44}PYE motif) decreases β-catenin binding (140). This evidence implies that MUC1-CD has the ability to mediate cross talk among different cellular signaling pathways.

There is growing evidence concerning the involvement of MUC1-CD in signal transduction via several oncogenic signaling pathways in tumor cells (135), (160), (23). In the normal (non-malignant) epithelial cells, however, MUC1-CD is not reported to be actively involved in the molecular interactions described above. This functional difference between MUC1 in normal and tumor cells and the key signaling pathways, which use MUC1 in the progression of breast cancer, are discussed below.

1.5.4. Roles of MUC1 in Breast Cancer

As an oncogene and tumor antigen, MUC1 has long been associated with progression of breast carcinomas (161). There was a significant delay in tumor formation when Muc1 (the MUC1 gene in mice) knockout mice (Muc1 -/-) were crossed with the strains of mice expressing mouse mammary tumor virus-driven

(MMTV) oncogenes, MMTV-Wnt-1 and MMTV-TGFα (162), (163). Conversely, the MMTV-MUC1 transgenic mice developed spontaneous tumors in the mouse mammary gland proving that MUC1, can in fact, function as an oncogene (24). Also, there was a significant delay in tumor progression when the mouse mammary tumor virus-driven transgenic mice (MMTV-PyV MT), which show increased tumorigenesis in the presence of Src, were crossed into a Muc1 null background (164)

Interestingly, the above mouse models identified that the loss of either the cytoplasmic domain (Muc1-CD) or the tandem repeats region of the extracellular domain (Muc1-ECD), prevented the oncogenic capacity of Muc1 molecule, suggesting that both of these domains are crucial for mammary tumor formation (165). This evidence indicates that MUC1 acts as a molecular sensor on the cell surface and has the ability to initiate outside-in signaling. For instance, binding of MUC1-ECD to other molecules such as ICAM-1 triggers phosphorylation of MUC1-CD (150). In normal epithelial cells unlike tumour cells, such binding events are prevented or spatially separated due to the apical localization and structure-function integrity of the MUC1 molecule (135).

The major functional differences of MUC1 between normal and malignant cells are reported to be due to its overexpression, loss of apical localization and underglycosylation, in malignant breast epithelial cells (80). These differences can be related to the following observations; viz. i) Since MUC1 is expressed at the apical surface of normal, polarized epithelial cells, it is spatially separated from other molecules that are not apically expressed and thus the interactions of MUC1 with such molecules are prevented; for instance, in cancer cells, non-

apically localized MUC1 can constitutively interact with growth factors such as ErbB1-4, FGFR3 and PDGFRβ that are expressed at the basolateral membrane (160), ii) it is possible that MUC1-CD is degraded in normal cells if it is separated from the extracellular domain (135) and/or when it is not modified by phosphorylation and/or not associated with other molecules, iii) since MUC1 is overexpressed and non-apically polarized in malignant cells, the propensity of MUC1-CD for aggregation may also be increased. Although upstream trigger signaling mechanisms are not revealed so far, there is evidence for aggregation of MUC1-CD via disulphide linkages at the CQC motif (136), which could serve as a mechanism to avoid degradation and thereby allow multiple phosphorylations and subsequent molecular associations (134). Overexpression of MUC1 is also associated with the accumulation of MUC1-CD in the cytoplasm and its targeting to the nucleus and mitochondria (166). Overexpression, aberrant glycosylation and loss of apical polarity of MUC1 molecule have been associated with its diverse roles in tumor cells including i) oncogenic signaling functions, ii) anti-adhesive functions, and iii) adhesive functions, all of which, are proven to contribute to tumor progression as illustrated in the upcoming subsections.

1.5.4.1. MUC1 as an Oncogenic Signaling Molecule

Numerous studies report the involvement of MUC1-CD in molecular interactions that involve the MUC1-CD in many different types of malignant cells (reviewed in, (135)). The most relevant interactions, directly related to breast cancer progression and metastasis, are discussed below.

The majority of oncogenic properties of the MUC1 molecule are due to

the phosphorylation dependent molecular interactions of MUC1-CD, as introduced in the previous section. Evidence from independent studies show that some external stimuli such as binding of bacteria or antibodies to the extracellular segment can lead to the phosphorylation of MUC1-CD and initiate downstream signaling pathways; e.g.(s) i) binding of MUC1-ECD to the bacterium, *Pseudomonas aeruginosa,* in human airway epithelial cells (167), (168), and, ii) binding of anti-CD8 antibodies to the extracellular domain of chimeric CD8-MUC1 (MUC1-CD fused to the extracellular and transmembrane domains of CD8), in CD8-MUC1 expressing COS-7 cells (169), have both led to the phosphorylation of MUC1-CD followed by activation of a downstream Grb2-Sos-Ras-MEK1/2-ERK1/2 signaling pathway, iii) binding of ICAM-1 to MUC1, in human T47D breast cancer cells are shown to initiate downstream cell migration pathways (150). The tyrosine phosphorylation of MUC1-CD could be inhibited by phosphatase inhibitors, suggesting that it is a reversible regulatory mechanism in cancer cells (136).

Phosphorylated tyrosine residues of MUC1-CD provide docking sites for SH2 domains of adaptor proteins and kinases (170), (139), (142), which often initiate multiple oncogenic signaling cascades. More specifically, of the seven highly conserved tyrosine residues, three are located within the consensus motifs that are experimentally proven to bind, once phosphorylated, to the SH2 domains of the following molecules and initiate downstream signaling in breast cancer cells; (i) p-Y^{20}HPM (phosphatidylinositol 3- kinase), (ii) p-Y^{46}EKV (Src Family Kinases), and (iii) p-Y^{60}TNP (growth factor receptor-bound protein 2, Grb2) (171). Phosphorylation of Y^{20}, in rat 3Y1 fibroblasts, allows MUC1-CD to bind

to the SH2 domain of the p85 subunit of PI3K via p-Y^{20}HPM motif, which activates PI3K-AKT pathway that attenuates mitochondrial apoptotic signaling leading to survival of cancer cells (154,154).

As demonstrated in human MCF7 breast cancer cells, the phosphorylation of Y^{60} at the Y^{60}TNP motif of MUC1-CD allows recruitment of the SH2 domain of Grb2 (Growth factor receptor-bound protein 2), which then binds to SOS (son of sevenless) (145). As a result, SOS activates the G-protein, Ras, leading to activation of a mitogen activated protein kinase (MAPK) signaling through the extracellular signal regulated kinases (ERK1/2) promoting cell proliferation, survival and motility (145). Further to this, the stimulation of EGFR in the presence of MUC1 also triggers the Ras/ERK/MAPK signaling cascade in the human MDA-MB-468 breast cancer cell line as well as in the mouse mammary gland (143).

The upregulation of Src family kinases such as Src and growth factor receptors such as EGFR play a significant role in invasion and migration of breast cancer cells (172), (173). It has been shown that MUC1-CD potentiates the oncogenic signaling triggered by Src and EGFR, both of which phosphorylate Y^{46} of MUC1-CD allowing the binding of Src-SH2 domain to p-Y^{46}EKV motif as shown in human ZR-75-1 breast cancer cells (139), (142). As discussed in a previous section, binding of Src increases binding of β-catenin to MUC1-CD (139) while binding of GSK3β to MUC1-CD decreases the binding of β-catenin (140). Therefore, phosphorylation of Y^{64} and interaction of Src and MUC1-CD may regulate the cytosolic pool of GSK3β and β-catenin molecules, both of which play major roles in the Wnt signaling pathway (81). This evidence suggests that

MUC1 functions in integrating the growth factor receptor and Wnt signaling pathways (174). As described in the next section, the decreased cytosolic pool of β-catenin, due to competitive binding of MUC1-CD, contributes to the destabilization of cadherin-catenin junctions and promotes the invasive capacity of cancer cells (146). The previous studies in our laboratory report that binding of ICAM-1 to MUC1-ECD trigger increased association of Src with MUC1-CD, which contributes to transendothelial cell migration, cellular calcium (Ca++) oscillations and cytoskeletal reorganization pathways (section 1.6.2) (72), (175), (150).

Other studies involving breast cancer cells have revealed that MUC1-CD facilitates the transmission of signals from the cell membrane to the nucleus: e.g. i) MUC1-CD directly binds and induces the nuclear localization of p120 (ctn) in human ZR-75-1 cells (147), ii) Interaction of MUC1 and ErbB-2, after stimulating breast cancer cells with heregulin (HRG), resulted in formation of MUC1-γ catenin complex formation and its targeting to the nucleus.

Taken together, MUC1-CD is a part of several signaling pathways that also involve Src, which has led to the speculation that MUC1 molecules in tumor cells may bring Src closer to its downstream targets such as β-catenin and FAK at the cell membrane (Fig. 1.6) (164). As discussed in the next two sections, the recruitment of Src contributes to inside-out signaling that increases anti-adhesive properties of MUC1, promoting the invasion, migration and metastasis of breast cancer cells. The recruitment of Src to MUC1-CD, on the other hand, is a result of outside-in signaling regulated by pro-adhesive properties of MUC1 that allow binding of ICAM-1.

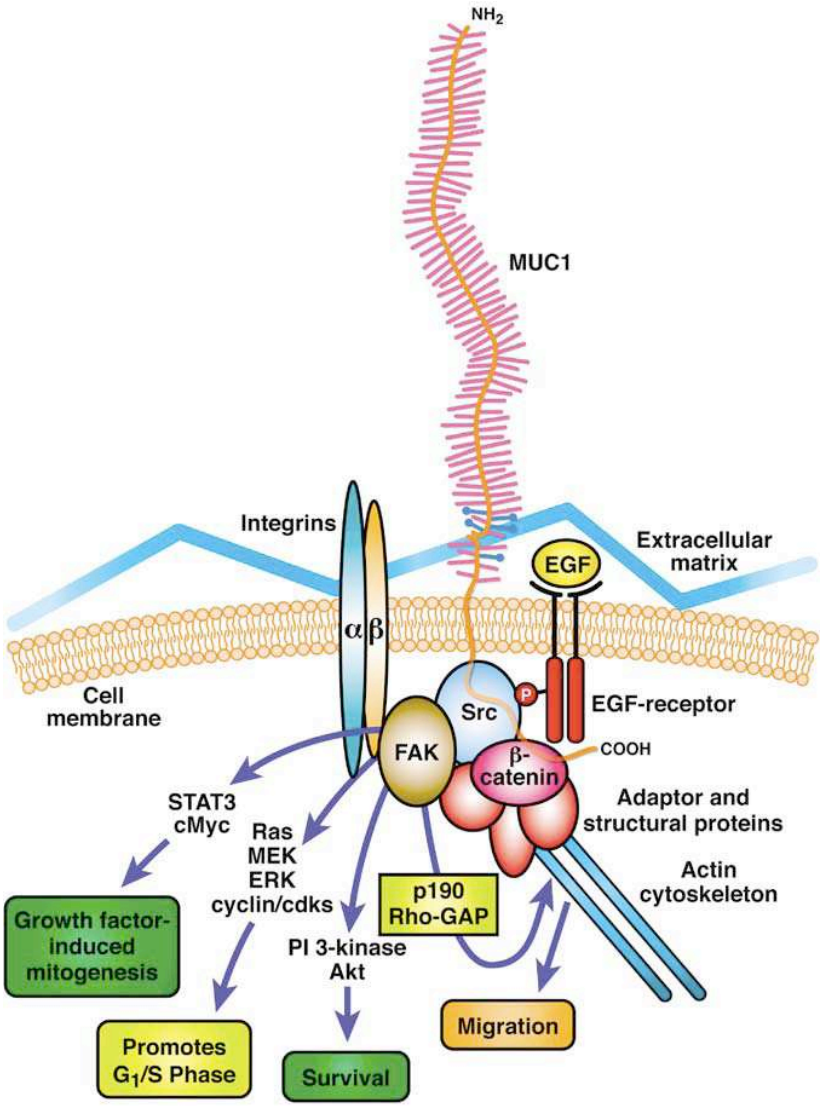


Figure 1.6. A model as shown in Masri *et al* (2005) (164), that illustrates possible ways that Muc1 could be influencing c-Src signaling. The model was originally adapted from Frame, 2002 (59)). According to Masri *et al* (2005) (164), the overexpression of Muc1 by tumor cells may potentiate c-Src signaling by bringing it closer to its downstream targets such as β-catenin, FAK, and p85 at the cell membrane. Activation of c-Src influences various aspects of cell behavior including growth, proliferation, survival, and migration.

1.5.4.2. MUC1 as an Anti-Adhesive Molecule

The MUC1 molecule demonstrates anti-adhesive properties on both structural/physical and functional levels. The highly sialylated, negatively charged, rod-like extracellular domain physically destabilizes cell-matrix connections and thus may sterically hinder cell-cell aggregation promoting detachment of cancer cells from the primary tumor mass (176), (118), (120), (177).

Also, the above-discussed interaction of MUC1-CD and β-catenin, which is upregulated by Src and PKCδ, promote anti-adhesive properties by disrupting cell-cell adhesion at the adheren junctions. The cadherins, adhesion molecules on the cell surface, interact with the cytoplasmic α-catenin, β-catenin and γ-catenin proteins (74) that are involved in cell-cell adhesions at the adherein junctions and maintain the structural integrity of the actin cytoskeleton. Particularly, β-catenin links E-cadherin to α-catenin, which in turn forms homodimers and interacts with the actin cytoskeleton. MUC1-CD interferes with E-cadherin and integrin mediated adhesions, by competitively binding to β-catenin (162), (146). Disruption of the adheren junctions affects the regular maintenance of polarity, shape and dynamics of the epithelial tissue (178), which allows the cancerous cells to gain migratory properties and detach from the epithelial tissue. The anti-adhesive property of MUC1 molecule also increases the survival of tumor cells from T-cell mediated cytolysis (179).

1.5.4.3. MUC1 as a Pro-Adhesive Molecule

In addition to the above-discussed roles of MUC1 in cancer, it plays an important role in adhesion, which may aid the attachment of circulating tumor cells to the blood vessels of distal organs, one of the rate limiting steps in tumor metastasis. Some studies have reported the adhesive ability of MUC1-ECD due to the presence of sialyl Lewis$^{x/a}$ carbohydrates on its O-linked glycans that promotes binding to selectin-like molecules on nearby cells (78). MUC1-ECD is underglycosylated in the cancerous breast tissue with less branched or truncated sugar chains (113), (180) and thus its protein backbone is relatively more surface exposed compared to that of the normal breast tissue. The previous studies in our lab revealed that the exposed immunodominant PDTRP epitope of the tandem repeat region of MUC1 molecule binds to ICAM-1 in breast cancer cells (71).

Overall, when the MUC1 molecule is non-apically localized, underglycosylated and present at high density on the cell surface, it can interact with its ligand, ICAM-1 and promotes invasion, migration, adhesion and metastasis of breast cancer cells thereby contributing to tumor progression (181).

1.6. Interaction of MUC1 and the Intercellular Adhesion Molecule, ICAM-1

1.6.1. Structure and Function of ICAM-1

The intercellular adhesion molecule 1 (ICAM-1) is a transmembrane glycoprotein of the immunoglobulin (Ig) superfamily and consists of a rod shaped extracellular domain that contains five Ig-like domains (453 aa), a hydrophobic transmembrane domain (24 aa) and a short cytoplasmic domain (28 aa) (182). On

the cell surface, ICAM-1 exists as a covalently associated homodimer (183). The crystal structure of ICAM-1 demonstrates that the intimate contact of domains 4 and 5 in the dimer provide a rigid stem optimally orienting the ligand binding sites in domains 1 and 3 (184), (185). In the dimeric Y shaped molecule, the domain 1 is located at the tips showing an increased avidity for ligand binding (186), (187). The dimerization interface in the extracellular domain 1 interacts with that of an adjacent dimer and forms W shaped tetramers although the most stable form of ICAM-1 on the cell surface is reported to be a dimer (184). ICAM-1 is expressed at basal levels in several cell types including endothelial and epithelial cells but is significantly upregulated by inflammatory responses and binds to β2 leukocyte integrins to promote the transit of leukocytes to inflamed tissues (69). Therefore it is possible that tumor cells may use the capability of ICAM-1 molecules to promote extravasation of tumor cells, through vessel walls, to distant metastatic sites (188).

1.6.2. MUC1-ICAM-1 Interaction and Downstream Signaling

Our laboratory was the first to report pro-adhesive properties of MUC1 showing that it binds to ICAM-1 (71), which has been confirmed by subsequent studies (189), (116). The immunodominant knob-like epitopes (PDTRP) of the tandem repeat region are shown to bind to the extracellular domain-1 of ICAM-1 (71), (189). Supporting evidence suggests that the treatment with benzyl-α-GalNAc, an inhibitor of O-glycan extension, increases MUC1 binding to ICAM-1 (116). The binding of ICAM-1 could be competitively inhibited by a synthetic peptide containing six 6 successive tandem repeats (20 amino acids each) of

MUC1 ECD (190).

Subsequent studies in our laboratory revealed that the MUC1/ICAM-1 interaction increases the transendothelial migration of MUC1 bearing breast cancer cells *in vitro*, through a monolayer of ICAM-1-expressing cells (72) and triggers Ca++ oscillations through a Src mediated PI3K=>PLCγ pathway (175). The most recent work in our lab shows that the binding of ICAM-1 to MUC1-ECD also triggers a Src and CrkL (CT10 regulator of kinase like) dependent signaling cascade that initiates cytoskeletal rearrangements in human T47D breast cancer cells. This study also confirmed that Src functions upstream of CrkL (150). Src may phosphorylate the tyrosine residues at the putative CrkL binding motifs, Y^{35}VPP and/or Y^{60}TNP, which may then allow recruitment of CrkL to MUC1-CD, leading to the formation of a Src/CrkL/MUC1-CD signaling complex (150). The adaptor protein CrkL belongs to the Crk family of adaptor proteins that mediate cell migration by associating with guanine nucleotide exchange factors (GEFs), which catalyze RhoGTPase activation leading to membrane protrusive motility (191).

After binding to ICAM-1, MUC1 may undergo physicochemical changes such as cleavage and shedding of ECD and/or dimerization via the juxtamembrane CQC motif that may facilitate Src recruitment, although the regulatory mechanisms of this are still unclear.

Since MUC1 does not have intrinsic kinase activity, it cannot trigger phosphorylation dependent downstream signaling events without the aid of a kinase and/or other signaling molecules. The abovementioned signaling pathways(s) initiated by MUC1/ICAM-1 interaction, in breast cancer cells, thus,

depend on the recruitment of non-receptor tyrosine kinase, Src, by MUC1-CD although the exact mechanism of Src recruitment is unknown. The following section addresses the structure-function relationships of Src, more specifically the structural basis of the SH3 domain-mediated molecular interactions of Src, as MUC1 may use a similar mechanism(s) to recruit Src.

1.7. The Non Receptor Tyrosine Kinase, c-Src (Cellular Src)

The human cellular Src (c-Src) gene, localized in chromosome 20, also known as pp60c-Src (**p**roto-oncogene **p**rotein with molecular weight of **60**KDa), is a proto-oncogene, and the cellular homologue of transforming Rous sarcoma viral oncogene, v-Src-tyrosine kinase, which was the first oncogene discovered (192), (193). The c-Src (Src hereafter) is the most widely studied member of the Src family of non-receptor tyrosine kinases (SFKs) (194) that play key roles in regulating cellular signaling pathways in multiple cellular environments, most of which are involved in tumor development (195), (196), (57).

1.7.1. Structure and Function of Src

Structurally, Src consists of a myristoylation site, a unique region followed by SH3 (Src homology-3), SH2 (Src homology-2) and kinase domains (197) (Fig. 1.7). The kinase activity of Src is repressed in normal cells where the dephosphorylated Tyr 419 (at the activation loop) is prevented from phosphorylation by a highly stabilized, inactive quaternary structure of Src molecule, primarily attained by intramolecular SH2 and SH3 domain interactions (198) (Fig. 1.7). However, in many types of human tumors and cell lines derived from tumors, including breast cancer, Src is activated by multiple mechanisms,

and thus promotes cell proliferation, survival, motility, invasion and metastasis by interacting with many other oncogenic signaling molecules (196), (199), (200), (201), (202), (203), (204), (205).

The switching between active (open) and inactive (closed) conformations of Src involves dramatic conformational changes in the molecule and is catalytically regulated by phosphorylation and/or dephosphorylation of two tyrosine residues, one located at the C terminal tail (Tyr 530 in human c-Src and Tyr 527 in chicken c-Src) and the other in a cleft between the N and C lobes of the kinase domain (Tyr 419 in human c-Src and Tyr 416 in chicken c-Src) (196). The catalytically inactive conformation of Src is attained by, i) phosphorylation of Tyr 530 by c-Src kinase (Csk), which allows the intramolecular binding of p-Tyr 530 (phosphorylated Tyrosine 530) to the SH2 domain, and by, ii) the positioning of the SH3 domain in contact with the polyproline (PPII) type region in the SH2-kinase linker (197) (Fig. 1.8-a).

Src is activated by i) trans-autophosphorylation of Tyr 419, ii) dephosphorylation of Tyr530 by protein tyrosine phosphatases (such as PTPα), which displace the tail from the SH2 domain and initiate an open conformation and/or iii) displacement of intramolecular interactions by high affinity SH3 (polyproline type II) ligands and/or SH2 (phosphotyrosine) ligands that unfold the inactive molecule and subsequently allow full activation by trans-autophosphorylation of Tyr 419, which results in a conformational change of the activation loop (Fig. 1.8).

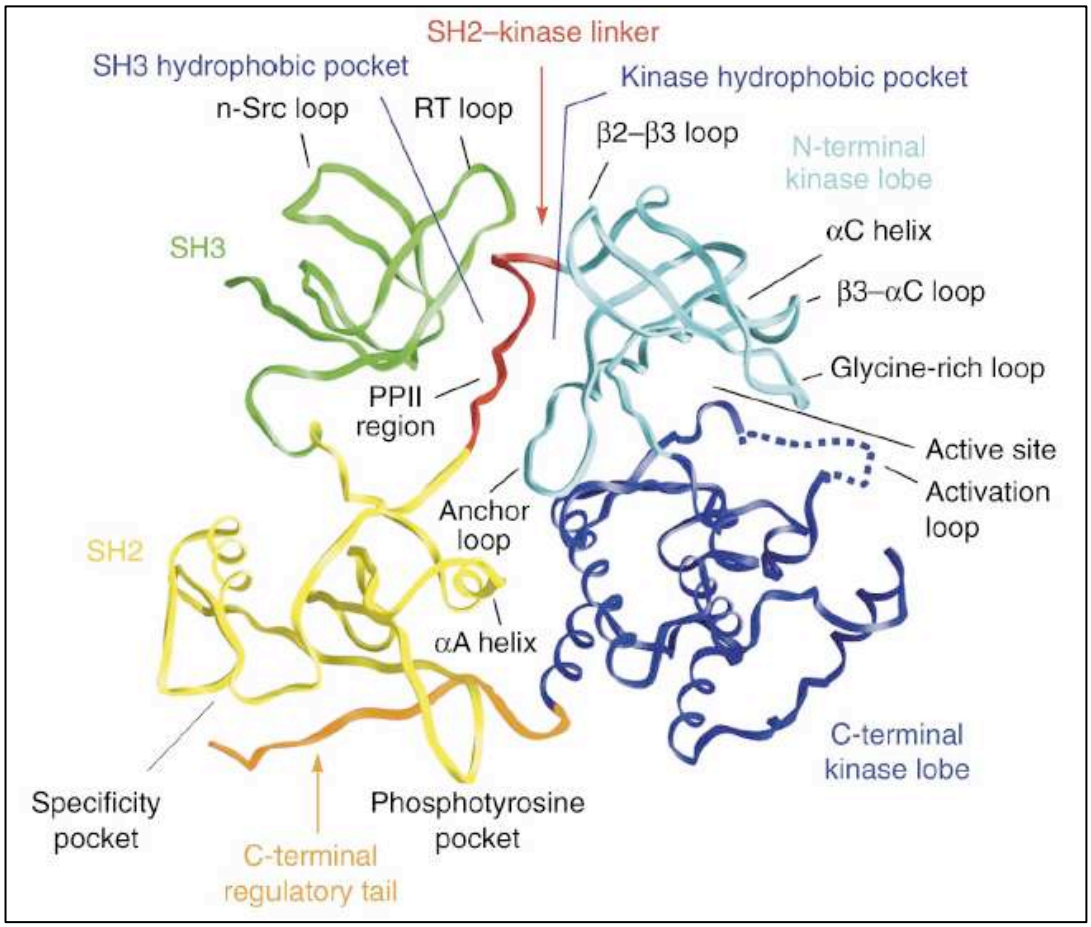


Figure 1.7. Structure of inactive Src kinase (side view) based on (206) adapted from (207). The SH2–kinase linker (red) intercalates between the SH3 domain (green) and the N-terminal kinase lobe (light blue). The SH2 domain (yellow) lies next to the C-terminal kinase lobe (dark blue). The C-terminal tail (orange) reaches across the interface between the kinase and SH2 domains to bind to the SH2 phosphotyrosine-binding pocket.

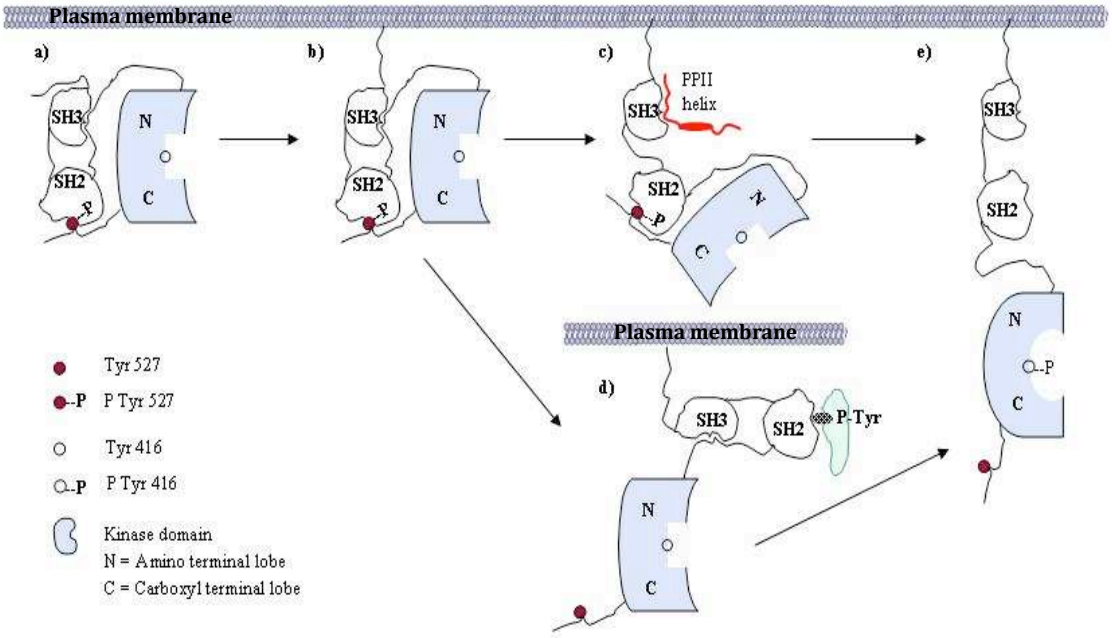


Figure 1.8. A schematic of the mechanisms of activation of Src; a) *Inactive Src*: the SH2 domain is bound to the phosphorylated tyrosine (p-Tyr$^{527/530}$) at the C terminal tail and the SH3 domain is bound to polyproline type II helix of the linker between the SH2 and kinase domains (208) ; b) *myristoylation and membrane anchoring of Src*; both SH2 and SH3 domains are still bound intramolecularly but the conformation is equilibrated towards activation (209) ; c) *partially active Src*: p-Tyr$^{527/530}$ is bound to the SH2 domain but SH3 domain is displaced by a PPII ligand (red) (210) ; d) *partially active Src*: Tyr$^{527/530}$ is dephosphorylated, SH2 domain binds to a p-Tyr of a ligand while the SH3 domain is still bound intramolecularly; e) *fully active Src*: p-Tyr$^{527/530}$ is dephosphorylated; Tyr$^{416/419}$ is phosphorylated.

The reduced expression of Csk contributes to a relatively higher level of active Src in colon and Hepatocellular carcinomas (211), (212). The mutation of Tyr530 at the C terminal tail (to phenylalanine) could result in constitutive activation of Src, but it has been confirmed only in colon carcinoma (213) suggesting that mutations are not the primary mechanism of Src activation (57).

In a majority of human tumors, including breast cancer, Src is activated by displacement of intramolecular interactions by a wide variety of SH3 and SH2 ligands, most of which are integrins, receptor tyrosine kinases (RTKs)/growth factor receptors, immune recognition receptors, adhesion receptors, G-protein coupled receptors, cytokine receptors, scaffold proteins etc. (214), (215), (216), (217). In breast cancer cells, the cytoplasmic signaling molecules such as β integrin, focal adhesion kinase (FAK), Crk-associated substrate (CAS) and, activated cell-surface growth factor receptors such as EGFR, PDGFR, FGFR, HER2 can displace the intramolecular interactions of SH3 and SH2, in different ways, leading to activation of Src (194).

Active Src then phosphorylates a wide variety of proteins in breast cancer cells and contributes to major cell signaling pathways that directly contribute to progression and metastasis of breast cancer (Fig. 1.9); e.g.(s) i) Src => PI3K => Akt…=> cell survival pathway, ii) Src => STAT3 => VEGF => angiogenesis pathway, iii) Src => FAK => Paxillin/$p130^{cas}$ => cell invasion/migration pathway, iv) Src…=> Ras…=> MEK=> ERK=> cell proliferation pathway, v) MUC1 => CrkL => Rac1/Cdc42…=> Actin Cytoskeletal reorganization and cell migration pathway (59), (164), (194), (150).

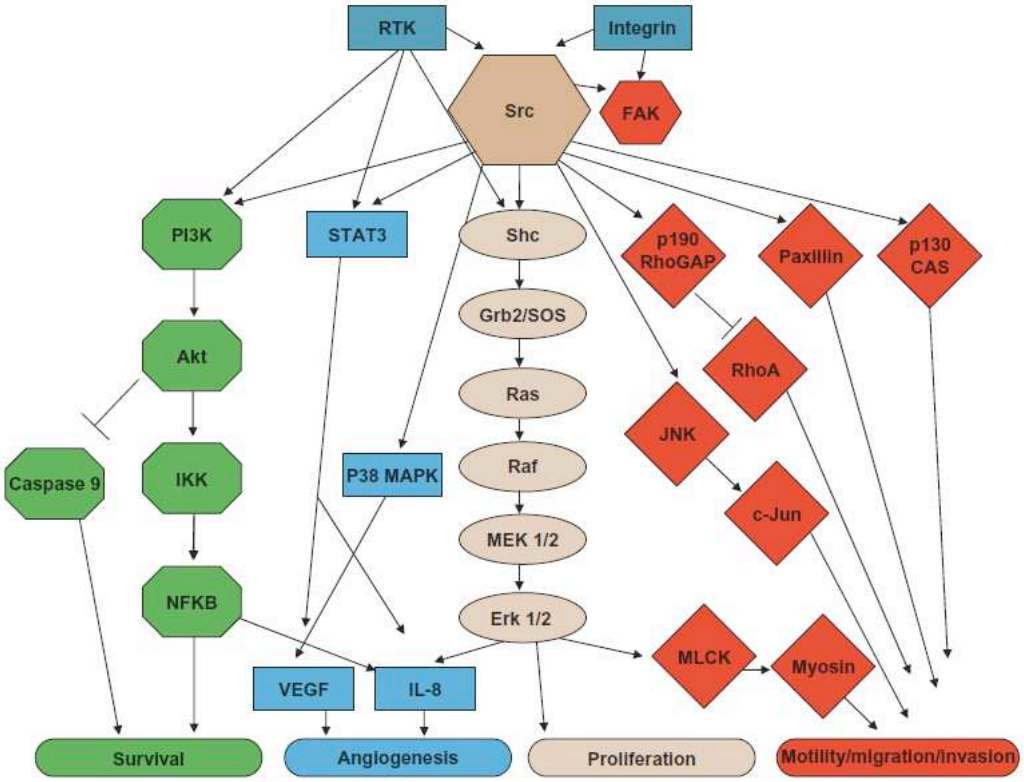


Figure 1.9. The major cellular signaling pathways triggered by Src that contribute to cancer cell growth, survival, motility, and angiogenesis (196). Adapted from (194).

The different modes of activation of Src by SH2-ligands and/or SH3-ligands may predetermine the type of downstream signaling events, which primarily seems to depend on the relative affinity and specificity of the ligand as well as dynamics of ligand binding. It has been shown that myristoylation and membrane anchoring (Fig. 1.8-b) primes Src for partial activation (208), (218) that could initiate the release of intramolecularly-bound SH3 and/or SH2 domains, in the presence of a relatively higher-affinity ligand (209) (Fig. 1.8-c,d). The affinity of the intramolecular interaction of the Src SH3 domain with the SH2-kinase linker has been reported to be weaker than the affinity of external SH3-ligands that typically contain PXXP motifs (219). Also, a phosphopeptide that mimicked the Tyr530 at the C-terminal tail has shown that the intramolecular SH2 domain interaction was 80-fold weaker than binding of an optimal phosphopeptide with the p-YEEI motif (220). Once the SH3 or SH2 domains are engaged elsewhere and Src becomes "partially" activated, the Tyr419 becomes more accessible for phosphorylation and can undergo trans-autophosphorylation by an adjacent partially activated Src kinase molecule. For instance, the ligand binding induced dimerization of the cytoplasmic domains of β integrin brings two partially activated SH3-bound Src molecules facilitating the trans-autophosphorylation of Tyr 419, followed by full activation of kinase activity (221).

The SH3 and SH2 domains are rigidly coupled by a very short linker (IQAEE) and thus, it may not be possible for both domains to interact with the targets that are not ideally spaced (222). Although individual SH3 or SH2 ligands

activate Src, the fixed relative orientation of the SH3 and SH2 domains by this linker (223) may favor unfolding of Src by synergistic activation of properly spaced tandem SH3 and SH2 ligands. For examples, the focal adhesion kinase and adaptor proteins involved in integrin signaling, Sin and p130cas, cooperatively activate Src via both SH2 and SH3 binding motifs (224), (225), (226).

In summary, Src is activated in response to a variety of oncogenic signals in different types of human tumors. Unlike normal human breast, c-Src activity is increased 4-30 fold in breast cancers (205) due to the elevation of specific activity of Src but not necessarily increased level of expression of the protein (204), (200). As described above, the competitive binding of ligands to SH3 and SH2 domains is the major cause of increased Src activity. The structural basis of SH3 domain interactions is discussed in the next subsection.

1.7.1.1. The SH3 Domain

The Src homology 3 (SH3) domain belongs to a family of small (55 – 70 amino acids) modular interaction domains that bind to proline rich peptides or segments of many other proteins (227). It was originally discovered as a conserved sequence in the viral adaptor protein, v-Crk and in the non-catalytic parts of several other cytoplasmic protein tyrosine kinases (228). The SH3 domain is a hallmark of SFKs but is also found in many other proteins. More than 50 SH3 domains have been identified so far, which belong to a wide variety of proteins including signal transduction enzymes (such as kinases, lipases, GTPases), cytoskeletal proteins (e.g. spectrin), cell adhesion molecules (e.g. FAK, Integrins) and signal transducing adapter proteins (e.g. GrB2) (229). Numerous

studies have established that the SH3 domain is a critical module of cellular signaling networks, which i) targets proteins to specific sub cellular locations, ii) mediates rapid assembly of protein complexes required for cellular signal transduction and iii) manages the conformational stability and activity of their mother molecules via intramolecular interactions (230).

Structurally, the SH3 domain has a beta-barrel fold (Fig. 1.10-a), which consists of five anti-parallel beta-strands packed to form two perpendicular beta-sheets (231). This particular fold brings most of the conserved aromatic residues close to each other (Fig. 1.11) to form the ligand binding site (Trp121, Tyr134, Tyr95, Try93, Tyr139 residues in Src-SH3) (232).

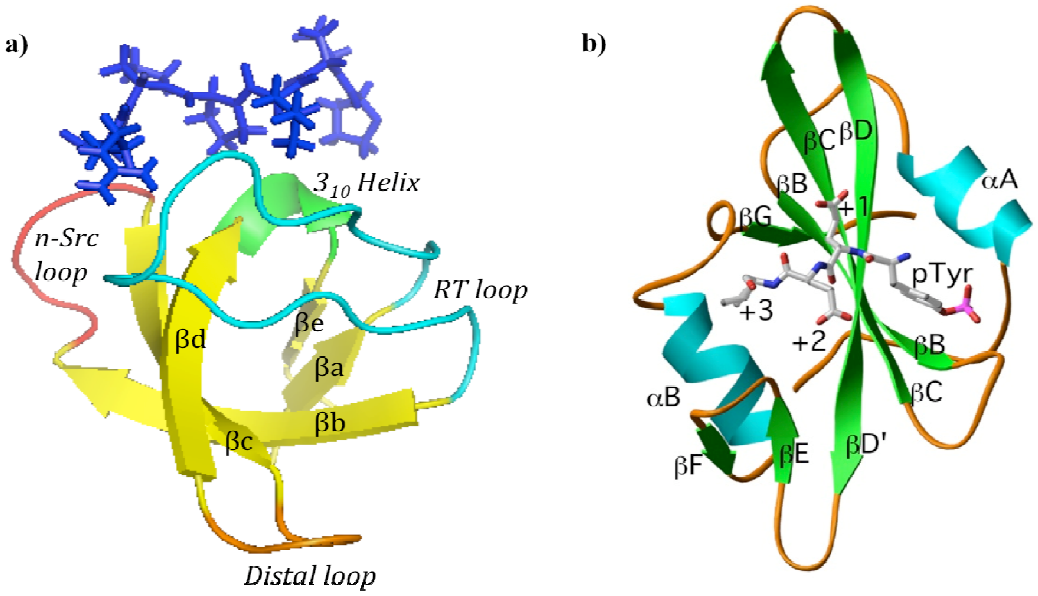


Figure 1.10. The ribbon diagrams of Src-SH3 and Src-SH2 domains: ***a)*** The Src-SH3 domain complexed with a class II ligand, APPLPPRNRPRL (blue) (233) adapted from the solution NMR structure (pdb code: 1QWE); A=Alanine; P=Proline; L=Leucine; N=Asparagine; R=Arginine. ***b)*** The Src-SH2 domain complexed with hmT (Hamster middle T antigen) phosphopeptide at the p-YEEI motif (234). The phosphotyrosine (pTyr), Glutamate (+1), Glutamate (+2), and Isoleucine (+3) of the hmT peptide are shown in licorice.

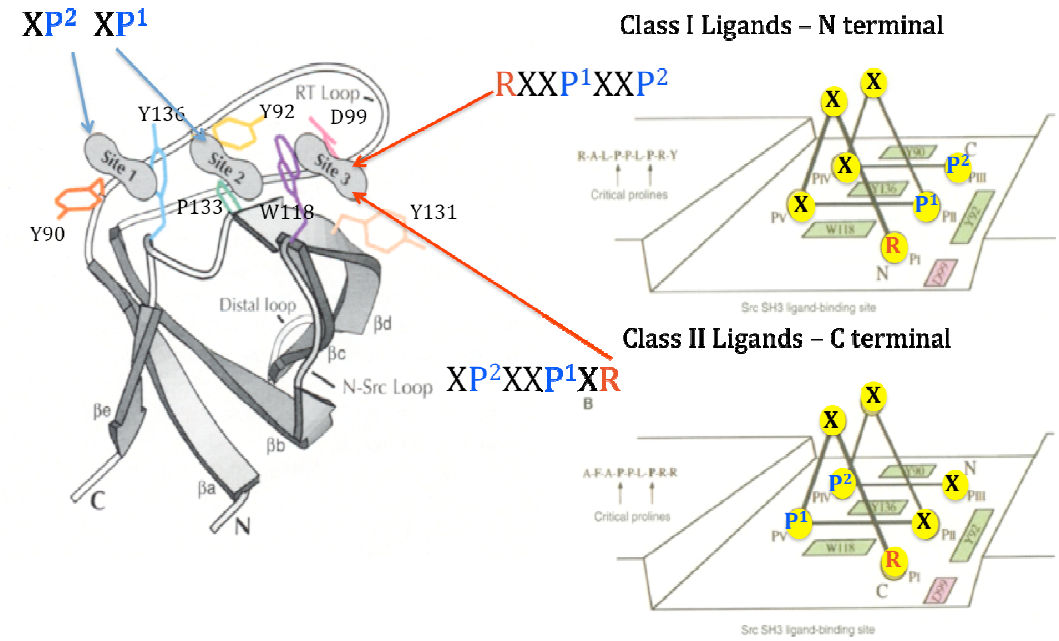


Figure 1.11. On the left: a schematic of the SH3 domain of Fyn showing the location of the ligand binding site relative to the RT and n-Src loops and the conserved aromatic residues that form the three binding grooves that accommodate two *XP* dipeptidyl moieties and the terminal arginine residue of PPII ligands, adapted from (235). On the right: schematics of the ligand-binding site of Src-SH3 domain that show the residues that form the three grooves of the ligand-binding site. The binding orientations and the contact sites of left handed polyproline II helix as in class I and class II ligands are also shown (adapted from (229)).

The SH3 domain typically binds to the ligands that form a left-handed polyproline type II (PPII) conformation with a minimum consensus sequence PXXP (P=Proline; X=any residue) (236). The unique properties of aromatic residues at the ligand binding site favor the binding of PPII-ligands (237); viz. *i)* the planar structure and near-parallel arrangement of such aromatic residues facilitate the formation of shallow grooves that are complementary to the base of the polyproline type II conformation of the ligand and *ii)* the bulky side chain of the aromatic residues offer a large van der Waals surface to contact with the ligand.

The structural studies of Src-SH3 domain bound to proline-rich ligands show that the ligand binding surface has three hydrophobic grooves (Fig. 1.11) (232). The XP dipeptide moieties of PPII helix (X=any hydrophobic residue and P=Proline)(229) fit onto the first two binding grooves (Fig. 1.11). The third site contains a conserved acidic residue (aspartate-99/D99 in Src-SH3 domain), which forms a salt-bridge with a characteristic basic residue of the ligand that is three residues upstream of the conserved PXXP motif (Fig. 1.11). The location of basic residue helps to determine the binding orientation of the ligand (Arg in Src-SH3-ligands) and classifies SH3-ligands into two typical classes; viz. a) Class I ligands with an N terminal conserved Arginine (**R***XX**P**XX**P**) that binds in plus orientation, while b) Class II ligands with a C terminal conserved Arginine (**P**XX**P**X***R**), binds in minus orientation (229). This is a highly conserved mechanism in the SH3 domains of SFKs, in which an acidic residue analogous to D99 interacts with a basic residue (Arg or Lys) at the same position, in the N terminal orientation (238).

It has been established that the 4th and 7th proline residues of class I ligands (**R**XX**P***XX**P***) as well as the 3rd and 6th proline residues in class II ligands (**P***XX**P***X**R**) are crucial for maintaining the PPII conformation of the ligand for binding. The critical proline residues (marked with asterisks) are generally preceded or followed by a hydrophobic residue (preferably, L, I, V or A) since it is shown to be essential to make hydrophobic contacts with the binding grooves of the Src-SH3 ligand-binding surface (229), (233), (236). Mutational studies have confirmed that the ligands with hydrophobic residues preceding the critical proline show a higher affinity due to ideal packing of this residue on the SH3 binding grooves (229). The classic (the highest affinity) Src-SH3 ligands (e.g. RALPPLP) contain LP (Leucine-Proline) moieties, of which the extended side chain of Leucine intercalates into the two hydrophobic grooves of the SH3 domain effectively (229).

The PXXP motif is not entirely sufficient for maintaining the binding specificity between different SH3-ligands and SH3 domains, as numerous ligands contain PXXP motifs (239). One way to establish specificity, as demonstrated by SFKs, has been the differential selection of ligands based on the basic residue that binds to the specificity pocket (third binding site) of SH3 domain (that contains D99). However, a majority of SH3-ligands markedly increases their affinity and specificity through tertiary interactions (233) established with the charged residues of RT and/or n-Src loops (232), both of which are in a very close proximity to the ligand binding site of the SH3 domain. The ligand binds to a valley between these two loops allowing the charged residues to make tertiary contacts to increase the specificity (Fig.(s) 1.10-a, 1.11). Many studies to date

report that SH3 domains also bind to unconventional, non-PXXP motifs of different ligands as summarized in Table 1 (237). Apart from the hydrophobic contacts with the ligands that form *PPII*, 3_{10} or *alpha* helices that are crucial for high-affinity binding, (229), (240), (241), some SH3 domains interact with their ligands exclusively through tertiary electrostatic interactions (242). Some ligands establish multiple but discontiguous interactions that enhance the binding affinity of the ligand many folds over the typical PXXP interaction; for instance, selective recognition of HIV_Nef protein by Hck-SH3 domain is determined by hydrophobic interactions of isoleucine (Ile) residue in the RT loop of Hck-SH3. Although HIV_Nef contains a PxxP motif, which is involved in the interaction, the highest binding affinity has only been observed with the full length HIV_Nef that establishes tertiary contacts with the RT loop of Hck-SH3 (243).

Table 1. The affinity and specificity of SH3-ligands (consensus data from literature, adapted from (237)). The upper-case letters = strong selection, lower-case letters = moderate selection. x = any amino acid. K_d= Dissociation Constant. NA=data not available.

The SH3 domain binding sequence (motif) of the ligand	Affinity (K_d, μM)	Ligand Structure	Reference(s)
(R/K)xxPxxP (class I)	1–200	PPII	(244), (229)
PxxPx(R/K) (class II)	1–200	PPII	(229)
RxxK	0.1–30	3^{10} helix	(240),
RKxxYxxY	20–60	NA	(245)
PxxDY	NA	NA	(246)
(R/K/G)XXPPGX(R/K)	10–200	PPII	(247)
R/K-rich	10–100	N/A	(248)
WQTDFEKLEKE	NA	α-Helix	(249)
RPSADLILNRCSESTKRKLAS A	2–13	α-Helix	(241)
PWTDQFEKLEKEVAEN…	NA	α-Helix	(250)
PX domain	50	PPII	(251)
SH2 domain	1–5	Tertiary	(252)
LIM domain	3000	Tertiary	(242)

Collectively, the above evidence suggests that SH3 domains can physically interact with a large number of diverse ligands but those with a higher affinity make better hydrophobic contacts with the binding grooves of the ligand binding surface of SH3 domain, whereas those with the lowest affinity make only tertiary contacts with the highly charged residues outside the ligand binding site. A few ligands that make both hydrophobic and tertiary contacts via multiple

interactions show the highest affinity. Nonetheless, all these interactions have been reported to be essential for assembly of numerous cellular signaling networks (253), (230), suggesting that ligand binding specificity plays a bigger role than affinity in SH3 domain interactions. As introduced in the next section, the SH2 domains select the ligands that are customized by phosphorylation dependent posttranslational modifications and thus, generally show a higher affinity to their ligands.

1.7.1.2. The SH2 Domain

The Src-homology 2 (SH2) domains are modules of ~100 amino acids that typically bind to phosphorylated tyrosine (p-Tyr) residues of the ligands (254), (255), (256). The p-Tyr residues serve as docking sites for SH2 domains of Src family kinases and various adapter proteins that trigger complex networks of cell-signaling pathways (257).

The SH2 domain of Src contains a central anti-parallel beta sheet surrounded by two alpha helices. The p-Tyr of the ligand generally binds as an extended beta strand that lies at right angles to the SH2 beta sheet (254) (Fig. 1.10-b). The range of binding affinities of SH2 domains (K_D = 4 – 500 nM) is higher than that of SH3 domains (K_D = 5 – 100 μM). The optimal high affinity phosphopeptides (K_D = 4 nM) interact with the Src-SH2 domain at six central residues, PQ(pY)EEI where p-Tyr and Ile (isoleucine) residues are tightly bound by two well-defined pockets on the ligand binding surface of SH2 domain (234). The Glu (glutamic acid) residues is crucial for binding as mutating either of this to Ala (alanine) greatly diminished the binding affinity (258). Conserved residues

such as Glu at the SH2 binding motif immediately adjacent to the phsphortyrosine contribute to the hydrophobic core or are involved in p-Tyr recognition while more variable C-terminal residues contribute to the specificity of the interaction (234). The Src-SH2 domain binding sequence of MUC1-CD has been identified as p-YEKV (139)

1.8. Interaction of MUC1 and Src

1.8.1 Evidence for MUC1-Src interaction mediated signaling in breast cancer

As discussed in section 1.6 above, our laboratory has demonstrated that ICAM-1 binding to MUC1 leads to Src dependent signalling cascades that initiate promigratory signalling in breast cancer cells. Treatment with a chemical inhibitor of the Src family kinases, PP2 significantly decreased the ICAM-1 binding induced phosphorylation of MUC1-CD and CrkL recruitment to MUC1-CD, in T47D breast cancer cells and MUC1 transfected-293T cells (150).

Other studies report that the inhibition of Src, in the MCF7 breast cancer cell line that overexpress MUC1, by dominant negative Src or siRNA (259) as well as small molecular inhibitors such as AZD0530 (173) or SKI-606 (260), directly affected integrin signaling pathways and actin-cytoskeletal dynamics, which resulted in reduced cell migration, adhesion and spreading (259). Src has been shown to interact with polyoma middle T-antigen transgenic mice (PyV-MT), and to play an integral role in MMTV-PyV MT-induced mammary tumorigenesis. When these mice were crossed onto a Muc1 null background there was a significant delay in tumor progression indicative of cooperative effects of MUC1 and Src (164).

However Src inhibition alone may not be responsible for the observed decrease in invasiveness and motility of breast cancer cells (194). As discussed in previous sections, MUC1 mediates some of these signaling events either by directly interacting with other molecules while already bound to Src (e.g. the MUC1/Src complex can increase β-catenin binding to MUC1 and promote cell invasion) (139) or by bringing Src closer to other molecules, such as FAK to form other signaling complexes (e.g. the Src/FAK complex may initiate several cancer cell migration pathways) (164).

Taken together, the above evidence suggests that MUC1-Src association plays a major role in pro-migratory signalling in breast cancer cells. However, in order to gain insights into the therapeutic-feasibility of blocking such interactions, it is crucial to explore the structural basis of the molecular recognition mechanism(s) involved in MUC1-Src interaction. The phosphorylation-dependent interaction of Src-SH2 domain and MUC1-CD is physiologically well established although the structural basis of the regulation is yet to be uncovered. Only a little is known about the direct recruitment of Src by MUC1-CD via the Src-SH3 domain.

1.8.2. Interaction of MUC1-CD with Src-SH3 domain

Evidence for the interaction of the Src-SH3 domain with MUC1-CD is based on *in vitro* GST pull down assays (139). Using purified His-tagged MUC1-CD and GST-Src-SH3 domain, Li *et al.* have assessed the direct binding of MUC1-CD to Src-SH3 domain using GST and GST-Src-SH3-De90/92 (a mutated Src-SH3 domain) as controls. The results revealed that MUC1/CD binds to wild-type Src-SH3 but not to the mutant Src-SH3 or GST (139). However, the specific

sequence motif(s) of MUC1-CD that interact with the Src-SH3 and the binding affinity or specificity have not been investigated.

The primary amino acid sequence shows that there are two putative SH3 binding motifs (RYVPPSS and PPSSTDR) in MUC1-CD that share some characteristics of typical SH3-ligands; *viz*. These putative motifs contain, i) prolines, serines and arginines that are indicative of a local PPII type structure, ii) conserved N and/or C terminal arginine residues, followed by at least one critical proline residues, **R***YV**P***PSS and P**P***SSTD***R** respectively (marked by asterisks), iii) a hydrophobic residue (V) preceding the 4^{th} proline in the N terminal orientation.

Since the SH3 domain and ligand interactions are highly promiscuous, the above assumptions that are based on the primary structure of MUC1-CD may not be entirely factual. There is a possibility that other/flanking residues or multiple discontiguous regions are responsible for the specificity of this interaction. The residue-specific details of the binding interface, in addition to the binding affinity, are thus essential to understand the MUC1-CD and Src-SH3 interaction.

1.9. Rationale, Hypothesis and Objectives

Rationale:

The recruitment of Src is a major functional step in ICAM-1 induced motility of MUC1 bearing breast cancer cells but the structural aspects of the molecular recognition events of MUC1 and Src are still unknown. Since the putative SH3 binding motif(s) of MUC1-CD are non-PXXP type, the interaction of MUC1-CD and Src-SH3 domain may be transient but specific. However, due to the fact that the SH2 and putative SH3 binding motifs of MUC1-CD are immediately adjacent to each other, binding of Src-SH3 domain to MUC1-CD may sterically hinder Src-SH2 domain binding (or vice versa). One possibility is rapid and transient interaction of MUC1-CD with Src-SH3 domain that may draw Src molecules close to MUC1-CD, allowing the subsequent phosphorylation of Tyr^{46} and binding of Src-SH2 domain to p-Y^{46}EKV motif. If disulphide-linked dimers of MUC1-CD are involved, as MUC1-CD is evidently flexible, Src-SH3 domain may bind to one dimer partner allowing the other partner to bind with the Src-SH2 domain at p-Y^{46}EKV.

The relative binding affinities and specificities of the putative SH3 binding motif(s) and phosphorylated SH2 binding motif of MUC1-CD may provide insights into the feasibility of direct recruitment of Src-SH3 domain vs. already established phosphorylation-dependent recruitment of Src-SH2 domain. Therefore, as the first step in understanding the interaction of MUC1 and Src, it is crucial to determine the binding affinity of the interaction of Src-SH3 with MUC1-CD and identify the MUC1-CD binding site on the Src-SH3 domain. These findings will then help to identify the consequences of Src-SH3 domain

binding in the regulation of Src-SH2 domain binding to MUC1-CD. Exploring the structural aspects of the interaction of MUC1-CD with Src-SH3 domain, therefore, is the focus of this thesis, which will provide a better understanding of this interaction for the first time.

Hypothesis:

The cytoplasmic domain of MUC1 (MUC1-CD) binds to Src-SH3 domain transiently through the putative N terminal SH3 binding motif, R^{34}YVPPSS

Objectives:

1: To map the binding site of MUC1-CD on the Src-SH3 domain using NMR spectroscopy and determine the binding affinity of the full-length cytoplasmic domain of MUC1 to the Src-SH3 domain.

2: To determine the differential binding affinities of monomeric and dimeric peptides of MUC1-CD and the Src-SH3 domain.

Chapter 2
EXPERIMENTAL METHODS AND RESULTS

2.1. Introduction

The structural studies of modular binding domains such as SH3 has been facilitated by their ability to fold independently into a stable three-dimensional modular structure, when isolated from the native protein (261). These modular domains also readily bind to short synthetic peptides *in vitro*, and thus, synthetic peptides have been identified as valuable tools for structural investigations of SH3 or SH2 domain interactions (261).

The putative SH3 binding motifs of MUC1 are non-PXXP type and may bind to Src-SH3 transiently but with a unique specificity if flanking residues are involved. Typically, the ligand binding affinities of SH3 domains fall in the micromolar range (K_D=1 – 100 μM) but those with K_D in the millimolar (mM) range can also be biologically important (242). Nuclear magnetic resonance (NMR) spectroscopy is a powerful analytical tool that is widely used to gain insight into such weak protein–ligand interactions in solution (262).

Unlike many other biophysical techniques available for protein-ligand binding studies, (e.g. Surface Plasmon Resonance (SPR), Isothermal Titration Calorimetry (ITC), Fluorescence Spectroscopy (FS)), which often provide only one or two types of information (e.g. kinetics and thermodynamics) about moderate to high affinity protein-ligand interactions, NMR spectroscopy can provide residue-specific, atomic-level details about a wide range of strong (K_D in the nanomolar range) to ultra weak (K_D in the millimolar range) molecular interactions (263), (262), (264), (265), (266); e.g. i) protein-ligand interface mapping (chemical shift mapping) and visualization of potential binding sites based on available 3D structures, ii) determination of the structure of protein-

ligand complexes, iii) structure-based mutational studies that determine the specificity and selectivity, iv) monitoring structural dynamics of an interaction, v) calculation of binding constants (kinetics) and deriving thermodynamic parameters, etc.

Given that specificity plays a major role in ligand binding to SH3 domains (237), binding affinity alone is not sufficient to distinguish between different SH3-ligand binding events. NMR spectroscopy has been particularly useful to this end, as it has produced data on site/residue-specific differences of a variety of SH3 domain and ligand interactions (229), (233), (267), (245), (241), (228), (240), (246).

2.11 Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR spectroscopy is based on a quantum property of atomic nuclei, called spin (I = spin angular momentum quantum number), which can be integral (I = 1, 2, 3… e.g. ^{2}H, ^{14}N etc.), fractional (I = 1/2, 3/2, 5/2…e.g. ^{1}H, ^{15}N, ^{13}C etc.) or zero (I = 0; e.g. ^{12}C, ^{16}O, ^{32}S etc.). Although NMR can detect isotopes with either integral or fractional nuclear spins, only those with a fractional nuclear spin of ½, provide sharp and meaningful resonance lines in NMR spectra (268).

The most commonly used nuclei for biomolecular NMR studies are ^{1}H, ^{15}N, ^{13}C isotopes, each of which has a spin of ½ and thus can have two possible energy states (+1/2 and -1/2) in a magnetic field, as determined by quantum mechanics (268). There are slightly more nuclei in the lower energy level than the higher level, proportional to the energy difference between two energy states (according to the Boltzmann distribution), which provides magnetization that is

measured in an NMR experiment (268). It is possible to excite the nuclei at lower energy state into the higher state with radio frequency electromagnetic radiation, where the frequency of radiation is determined by the difference in energy between these energy levels (268).

2.12 Chemical Shift Mapping

During an NMR experiment, the electrons of a spinning nucleus shield the static magnetic field. Therefore, the nuclei (protons) absorb electromagnetic radiation at different frequencies depending on the electron density around the nucleus. These shifts in resonance frequencies are called chemical shifts (268). Since the chemical shift (δ) is changed based on the static magnetic field strength, these values are standardized using a reference compound such as DSS (2,2-dimethyl 2-silapentane 4-sulphonate). Due to low electronegativity of the silicon atom, the nine identical methyl protons in DSS are highly sheilded resulting a high intensity proton signal at the most upfield position of an NMR spectrum and thus can be easily used to assign the chemical shift to zero. If the frequency of a given resonance line is v and the frequency of the line from DSS is v_{dss}, the chemical shift is given as, $\delta = (v - v_{dss})/(v_{dss} \times 10^6)$. The chemical shift is expressed as parts per million (*ppm= 1/1000,000*) in frequency since the numerator of this equation represents the frequency of resonance lines (v and v_{dss}) in hertz but the denominator, which is the operating frequency of a given spectrometer, is in megahertz (1 MHz= 10^6 Hz). Measuring chemical shift in *ppm* makes all frequencies scale with the applied magnetic field and thus independent of the static magnetic field strength (268).

The atoms in a well-folded protein experience many different chemical environments due to different inter-atomic interactions, while most of the atoms in an unfolded molecule experience similar chemical environments. The residues at the binding site of a protein and/or a ligand that make strong contacts with each other also experience different chemical environments. This conformational heterogeneity produces clearly distinguishable peaks in NMR spectra in contrast to mostly overlapping peaks generated by the atoms of an unstructured molecule or a low-affinity molecular interaction interface.

The NMR titration or chemical shift perturbation experiments that are based on Heteronuclear Single Quantum Coherence (HSQC) NMR spectroscopy are commonly used to monitor the changes that occur in the chemical environment of the binding partners (chemical shift mapping), at a residue specific level, during a protein-ligand interaction (269). The two dimensional (2D) ^{1}H-^{15}N HSQC spectrum provides a quick diagnostic fingerprint of a uniformly ^{15}N-labelled protein as the backbone amide group of each non-proline residue creates a single ^{1}H-^{15}N cross peak (263). The pattern of dispersion, intensity, and the number of observed cross peaks in a 2D ^{1}H-^{15}N HSQC spectrum are directly correlated with the chemical shift heterogeneity (folded or unfolded state) of the protein as well as with the overall sample quality (270).

Overlaying a series of 2D ^{1}H-^{15}N HSQC spectra obtained at consecutive titration points shows the chemical shift changes that occur in a protein in response to binding of increased amounts of a particular ligand. The dependence of chemical shift changes on the ligand concentration facilitates calculation of the dissociation constant of the protein-ligand complex (271).

In the current study, the ^{15}N labeled Src-SH3 domain was used to monitor the ^{1}H-^{15}N amide chemical shift changes upon binding of MUC1-CD peptides. Although the stable ^{15}N isotope, which has a nuclear spin of ½ simplifies its observation by NMR (272), its natural abundance is only 0.37%. In order to obtain strong signals and clear cross peaks in a 2D ^{1}H-^{15}N HSQC experiment, it is necessary to label the nitrogen atoms in the protein with the ^{15}N isotope. This requires recombinant expression of the protein in the presence of a ^{15}N isotope-enriched chemical compound.

2.13 The GST Gene Fusion System

The GST (Glutathione-S-Transferase) gene fusion system was chosen for expression and purification of ^{15}N isotope labeled Src-SH3/SH2 domains since the GST tag promotes the solubility of a recombinant protein and allows easy purification by affinity chromatography (using Glutathione Sepharose 4B) and mild elution conditions, which help to preserve the native fold of the expressed protein (273). All GST fusion vectors contain a *tac* promoter, open reading frame encoding GST, a protease cleavage site followed by restriction endonuclease sites (BamHI, SmaI, and EcoRI in PGEX-2T vector) and termination codons (Fig 2.1). The *tac* promoter allows chemical induction of protein expression with isopropyl-β-D-thiogalactopyrenoside (IPTG). The protease cleavage site (thrombin cleavage site in PGEX-2T) facilitates cleavage of the target protein from the GST tag.

2.2. Materials and Methods

2.2.1. Subcloning of Human Src cDNA

The complementary DNAs (cDNAs) encoding human c-Src-SH3, SH2, and SH3+SH2 (SH3-linker-SH2) domains were subcloned from a recombinant pUASEMP_Src plasmid, into the pGEX-2T, GST fusion vector (Fig 2.1), in between EcoRI and BamHI sites. The recombinant human Src plasmid (pUASEMP_Src) was a generous gift from Dr. Tony Pawson, (Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Ontario). The PGEX-2T vector was a donation from Mr. Gary Ritzel (Department of Biological Sciences, University of Alberta).

The PCR primers with 6 flanking bases were designed to amplify SH3 and SH2 domains of Src for insertion into the pGEX-2T vector, as follows (Blue=EcoRI site; Red =BamHI site; F=Forward; R=Reverse);

SH3_F => 5'- GGC CCG GGA TCC ATG GGT GGA GTG ACC ACC TTT - 3'

SH3_R => 5'- GCG CCG GAA TTC TTA GGA GTC GGA GGG CGC CAC - 3'

SH2_F => 5'- GCG CGG GGA TCC ATG TGG TAT TTT GGC AAG ATC - 3'

SH2_R => 5'- GCG CCG GAA TTC TTA GCA CAC GGT GGT GAG GCG - 3'

The recombinant plasmids were sequenced to verify the insertion of the correct coding sequence with reference to Swiss-Prot sequence for human c-Src (Accession No. P12931).

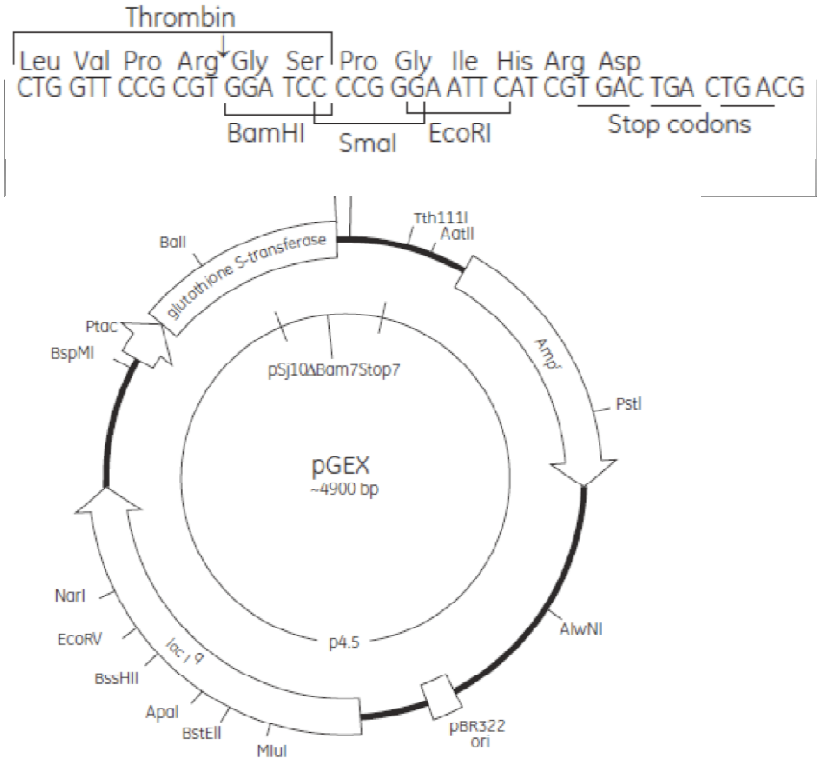


Figure 2.1. Map of the glutathione S-transferase (GST) fusion vector showing the open reading frame of PGEX2T, GST followed by thrombin cleavage site, the restriction sites and stop codons (274).

2.2.2. Expression and Purification of ^{15}N labeled Src-SH3 Domain

To produce expression clones, the recombinant PGEX-2T-SH3, SH2 and SH3+SH2 plasmids were transformed into BL21 (DE3) PlySs cells (*E. coli* host). The recombinants were expressed in M9 minimal media enriched with $^{15}NH_4Cl$ to a cell density (Optical Density at λ^{600}) of 0.6 - 0.8 and induced with 1.0 mM IPTG (Appendix 1). As determined by a series of test expressions, growing *E.coli* for 3-4 hrs after induction was sufficient to get a good yield of target protein (1 - 2 mg/ml per Litre) with minimal expression of untargeted proteins. Accordingly,

the cells were harvested after 3-4 hours by centrifugation and the cell pellets were lysed according to the standard protocols (Appendix 2) using an Emulsiflex (available at Dr. Joel Weiner's lab, Department of Biochemistry, University of Alberta). Tris-Tricine SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) was used to analyze the cell lysates (Appendix 3). The cell lysates were then purified using GST affinity columns, cleaved with thrombin and desalted using a Sephadex-G25 column with 10 mM NH_4HCO_3. The SH3 domain (6.88 KDa) and GST (26 KDa) were then separated by size exclusion chromatography and desalted again. The pure SH3 domain was lyophilized (Appendix 4).

The pre-packed GSTrapFF-1ml columns (GST tagged protein purification columns), Glutathione Sepharose 4B (GS4B) beads, Thrombin, Superdex-75 (for size exclusion chromatography), and 15 ml glass chromatography columns were purchased form Pharmacia/GE healthcare. The Sephadex-G25 desalting columns, peristaltic pumps, size exclusion chromatography columns, lyophilization Jars and all other equipment necessary for protein purification were generously provided by Dr. Brian Sykes at the Department of Biochemistry, University of Alberta.

2.2.3. Design of MUC1 Peptides and Synthesis of Full-length MUC1-CD

Native, Mutant and Dimer Peptides of MUC1-CD:

A native 23-residue synthetic MUC1-CD peptide (Fig. 2.2) that contain both SH2 and putative SH3 motifs was designed to study direct binding of the Src-SH3 domain. Control peptides (23-residue) with the same sequence but with point mutations, $R^{34}A$, $R^{43}A$ and $P^{37}A$, were designed to test the putative terminal

arginines and the critical proline, respectively (Fig. 2.2). A synthetic 48-residue-dimer MUC1 peptide (two 23-residue peptides linked via 2 cysteine residues, with one arm phosphorylated at Y^{46}) was designed for comparison of SH3 and SH3+SH2 domain binding studies of 23-residue monomer vs. a dimer. All peptides were obtained through the IBD (Institute of Biomolecular Design), University of Alberta.

Full Length MUC1-CD:

The recombinant, full length, 72-residue His^6 tagged MUC1-CD was obtained from GenScript Inc. The SDS-PAGE analyses showed monomers, spontaneous dimers (and oligomers) that were verified by in gel digestion/mass spectrometry and MALDI-TOF analyses (Fig. 2.3), suggesting that 72-residue His^6 tagged MUC1-CD may not be suitable to distinguish the interaction between monomeric MUC1-CD and Src-SH3 domain. Therefore, a 69-residue full-length synthetic peptide of MUC1, without the CQC motif, was obtained (GenScript Inc.).

The amino acid sequences of all peptides were checked against the UniProtKB/Swiss-Prot sequence for human MUC1 (Acc No. P15941).

Rationale for comparisons of binding data based on the MUC1 peptides

1) **23-residue monomer peptide vs. the 69-residue full-length monomer peptide**: - If the 23-residue peptide does not bind to Src-SH3, the binding site could be located somewhere else or multiple discontiguous regions of MUC1-CD may be involved in binding. To rule out this possibility, a 69-residue full length MUC1-peptide was used to compare the binding affinity and chemical shift perturbations of the residues at each binding event.

2) **23-residue monomer peptide vs. the 48-residue dimer peptide**: - The 48-residue-dimer is a non-natural peptide. It was used only for the purpose of testing whether any type of dimerization at the vicinity of the putative binding site causes a significant change in chemical shift perturbations (differential binding), compared to a native, monomeric peptide with the same residues (23-residue monomer).

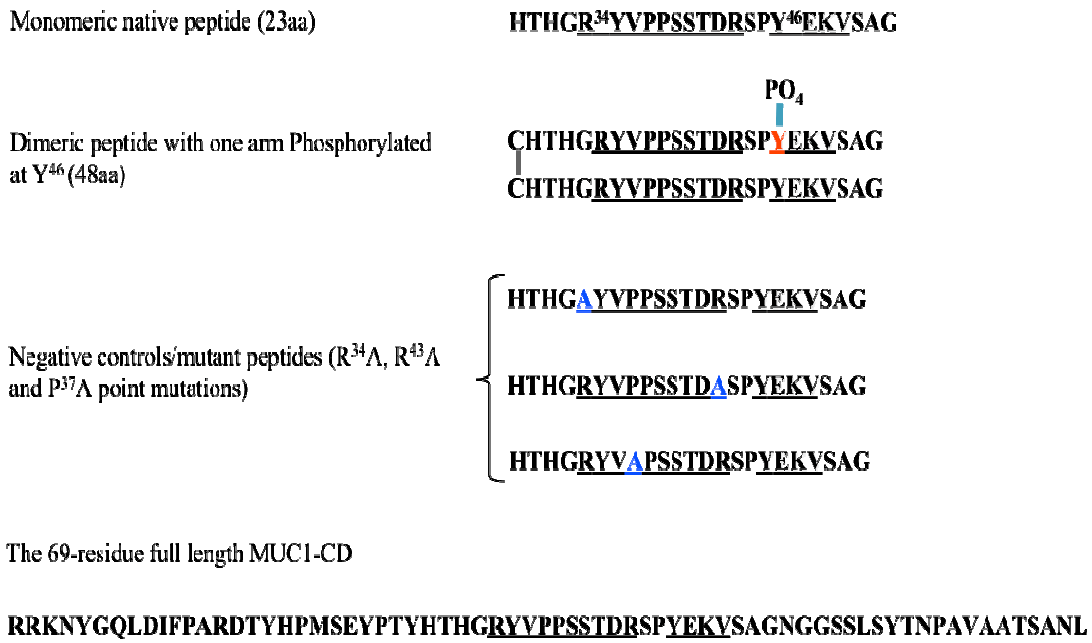


Figure 2.2. The synthetic peptides of MUC1-CD that were designed for NMR titrations with ^{15}N labeled Src-SH3 and SH3+SH2 domains. The putative SH3 binding motif and the SH2 binding motif are underlined (blue = mutated residues; red = phosphorylated Y^{46}; PO_4 = phosphate group). The 23-residue native, mutant and 48-residue dimer peptides were obtained through Institute of Biomolecular Design (IBD), University of Alberta and the 69-residue MUC1-CD (without CQC) was obtained from GenScript (Inc).

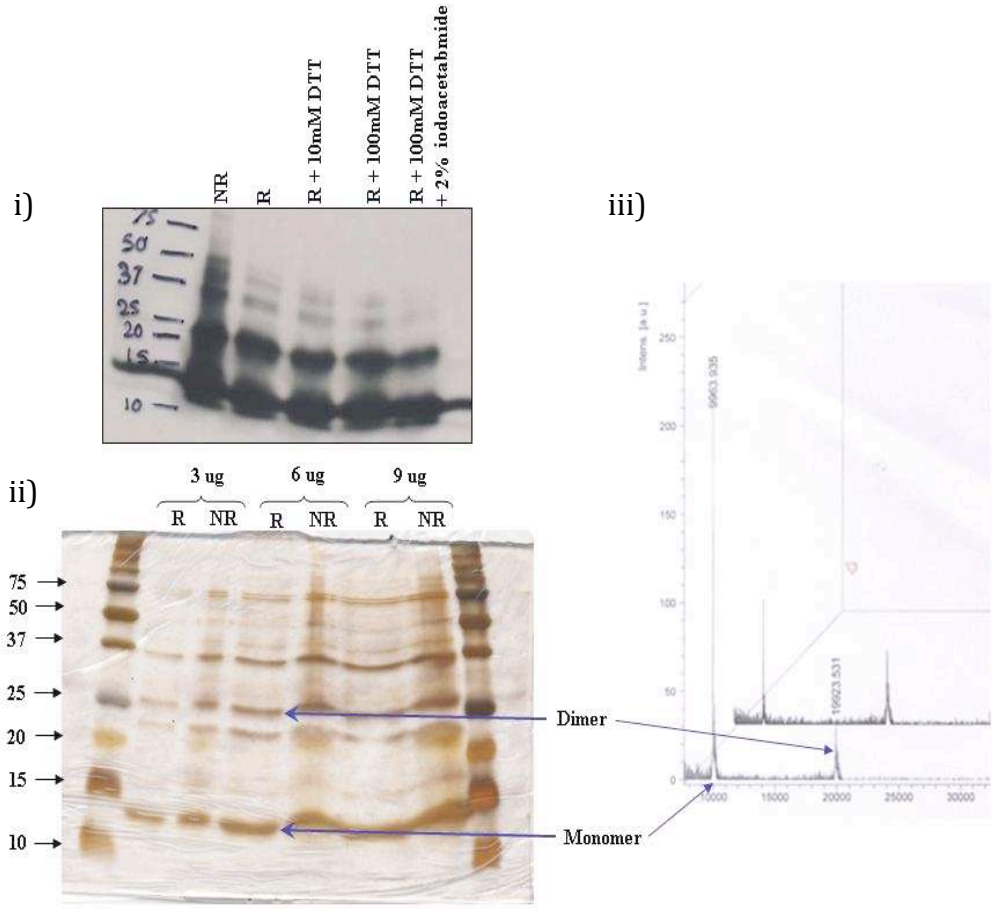


Figure 2.3. The SDSPAGE analyses of recombinant 72-residue His^6 tagged MUC1-CD (GenScript). ***i)*** Western blot probed for CT2 (cytoplasmic tail-2) antibody (recognize the last 15 residues of MUC1-CD) that shows robust spontaneous dimers formed *in vitro*. (NR=Non reduced sample; R = reduced sample (BME + boiling for 5 minutes)); BME=beta mercaptoethanol; DTT=Dithiothreitol. The alkylating agent, 2% iodoacetamide was added to prevent reformation of disulfide bridges. ***ii)*** a silver stained gel of His^6 tagged MUC1-CD showing monomers and dimers confirmed by in gel digestion-mass spectrometry. ***iii)*** The MALDI-TOF (Matrix Assisted Laser Desorption Ionisation-Time of Flight) analysis of full length MUC1-CD showing covalently linked monomers/dimers (*x axis*:- mass (m/z) and *y axis*:- peak intensity).

2.2.4. NMR Titrations of Src-SH3 Domain with MUC1-CD Peptides

The NMR titrations of Src-SH3 domain with i) a native 23-residue MUC1-CD peptide, ii) a $R^{34}A$ mutant 23-residue MUC1-CD peptide, iii) a 69-residue full length MUC1-CD and, iv) a 48-residue-dimer MUC1-CD peptide (Tables, 2-5) were carried out as described below.

Calculated amounts of MUC1 peptide aliquots were titrated into 500 µl of 0.25 mM ^{15}N labeled Src-SH3 domain in a high-salt phosphate NMR buffer [50 mM Na_2HPO_4, 100 mM NaCl, 10 µM EDTA, 1 mM Imidazole and 4.6 mM DSS (containing 0.196% NaN_3 and 98% v/v D_2O)]. The buffer was chosen according to previous NMR studies that addressed ligand binding to the Src-SH3 domain (233), (229). The pH of the protein solution was initially adjusted to 6.8 but it was in the range of 6.5 - 6.7, according to the internal pH indicator, imidazole. The pH was maintained at a fairly constant value throughout the titration by adding 0.5 – 3.0 µl of 1 M NaOH, where necessary, if the solution became more acidic with addition of the peptide aliquots. The initial protein and peptide concentrations were determined by weight/volume method but confirmed and adjusted by using subsequent amino acid analyses (Institute of Biomolecular Design, University of Alberta). Changes in protein concentration due to the addition of peptide were recorded at each titration point and included in the calculation of subsequent volume of peptide solution (Tables 2-5).

Two-dimensional (2D) ^{1}H-^{15}N HSQC NMR spectra of MUC1-CD bound and unbound Src-SH3 domain were acquired at each titration point of MUC1 peptide at a constant temperature (30^{0} C). The published ^{1}H and ^{15}N chemical shifts of the Src-SH3 (275) available through Biological Magnetic Resonance

Data Bank (BMRB) were used to assign the HSQC spectrum. The NMR experiments and data analyses were carried out in collaboration with Dr. Brian Sykes (Department of Biochemistry, University of Alberta) by using the Varian 500 or 600 MHz NMR spectrometers in his lab.

Table 2. The NMR titration of 23-residue MUC1-CD peptide wih Src-SH3 domain (* = pH adjustments)

Titration points	Volume of 20 mM MUC1 peptide solution added (ul)	Calculated MUC1 peptide Concentration (mM)	Calculated Volume of SH3 domain (µl)	Calculated. SH3 domain Concentration	Molar Ratio [MUC1]/ [SH3]	pH	File names of 1D NMR experiments	File names of 2D ^{1}H-^{15}N HSQC NMR Experiments
0	0.00	0.00	500.0	0.160	0.0	6.7	SH3_S2_1D_01	SH3_S2_Nhsqc_01
1	8.0	0.16	508.0	0.157	1.0	6.7	SH3_S2_1D_02	SH3_S2_Nhsqc_02
2	7.6	0.31	515.6	0.155	2.0	6.6	SH3_S2_1D_03	SH3_S2_Nhsqc_03
3	7.7	0.46	523.4	0.153	3.0	6.6	SH3_S2_1D_04	SH3_S2_Nhsqc_04
4	7.3	0.60	530.7	0.151	4.0	6.6	SH3_S2_1D_05	SH3_S2_Nhsqc_05
5	8.0	0.75	538.6	0.149	5.0	6.6	SH3_S2_1D_06	SH3_S2_Nhsqc_06
6	7.00	0.88	545.6	0.147	6.0	6.5	SH3_S2_1D_07	SH3_S2_Nhsqc_07
7	7.1	1.01	552.7	0.145	7.0	6.53*	SH3_S2_1D_08	SH3_S2_Nhsqc_08
8	7.2	1.14	559.9	0.143	8.0	6.52*	SH3_S2_1D_09	SH3_S2_Nhsqc_09
9	7.3	1.27	567.2	0.141	9.0	6.54*	SH3_S2_1D_10	SH3_S2_Nhsqc_10
10	6.81	1.39	574.0	0.139	10.0	6.55*	SH3_S2_1D_11	SH3_S2_Nhsqc_11
11	33.9	1.98	607.9	0.132	15.0	6.5	SH3_S2_1D_12	SH3_S2_Nhsqc_12
12	31.6	2.50	639.5	0.125	20.0	6.5	SH3_S2_1D_13	SH3_S2_Nhsqc_13

Table 3. The NMR titration of 23-residue R34A mutant MUC1 peptide wih Src-SH3 domain (AAA=Amino Acid Analysis)

Titration points	Volume of 15.4 mM MUC1 peptide solution added (µl)	Calculated MUC1 peptide Concentration (mM)	Calculated Volume of SH3 domain (µl)	Calculated. SH3 domain Concentration	Molar Ratio [MUC1]/ [SH3]	Estimated [MUC1]/[SH3] Molar Ratio (before AAA)	pH	File names of 1D NMR experiments	File names of 2D ^{1}H-^{15}N HSQC NMR Experiments
0	0.00	0.00	500.0	0.190	0.0	0.0	6.68	SH3_S3_1D_01	SH3_S3_Nhsqc_01
1	5.3	0.16	505.3	0.188	0.9	1.0	6.67	SH3_S3_1D_02	SH3_S3_Nhsqc_02
2	5.3	0.32	510.6	0.186	1.7	2.0	6.66	SH3_S3_1D_03	SH3_S3_Nhsqc_03
3	5.1	0.48	515.7	0.184	2.6	3.0	6.65	SH3_S3_1D_04	SH3_S3_Nhsqc_04
4	4.9	0.62	520.6	0.182	3.4	4.0	6.64	SH3_S3_1D_05	SH3_S3_Nhsqc_05
5	5.0	0.77	525.5	0.181	4.3	5.0	6.63	SH3_S3_1D_06	SH3_S3_Nhsqc_06
6	4.73	0.91	530.3	0.179	5.1	6.0	6.62	SH3_S3_1D_07	SH3_S3_Nhsqc_07
7	5.0	1.06	535.3	0.177	5.9	7.0	6.61	SH3_S3_1D_08	SH3_S3_Nhsqc_08
8	4.8	1.19	540.1	0.176	6.8	8.0	6.60	SH3_S3_1D_09	SH3_S3_Nhsqc_09
9	5.1	1.34	545.2	0.174	7.7	9.0	6.58	SH3_S3_1D_10	SH3_S3_Nhsqc_10

Table 4. The NMR titration of 69-residue full length MUC1-CD peptide with Src-SH3 domain (AAA=Amino Acid Analysis)

Titration points	Volume of 9.0 mM MUC1 peptide solution added (µl)	Calculated MUC1 peptide Concentration (mM)	Calculated Volume of SH3 domain (µl)	Calculated. SH3 domain Concentration	Molar Ratio [MUC1]/ [SH3]	Estimated [MUC1]/[SH3] Molar Ratio (before AAA)	pH	File names of 1D NMR experiments	File names of 2D ^{1}H-^{15}N HSQC NMR Experiments
0	0.00	0.00	500.0	0.075	0.0	0.0	6.52	M_1D-01	M_15NHSQC-01
1	12.5	0.23	512.5	0.073	3.1	1.0	6.51	M_1D-02	M_15NHSQC-02
2	23.8	0.64	536.3	0.070	9.2	3.0	6.42	M_1D-03	M_15NHSQC-03
3	33.4	1.20	569.7	0.066	18.3	6.0	6.40	M_1D-04	M_15NHSQC-04
4	31.9	1.71	601.6	0.062	27.4	9.0	6.37	M_1D-05	M_15NHSQC-05
5	20.0	2.01	621.6	0.060	33.3	10.0	6.36	M_1D-06	M_15NHSQC-06

Table 5. The NMR titration of 48-residue dimer MUC1 peptide with Src-SH3 domain (AAA=Amino Acid Analysis)

Titration points	Volume of 10.9 mM MUC1 peptide solution added (µl)	Calculated MUC1 peptide Concentration (mM)	Calculated Volume of SH3 domain (µl)	Calculated. SH3 domain Concentration	Molar Ratio [MUC1]/ [SH3]	Estimated [MUC1]/[SH3] Molar Ratio (before AAA)	pH	File names of 1D NMR experiments	File names of 2D ^{1}H-^{15}N HSQC NMR Experiments
0	0.00	0.00	500.0	0.075	0.0	0.0	6.52	D_1D-01	D_15NHSQC-01
1	12.5	0.27	512.5	0.073	3.7	1.0	6.46	D_1D-02	D_15NHSQC-02
2	23.8	0.78	536.3	0.070	11.2	3.0	6.37	D_1D-03	D_15NHSQC-03
3	22.5	1.24	558.8	0.067	18.4	5.0	6.35	D_1D-04	D_15NHSQC-04
4	32.8	1.88	591.6	0.063	29.6	8.0	6.33	D_1D-05	D_15NHSQC-05
5	24.0	2.32	615.6	0.061	38.0	10.0	6.3	D_1D-06	D_15NHSQC-06

2.2.5. Analyses of NMR spectral data

All ^{1}H-^{15}N HSQC NMR spectra acquired during a particular titration were equally processed using VNMRJ software and then overlayed using a VNMRJ script. Each *fid* (free induction decay) was processed using an NMRPipe script (*fid_nmrview.com*) (276), in order to convert it into a **.nvj* file that was accessible by NMRViewJ software, which was then used to create peaklists of the assigned residues. A peaklist of a particular ^{1}H-^{15}N HSQC spectrum contained the chemical shift values corresponding to all peaks of that spectrum. The observed chemical shift changes per residue ($\Delta\delta$) in ^{1}H-^{15}N HSQC spectra of SH3 is calculated by the software, according to the equation, $\Delta\delta=[(\Delta\delta H)^{2}+(\Delta\delta N/5)^{2}]^{½}$, where $\Delta\delta H$ is the chemical shift change in ppm in ^{1}H dimension and $\Delta\delta N$ is the chemical shift change in ppm in ^{15}N dimension. The coefficient of 0.2 is applied in the equation to compensate for the scaling differences between ^{15}N and ^{1}H chemical shifts. The peaklists from each consecutive HSQC spectrum was concatenated. The combined data were used to generate an input file for the *Xcrvfit* software (277) using the script, *xpk-to-xcrvfit-nhsqc.pl* (Olivier Julien, Department of Biochemistry, University of Alberta), that extracts and lists all chemical shift values per residue along with the corresponding molar concentrations of peptide and protein (used to calculate the molar ratio of [peptide]: [protein]).

2.2.6. Determination of dissociation constants of MUC1/Src-SH3 interaction

The interaction of MUC1 and Src-SH3 domain was assumed to follow a single site binding model with 1:1 stoichiometry given by, P+L↔PL (P=protein;

L=ligand). The *Xcrvfit* software (277) was used to fit an appropriate binding model for the observed chemical shift data based on all residues as well as those residues with total chemical shift >0.04 ppm (the most mobile residues in terms of chemical shifts), selected based on literature (reviewed in (263)), which suggested a threshold value of 0.04 ppm is considered to indicate that the corresponding amino acid is involved in binding to the ligand. The software calculates dissociation constant (K_D) based on the equilibrium constant, the rate of formation (k_{on}) and dissociation (k_{off}) of the molecular complex ($K_D = k_{off}/k_{on}$) using chemical shift changes ($\Delta\delta$) of SH3 backbone amide protons as a function of MUC1-CD peptide concentration (details in Appendix 5).

2.2.7. Chemical Shift Mapping of the MUC1-CD Binding Site on Src-SH3 Domain

The 3D (three dimensional) NMR solution structure of Src-SH3 domain (1QWE) and the X-ray Crystal Structure of inactive Src molecule (2SRC) available through RSCB (Research Collaboratory for Structural Bioinformatics) protein data bank (*pdb*) (http://www.pdb.org/pdb/home/home.do) were used for chemical shift mapping. The residues of Src-SH3 domain that showed the highest chemical shift changes (>0.04 ppm) upon addition of MUC1-peptides were then mapped onto the surface of the 3D structures of Src-SH3 domain and inactive Src using the software, *MacPyMOL*.

2.3 Results

2.3.1. Expression and Purification of uniformly ^{15}N labeled Src-SH3 Domain

The ^{15}N labelled GST tagged Src-SH3 (32 KDa), Src-SH2 (37 KDa) and Src-SH2+SH3 (44.7 KDa) domains were cloned into PGEX-2T vector and successfully expressed using BL21-DE3-PlySs competent cells (Fig. 2.4) in a modified M9 minimal media with Thiamin-HCl and Biotin, as determined by a series of test expressions. The GST tagged Src domains were cleaved properly (Fig. 2.5-i and ii) except for the SH2 domain (Fig. 2.5-iii). The pure proteins were positively identified by mass spectrometry.

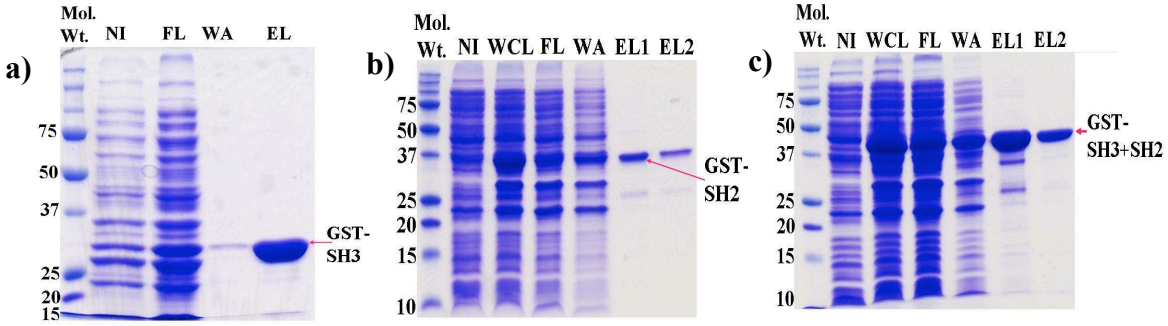


Figure 2.4 The expression and purification of GST tagged Src-SH3, SH2 and combined (SH3+SH2) domains. The ^{15}N labeled purified GST-Src a) SH3 (32 KDa), b) SH2 (37 KDa) and c) SH3+SH2 (44.7 KDa) domains expressed in E. coli BL21 (DE3) PLysS cells. The lanes are, non-induced (NI), whole cell lysate (WCL), flow-through (FL), wash (WA) and eluted (EL) fractions of GST-Src lysates passed through a GSTrapFF™ column.

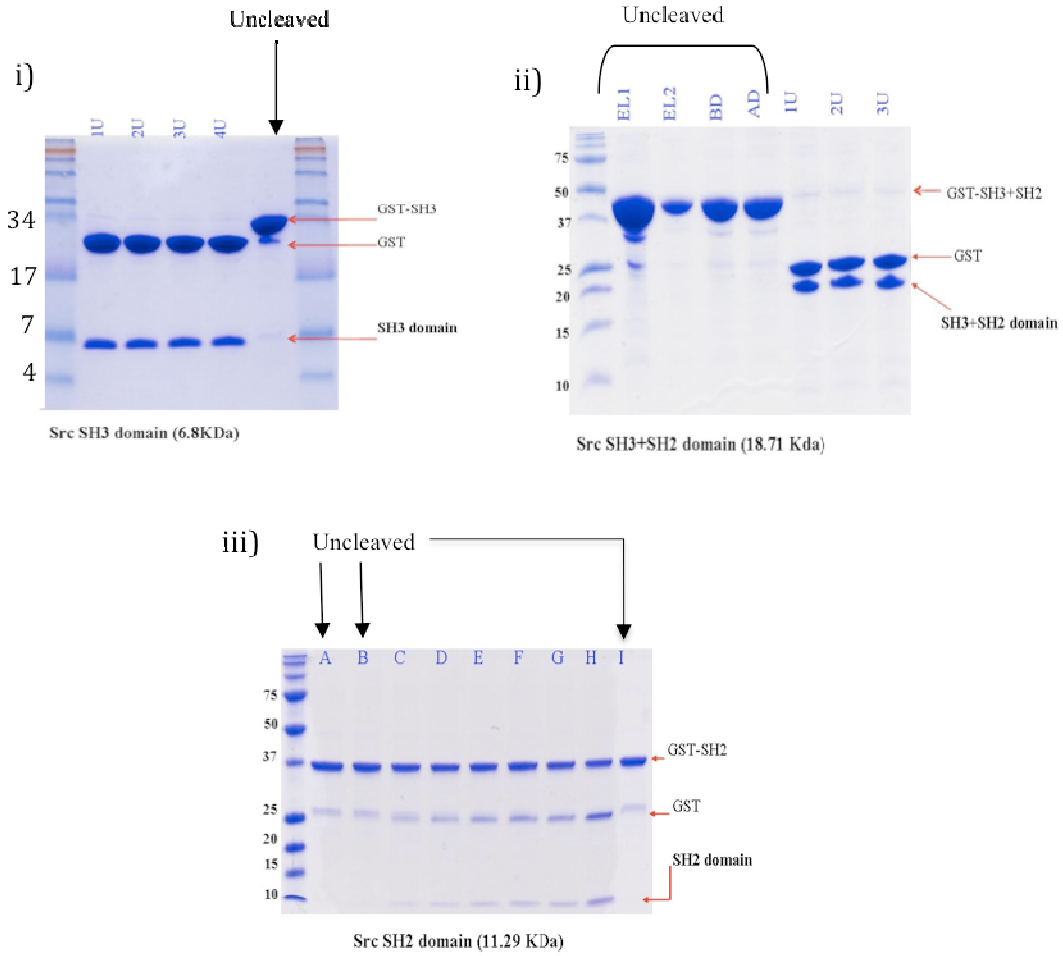


Figure 2.5. Thrombin cleavage of ^{15}N labeled GST-tagged Src domains; ***i)*** SH3 domain (6.88 KDa)- cleavage with 1 – 4 Units (U) of thrombin (16 hr. incubation), ***ii)*** SH3+ SH2 domain (18.71 KDa) cleavage with 1 -3 units of thrombin (16 hr. incubation); eluted fractions (EL1 and EL2), protein sample before dialysis (BD) and after dialysis (AD) and ***iii)*** SH2 domain (11.29 KDa) cleavage time course with one unit of thrombin; A = before dialysis; B=after dialysis; C – H = 1, 2, 3, 4, 5, and 16 hrs incubation with 1unit of thrombin; I = control.

2.3.2. The ^{1}H-^{15}N HSQC NMR Spectra of MUC1-CD/Src-SH3 Interaction

The 2 dimensional (2D) ^{1}H-^{15}N HSQC NMR spectrum of the Src-SH3 domain showed well resolved, assignable peaks with few ambiguities (Fig. 2.6-a). It was assigned using the published ^{1}H and ^{15}N chemical shift values available through BMRB (275). All non-proline residues except, G84, S126, S145 and T88 (Fig. 2.6-b), were assigned referring to the published chemical shift values.

The overlay of ^{1}H-^{15}N HSQC spectra acquired during the NMR titrations of, i) 23-residue native peptide, ii) 69-residue full-length MUC1-CD peptide and iii) 48-residue-dimer peptide, all showed the same residue shifts, of which the changes in chemical shift ($\Delta\delta$) were very small (<0.1 ppm) (Fig.(s) 2.7 and 2.8). The overlayed HSQC spectrum of the 23-residue peptide was compared against the 69-residue full length and 48-residue-dimer peptides based on the rationale outlined in the section 2.2.3. The residues with the highest chemical shift changes, ($\Delta\delta$ >0.04 ppm) in all three sets of overlayed HSQC spectra were R98, E100, H125, T132, G130, Y134 and L103. The other residues that showed minor changes in chemical shift were, E118, N138 and S97.

The only differences among three sets of HSQC spectra were, i) the residues, T99 and Q112 did not show chemical shift perturbations upon addition of the 69-residue full length MUC1 peptide whereas a small change in chemical shift was detected after adding 23-residue monomeric and 48-residue dimeric peptides, in separate titrations; ii) the direction of chemical shift change in D99 (labelled as D102 in this study) was clearly different in the HSQC spectra obtained for the titration of 69-residue peptide compared with that of the other two titrations (23-residue and 48-residue peptides).

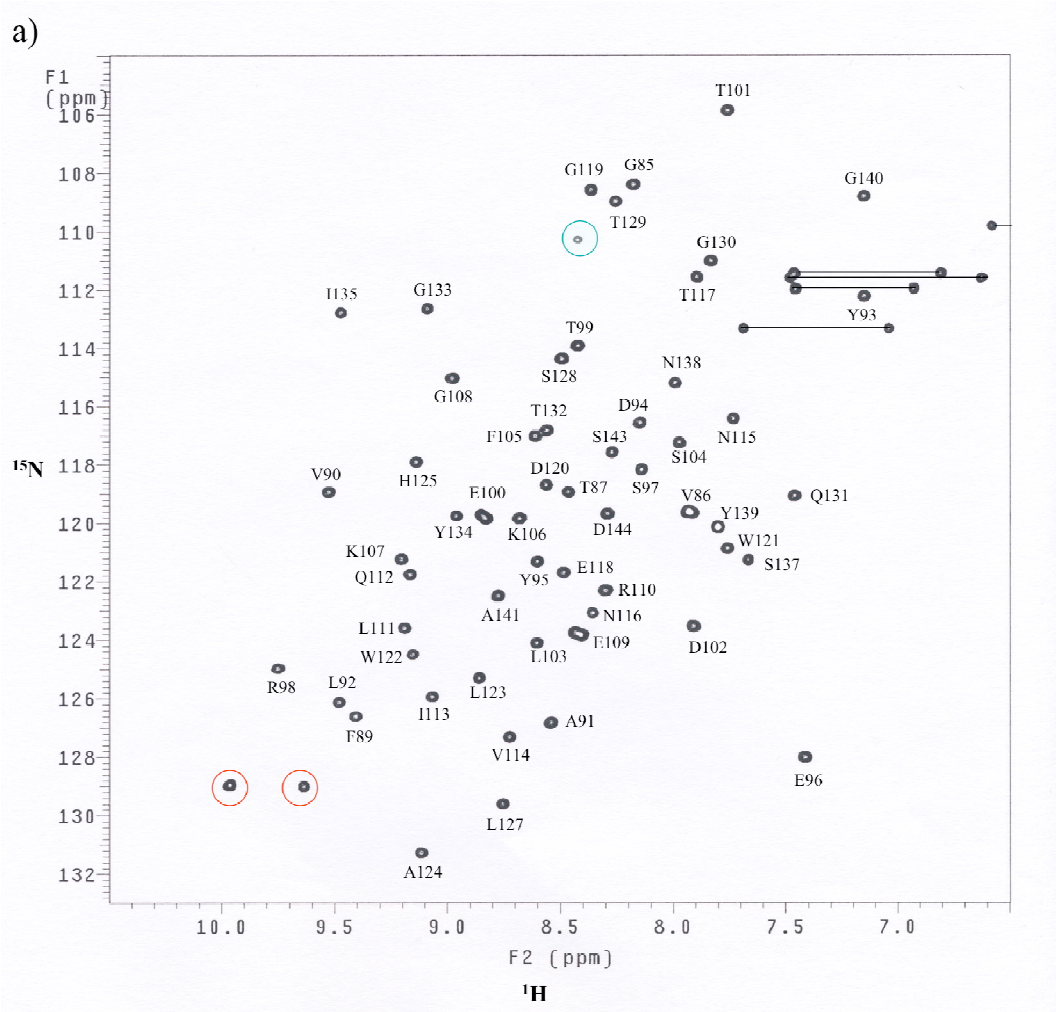


Figure 2.6. ***a)*** The 2D ^{1}H-^{15}N HSQC NMR spectrum of the Src-SH3 domain. Residues were assigned based on the published ^{1}H and ^{15}N chemical shift values available through BMRB (275). Bars connect the side chain amide protons (N-H) of Asparagine (N) and Glutamine (Q) residues. The side chain amide protons of two tryptophan residues are circled in red. Unidentified peaks are circled in green. ***b)*** The amino acid sequence of the Src-SH3 domain. Assigned residues are in red. Unassigned residues are in blue. The first residue of the protein sequence and the proline residues that do not produce peaks in ^{1}H-^{15}N HSQC spectra are in black.

The overlay of ^{1}H-^{15}N HSQC spectra of Src-SH3 domain acquired during the titration of $R^{34}A$ mutant MUC1 peptide showed a reduction in chemical shift changes in some of the same residues (E100, H125, G130, Y134) that were perturbed during the titration of 23-residue peptide with the native sequence (Fig. 2.9). Notably, there was no difference in the chemical shift perturbations of R98 in the mutant peptide titration compared to that of the native peptide titration (Fig. 2.9).

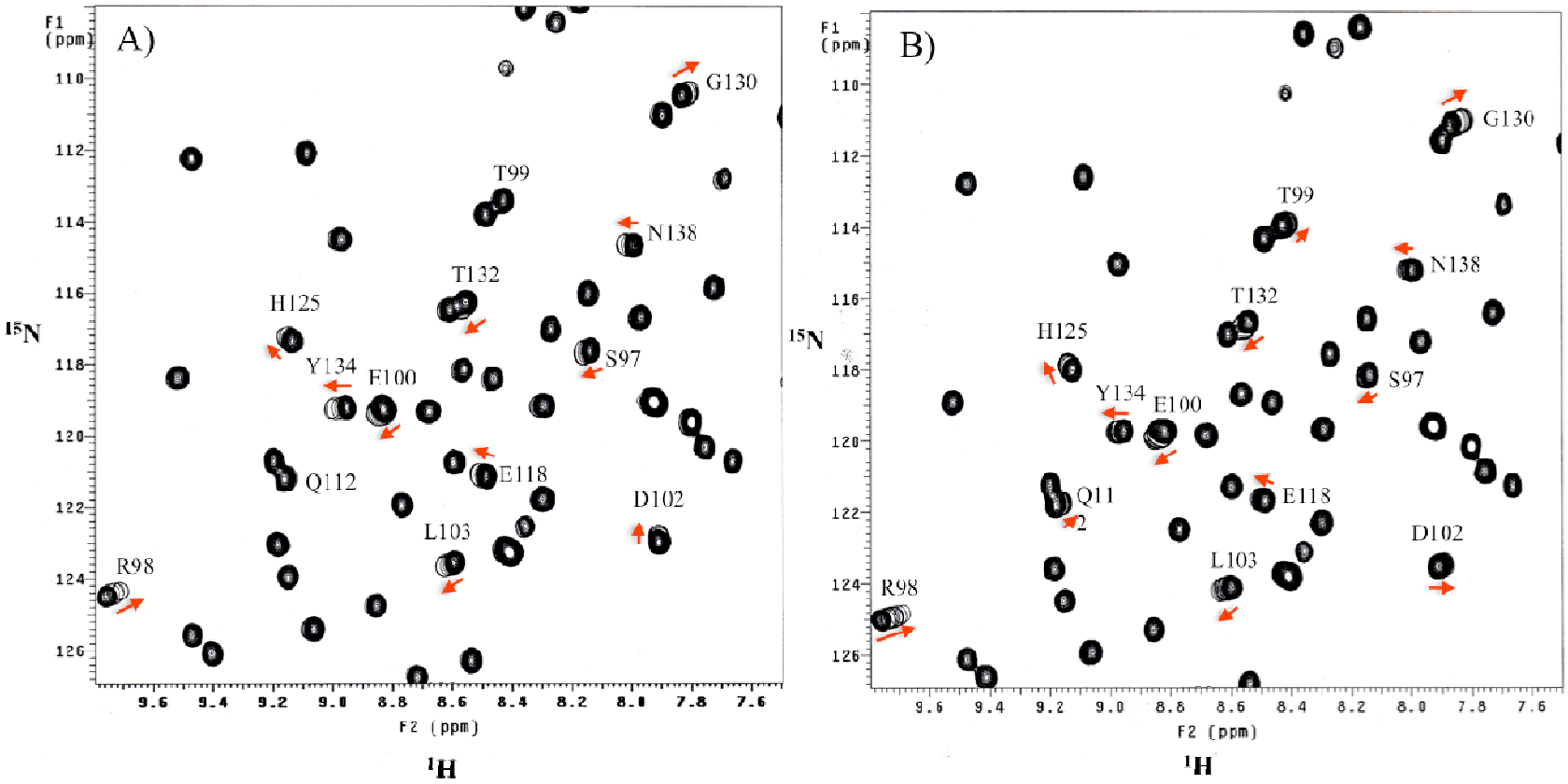


Figure 2.7. Overlayed 2D ^{1}H-^{15}N HSQC NMR spectra of Src-SH3 obtained by titrating the 69-residue full length (A) and the 23-residue (B) MUC1-CD peptides into Src-SH3

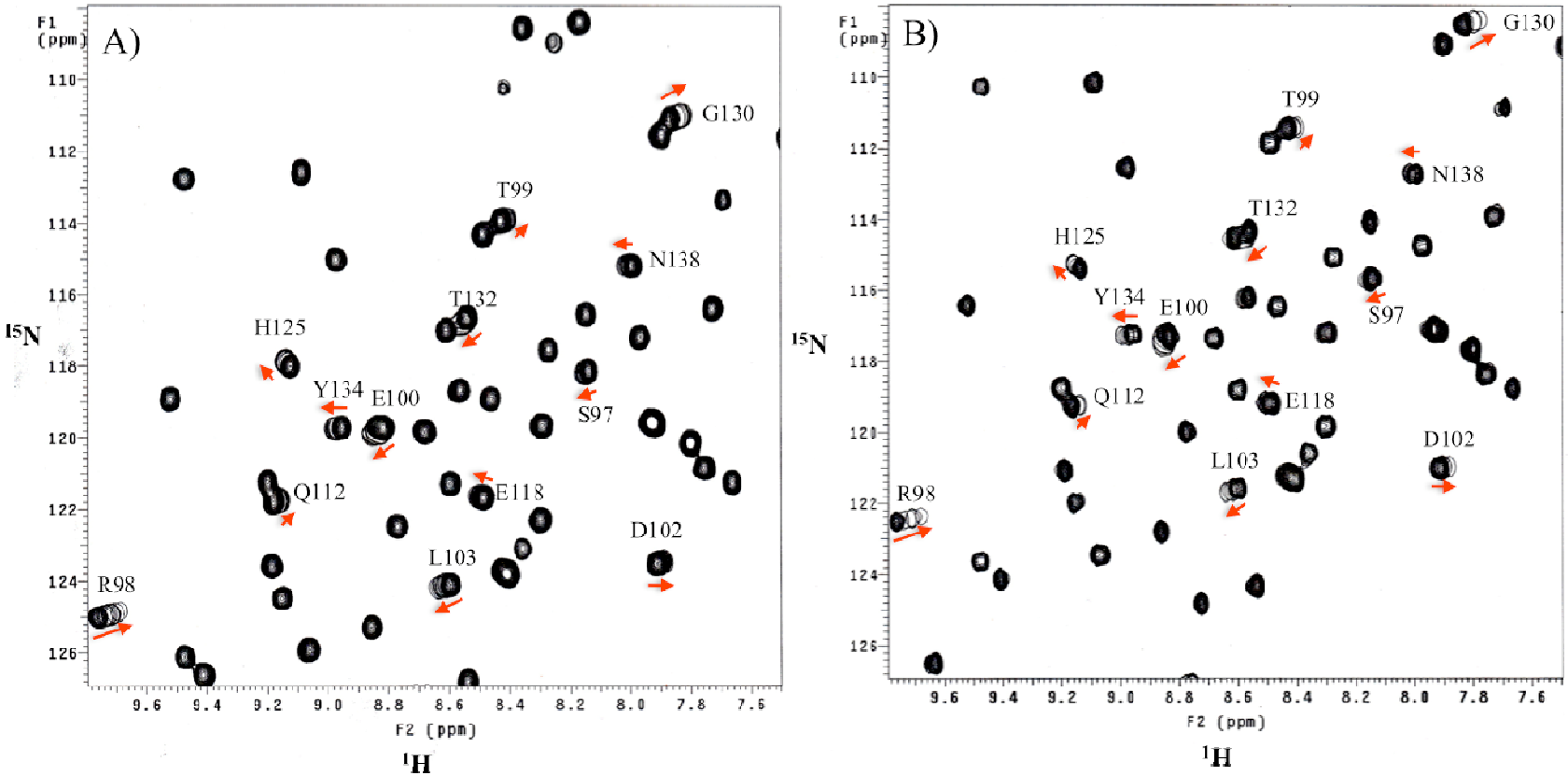


Figure 2.8. Overlayed 2D ^{1}H-^{15}N HSQC NMR spectra of Src-SH3 obtained by titrating the 23-residue monomer (A) and the 48-residue-dimer (B) MUC1 peptides into Src-SH3.

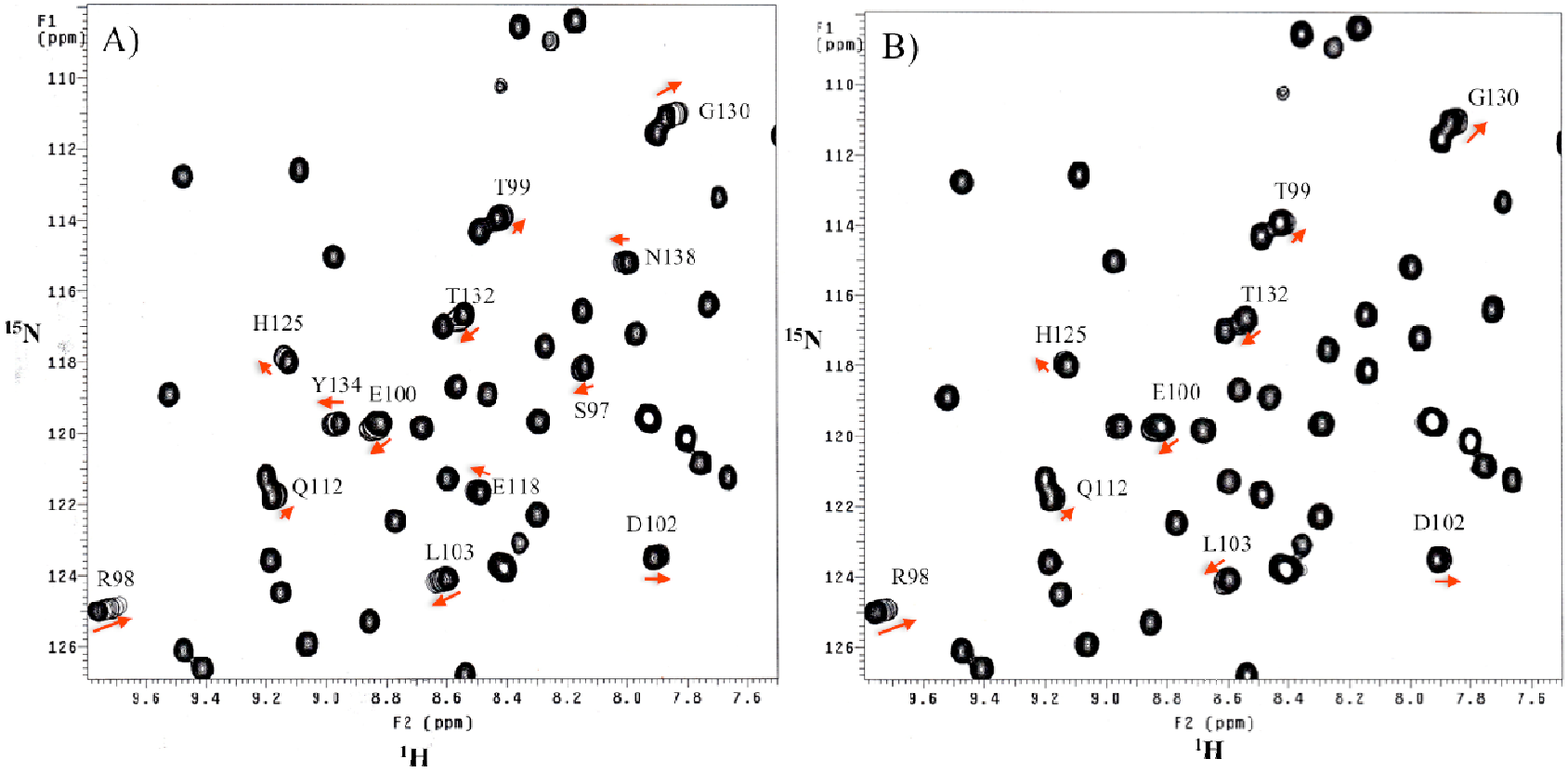


Figure 2.9. Overlayed 2D ^{1}H-^{15}N HSQC NMR spectra of Src-SH3 obtained by titrating the 23-residue monomer (A) and the 23-residue R34A Mutant (B) MUC1 peptides into Src-SH3.

2.3.3. The Dissociation Constant (K_D) of MUC1-CD/Src-SH3 Interaction

The observed chemical shift data were fit into 1:1 binding model but did not fit into 1:2 binding model. The estimated dissociation constants (K_D) at the lowest sums of squares of error (SSE), based on all residues (global K_D), were as follows, i) 2.9 mM for 23-residue native peptide (Fig. 2.10), ii) 2.4 mM for 69-residue full-length peptide (Fig. 2.11), iii) 2.3 mM for 48-residue-dimer peptide (Fig. 2.12) and iv) > 5.00 mM for $R^{34}A$ mutant peptide (Fig. 2.13). The sums of squares of error (SSE) plots indicate that the distribution of SSE for K_D calculations of 23-residue (Fig. 2.10-b), 69-residue (Fig. 2.11-b) and 48-residue-dimer (Fig. 2.12-b) peptides were, all, somewhat similar to one another (The $K_D$at lowest SSE were 2.x mM in all calculations), even after setting the upper limit of K_D as high as 10 mM (data not shown). In contrast, the SSE plots of $R^{34}A$ mutant did not find a converge and thus did not demonstrate a lowest value for the SSE that represent even a closely fitting K_D, when the upper limit was set to 10 mM.

The residues with $\Delta\delta$ > = 0.04 ppm demonstrated a somewhat better fit to the model with a lower local K_D compared to the global K_D although the dispersion of the SSE (Fig. 2.14-g) were still not symmetrical based on all titrations. The 23-residue peptide titration yielded data that showed the best fit with a K_D of 1.05 mM, based on the residues R98, E100, H125, T132 and Y134, followed by 69-residue peptide (K_D =1.85 mM) and 48-residue-dimer peptide (K_D =2.07 mM) (Fig. 2.14). The plots of [MUC1]:[SH3] molar ratio vs. chemical shift change for 23-residue peptide titration (Fig. 2.14-*a* to *e,* in red) show that the binding model fit well with the NMR data except for the last two titration points, partly due to the variations in the data since a higher number of consecutive

titration points were used in between 1:1 and 1:10 compared to only two titration points that were used in between 1:10 and 1:20. The residues E100 (K_D =1.12 mM), H125 (K_D =0.75 mM), G130 (K_D =0.68 mM) and T132 (K_D =0.68 mM) showed the tightest binding based on the 23-residue monomer peptide titration (Fig. 2.14-*a* to *e,* red). The 48-residue-dimer peptide titration showed a similar trend (Fig. 2.14-*a* to *e,* green). Apart from the residues E100, G130, H125, the 69-residue peptide showed tightest binding for Y134, unlike that of the other two peptides (Fig. 2.14-*a* to *e,* blue). The residue, R98, did not seem to saturate even at the highest MUC1 peptide concentrations used for all four titrations.

The first 9 consecutive titration points that generated chemical shift changes corresponded to more or less equivalent molar ratios of both 23-residue and $R^{34}A$ mutant peptides were compared to identify any trends in chemical shift changes based on the individual residues with $\Delta\delta >= 0.04$ ppm (R98, E100, H125, T132 and Y134) (Fig. 2.15). The K_D based on all residues was > 5.00 mM except for the residue, R98 (K_D > 4.63). Overall, the local K_D based on the residues with $\Delta\delta >= 0.04$ ppm was 1.74 mM for 23-residue peptide titration whereas it was >5.00 mM for the titration of $R^{34}A$ mutant indicating that mutating R^{34} to Alanine significantly increased the K_D.

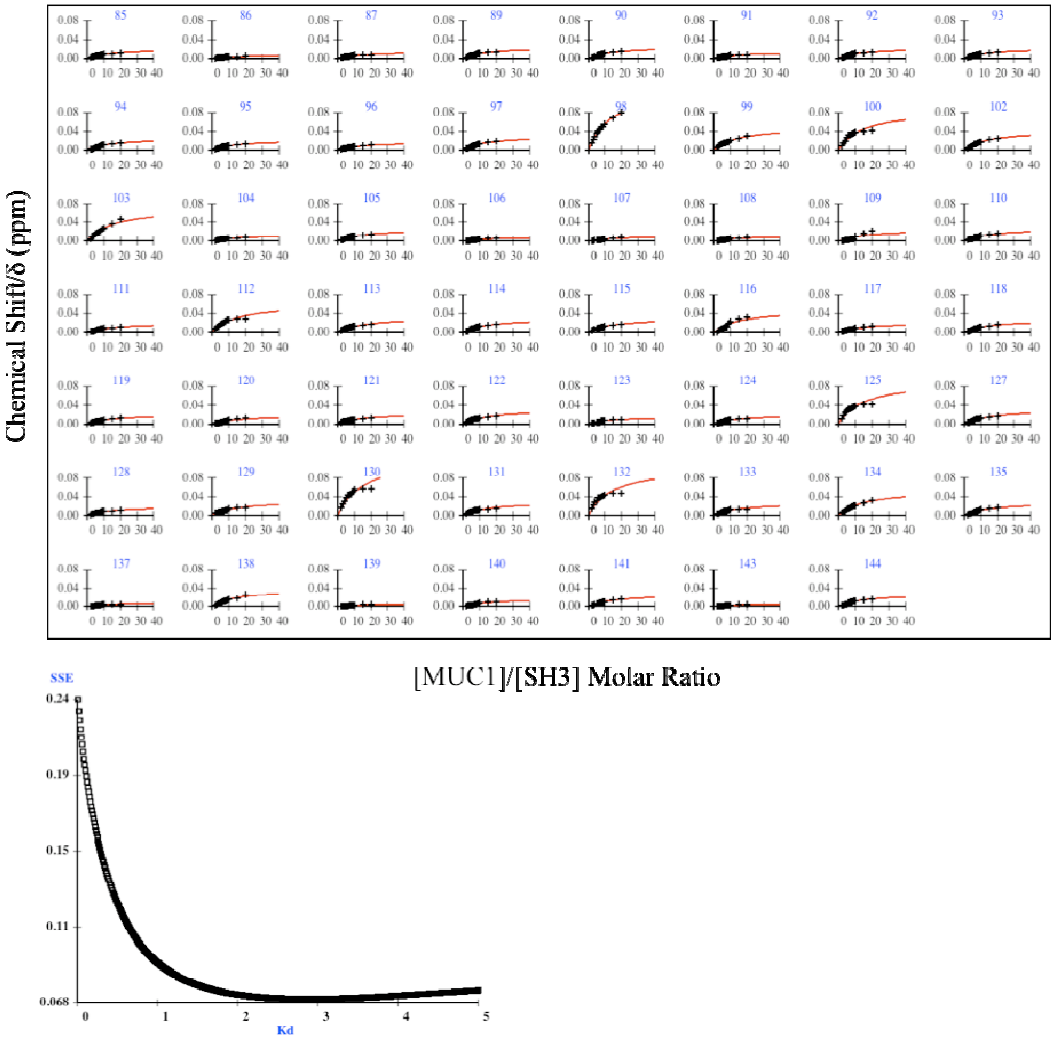


Figure 2.10. ***(a)*** The molar ratio of [MUC1]/[SH3] vs. chemical shift changes per-residue extracted by overlaying 2D ^{1}H-^{15}N HSQC NMR spectra after titrating the 23-residue native MUC1 peptide into Src-SH3 domain ([SH3]: [MUC1] concentration 1:1 through 1:20). The numbers on the plots represent the assigned residues of the 2D ^{1}H -^{15}N HSQC spectrum of Src-SH3 (Fig. 2.6 a). The minimum and maximum values of the dissociation constant (K_D) (1:1 binding model) were set to 0.1 mM and 5.0 mM respectively to obtain a better fit for the observed data. ***(b)*** The estimated K_D values vs. sums of squares of error (SSE) for the model showing the ***global K_D =2.9 mM*** at lowest SSE based on the chemical shift changes of all residues.

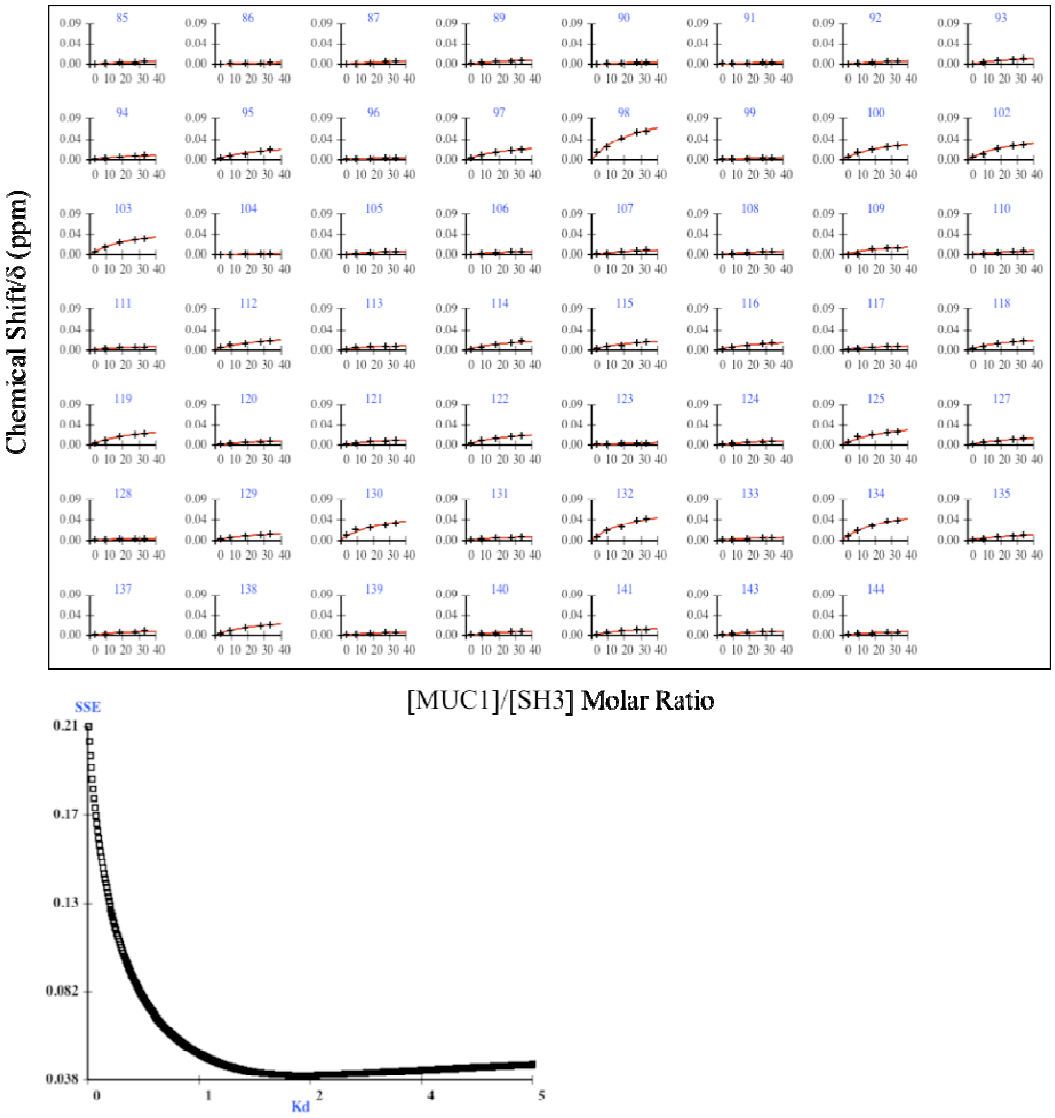


Figure 2.11. ***(a)*** The molar ratio of [MUC1]/[SH3] vs. chemical shift changes per-residue extracted by overlaying 2D ^{1}H-^{15}N HSQC NMR spectra after titrating the 69-residue full length MUC1 peptide into Src-SH3 ([SH3]:[MUC1] concentration 1:1 through 1:33). The numbers on the plots represent the assigned residues of the 2D ^{1}H -^{15}N HSQC spectrum of Src-SH3 (Fig. 2.6 a). The minimum and maximum values of the dissociation constant (K_D) (1:1 binding model) were set to 0.1 mM and 5.0 mM respectively to obtain a better fit for the observed data. ***(b)*** The estimated K_D values vs. sums of squares of error (SSE) for the model showing the ***global K_D =2.4 mM*** at lowest SSE based on the chemical shift changes of all residues.

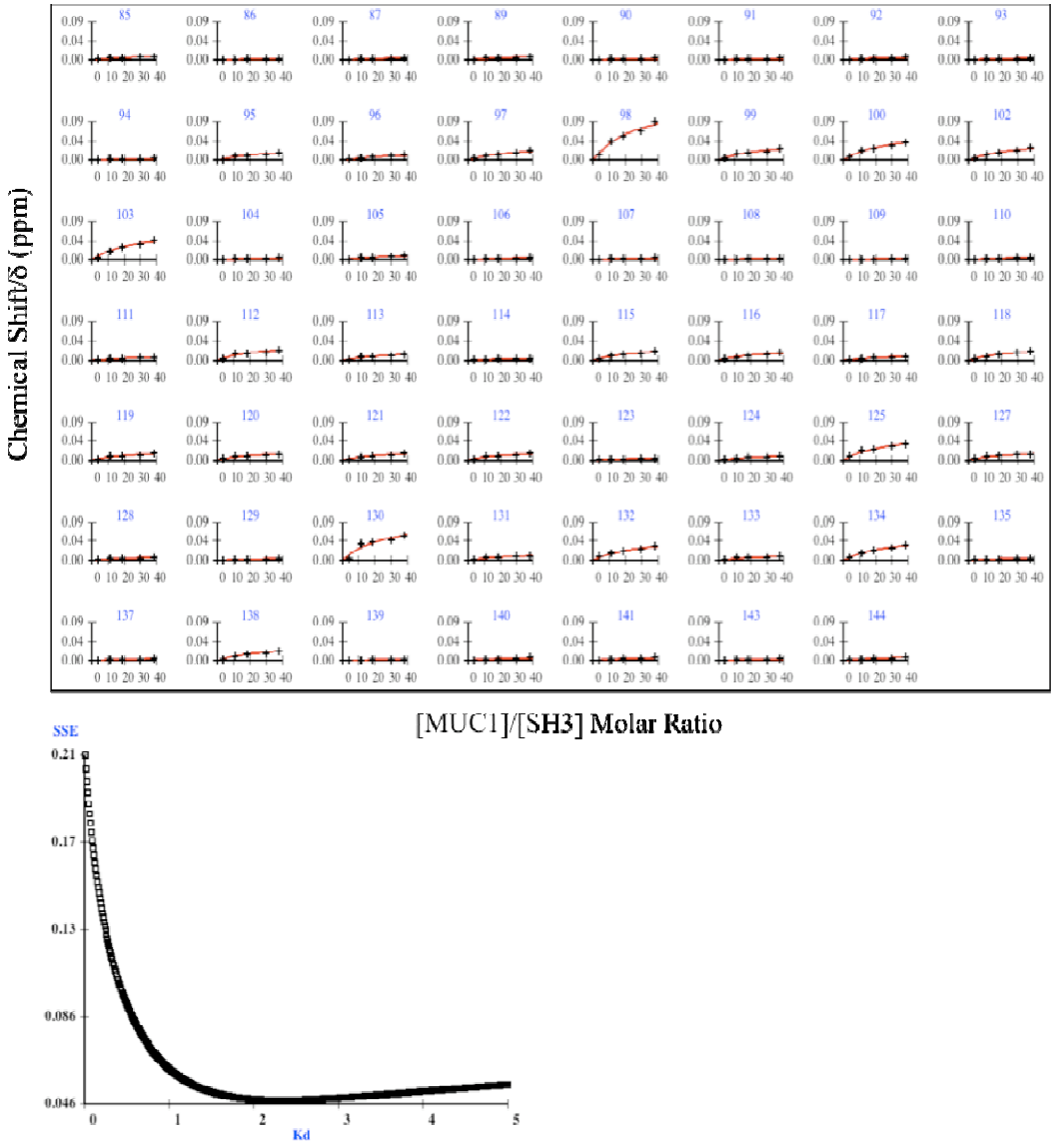


Figure 2.12. ***(a)*** The molar ratio of [MUC1]/[SH3] vs. chemical shift changes per-residue extracted by overlaying 2D ^{1}H-^{15}N HSQC NMR spectra after titrating the 48-residue dimer MUC1 peptide into Src-SH3 ([SH3]: [MUC1] concentration 1:1 through 1:38). The numbers on the plots represent the assigned residues of the 2D ^{1}H -^{15}N HSQC spectrum of Src-SH3 (figure 2.6a). The minimum and maximum values of the dissociation constant (K_D) (1:1 binding model) were set to 0.1 mM and 5.0 mM respectively to obtain a better fit for the observed data. ***(b)*** The estimated K_D values vs. sums of squares of error (SSE) for the model showing the ***global K_D =2.3 mM*** at lowest SSE based on the chemical shift changes of all residues.

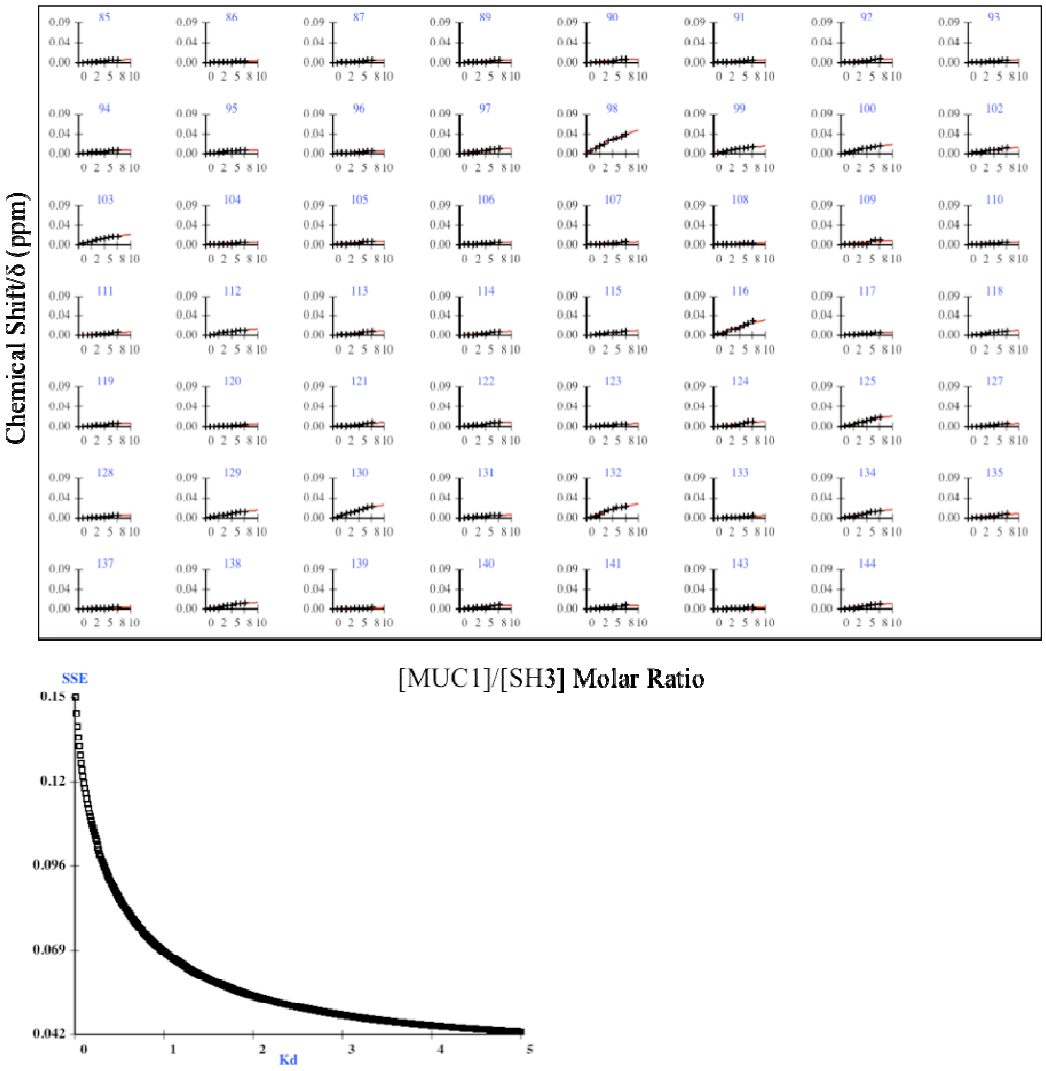


Figure 2.13. ***(a)*** The molar ratio of [MUC1]/[SH3] vs. chemical shift changes per-residue extracted by overlaying 2D ^{1}H-^{15}N HSQC NMR spectra after titrating the 23-residue R^{34}A mutant MUC1 peptide into Src-SH3 ([SH3]:[MUC1] concentration 1:1 through 1:8). The numbers on the plots represent the assigned residues of the 2D ^{1}H -^{15}N HSQC spectrum of Src-SH3 (figure 2.6a). The minimum and maximum values of the dissociation constant (K_D) (1:1 binding model) were set to 0.1 mM and 5.0 mM respectively to obtain a better fit for the observed data. ***(b)*** The estimated K_D values vs. sums of squares of error (SSE) for the model. The estimated ***global K_D*** was ***>5.0 mM*** based on the chemical shift changes of all residues.

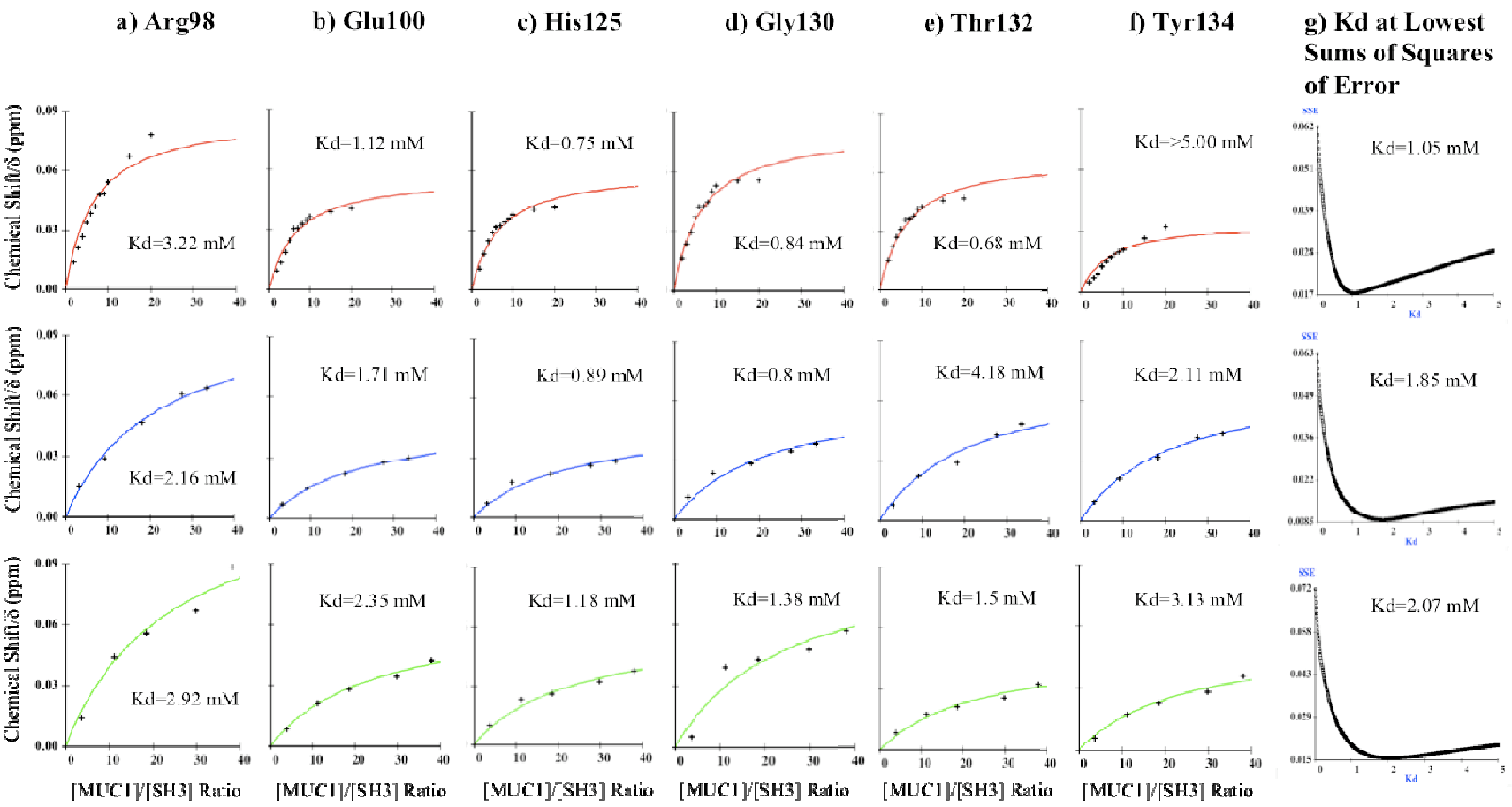


Figure 2.14. The molar ratio of [MUC1]/[SH3] vs. chemical shift changes for the residues with the total chemical shift > 0.04 ppm, extracted by overlaying 2D ^{1}H-^{15}N HSQC NMR spectra of Src-SH3, after titrating the 23-residue monomer (red) the 69-residue full length-monomer (blue) and the 48-residue dimer (green) MUC1 peptides into Src-SH3

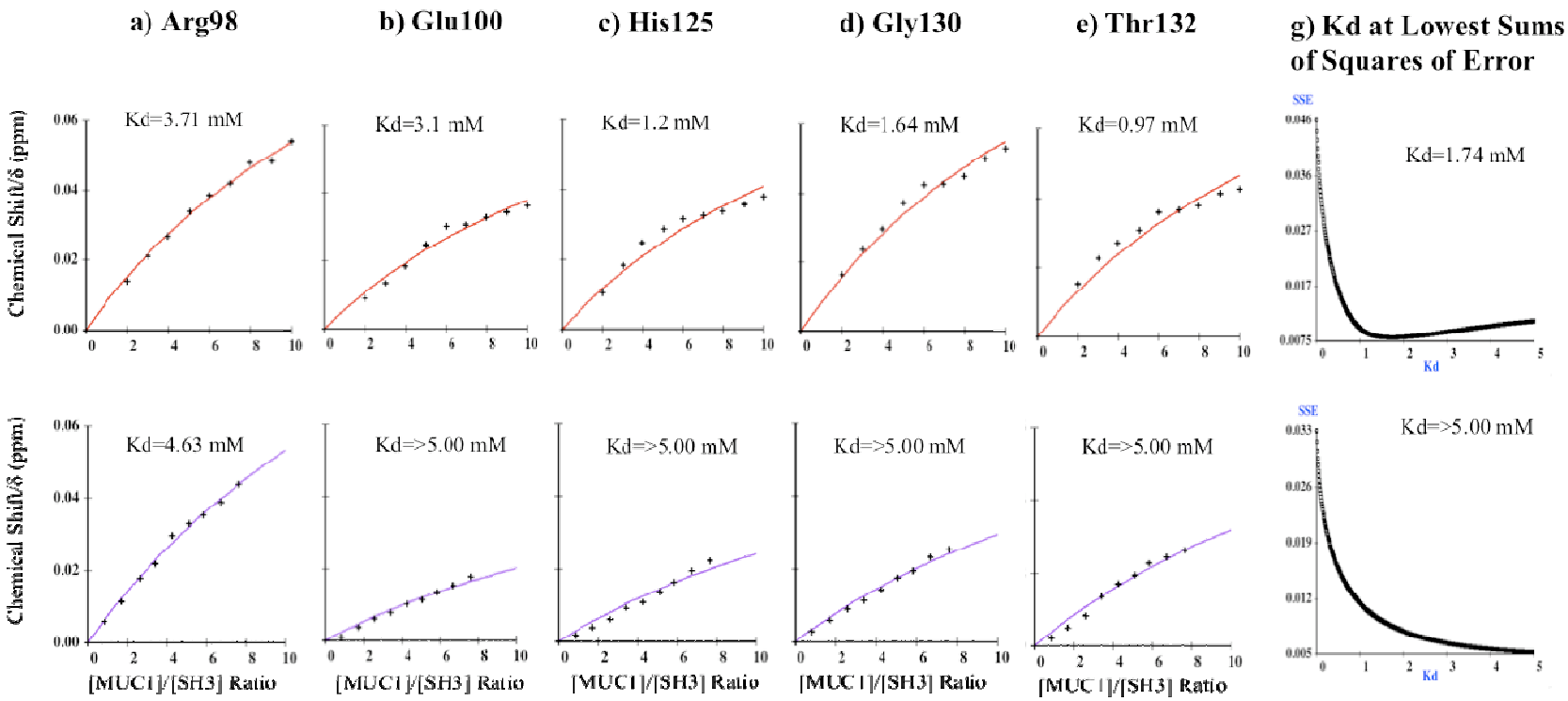


Figure 2.15. The molar ratio of [MUC1]/[SH3] vs. chemical shift changes for the residues with the total chemical shift > 0.04 ppm, extracted by overlaying 2D ^{1}H-^{15}N HSQC NMR spectra of Src-SH3, after titrating the 23-residue native (red) and the 23-residue R34A mutant (purple) MUC1 peptides into Src-SH3

2.3.4. Mapping the MUC1-CD binding site on Src-SH3 Domain

The chemical shift mapping clearly showed that the residues at the canonical binding site of Src-SH3 domain (W121, Y134, Y95, Y93, Y139 and D99) were not perturbed (Fig. 16), except Y134 and D99 that showed clear changes in 69-residue peptide titration [Fig. 2.7, Fig. 2.14 (blue)]. The residues with the highest chemical shift changes were mapped onto the following locations of the Src-SH3 domain; R98 and E100 on the RT loop, H125 on the β-sheet-c, T132 and Y134 on the β-sheet-d and G130 on the distal loop (Fig. 2.16). The location and orientation of Src-SH3 domain in the inactive Src molecule shows that the residues involved in binding with MUC1 peptides are located on the top of the molecule with R98 intercalated into the closed edge of the site of the intramolecular interaction of SH3 and SH2-kinase linker (Fig. 2.17).

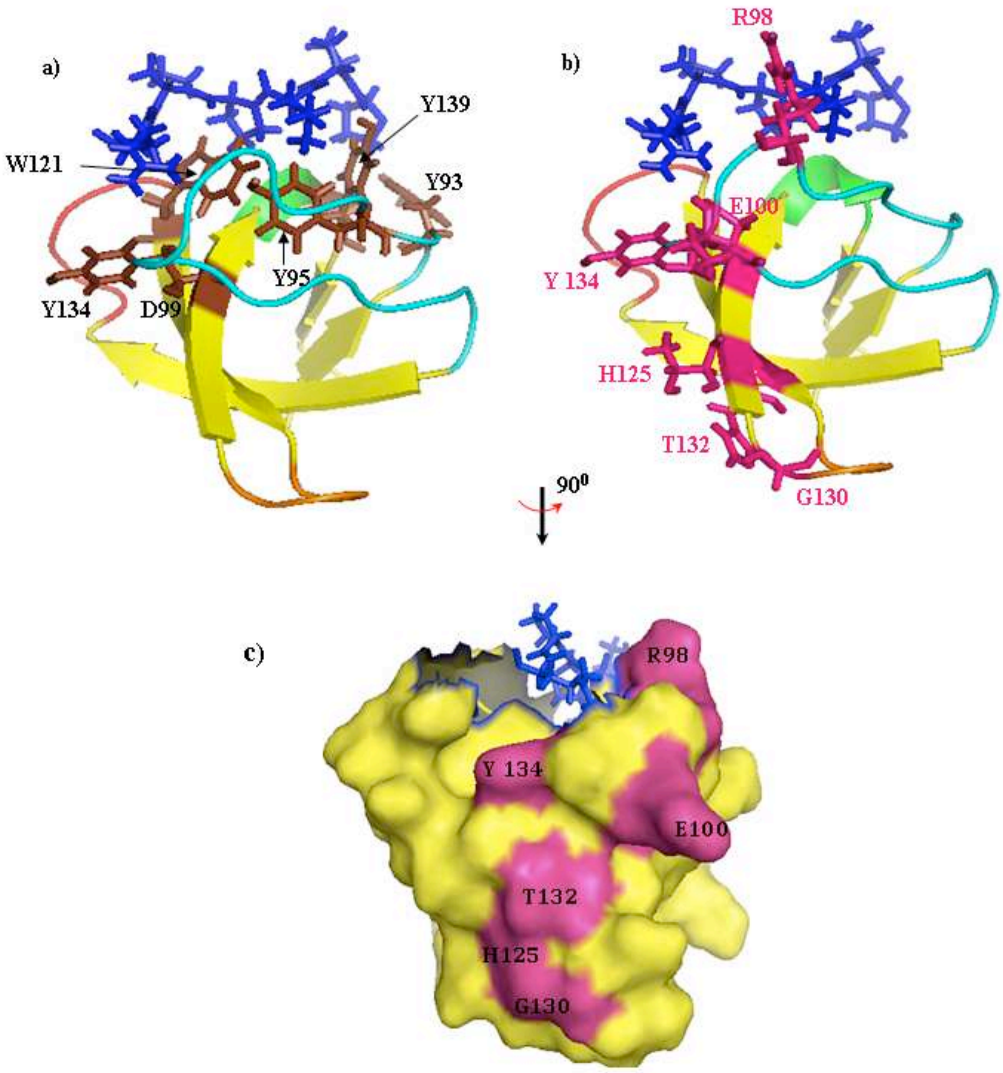


Figure 2.16. ***a)*** A ribbon diagram of Src-SH3 domain complexed with a class II ligand (blue), modified from the solution NMR structure (pdb code: - 1QWE) (233) using MacPymol. The *XP* dipeptidyl moieties of the class II ligand binds to the hydrophobic clefts formed by conserved aromatic residues (brown) that are located in between RT (cyan) and n-Src (red) loops. ***b)*** The chemical shift mapping of potential MUC1-CD binding site on Src-SH3 domain based on the residue-shifts > 0.04 ppm (magenta) obtained by titrating the 23-residue, 69-residue and 48-residue MUC1 peptides into Src-SH3. ***c)*** Molecular surface representation of Src-SH3 domain (rotated 90^0 clockwise with respect to *a* and *b*). The same residues with total chemical shift >0.04 ppm (magenta) are mapped onto the surface.

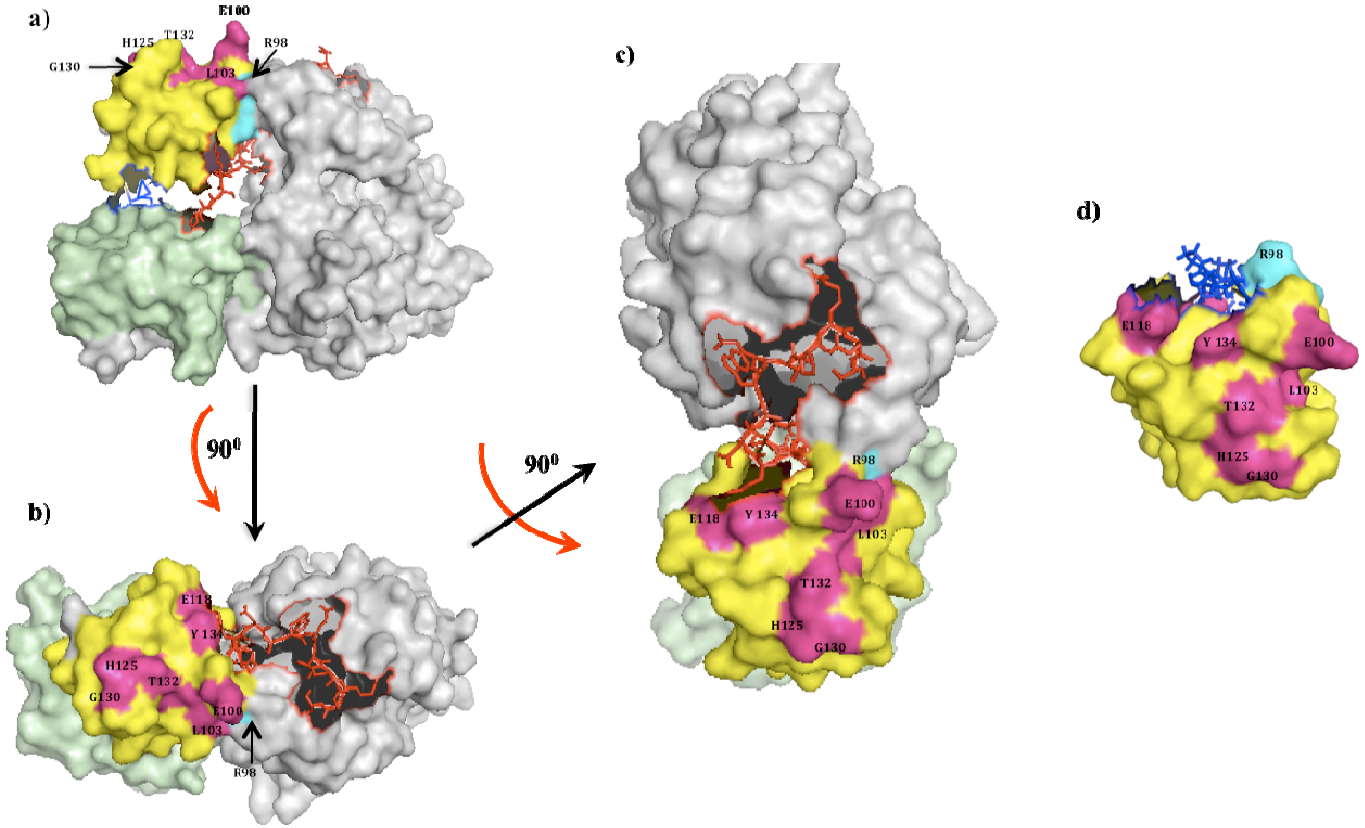


Figure 2.17. The molecular surface representation of the crystal structure of inactive Src (pdb code: - 2SRC), modified to show the relative location and orientation of its SH3 domain with the binding site of MUC1-CD mapped onto its surface: ***a)*** side view of Src showing the SH3 domain (yellow), SH2 domain (pale green), kinase domain (gray), SH3-SH2 linker (Blue) and SH2-kinase linker (red). The residues perturbed by adding MUC1 peptides are numbered (colored in Magenta except R98 in cyan), ***b)*** and ***c)*** the top-view of Src, clearly showing the intramolecular binding of SH3 domain to SH2-kinase linker relative to the perturbed residues (Magenta except R98 colored in cyan) by MUC1 peptides; Note that R98 (cyan) is intercalated into the closed conformation. ***d)*** Src-SH3 domain complexed with a class II PPII ligand (pdb code: - 1QWE), for comparison.

Chapter 3
DISCUSSION AND CONCLUSIONS

3.1. Introduction

The interactions of SH3 domains with their ligands are often weak and short-lived but have proven to be crucial for mediating many important cellular signaling events, such as rapid and transient assembly of signaling complexes for cytoskeletal reorganization and cell migration (237). Unlike the ligands that bind to the SH2 domain, those that interact with the SH3 domain need not be phosphorylated prior to binding and thus, possess an advantage of direct binding to Src. In other words, the SH3-ligands have the ability to lead and initiate signaling cascades via direct Src recruitment even in the absence of posttranslational modifiers (e.g. kinases).

The most recently published work from our lab states that recruitment of Src to MUC1-CD is the first step that triggers downstream signaling towards cancer cell migration (150). If no other kinases or growth factor receptors (such as EGFR) are involved, MUC1-CD which does not have intrinsic kinase activity, has only two possible mechanisms to trigger phosphorylation dependent recruitment of Src via the SH2 domain; viz. i) MUC1-CD may recruit unbound or partially active Src molecules directly through the Src-SH3 domain allowing subsequent phsphorylation of Y^{46} and binding of Src-SH2 domain, or ii) already kinase-active Src may be in the vicinity of MUC1-CD to phosphorylate Y^{46} and bind to it.

The present study was designed to investigate the relative binding affinities and specificities of SH3 domain to MUC1-CD, which would provide invaluable insights into the putative mechanisms of Src recruitment by MUC1-CD. The results revealed important structural aspects of the interaction of MUC1-

CD with Src-SH3, including the binding affinity and a potential binding site of MUC1-CD on Src-SH3. As reviewed and discussed below, these findings may serve as a foundation and provide insights for further studies on the MUC1/Src interaction.

3.2. Review and Discussion of Experimental Data

i) The ^{1}H-^{15}N HSQC spectra of Src-SH3 domain

The successful expression and purification of ^{15}N labeled Src-SH3 domain in milligram quantities, was the first step that facilitated this study. Since SH3 is a well-folded modular binding domain that is relatively small (55-60 amino acids) and stable in solution, it has been ideal for *in vitro* biophysical studies such as NMR. In fact, the structure and ligand binding dynamics of Src-SH3 domain have been well characterized (229), (233), (275). As a result, the 3D NMR solution structures and the assigned chemical shift values for ^{1}H and ^{15}N isotopes of Src-SH3 are available through the BMRB, which has accelerated the current study since there was no need to carry out a set of ^{15}N-NOESY (Nuclear Overhauser Effect SpectroscopY) and ^{15}N-TOCSY (Total Correlation SpectroscopY) experiments to assign ^{1}H and ^{15}N backbone resonances in 2D ^{1}H-^{15}N HSQC spectra. Thus, the 3D structures of inactive Src and Src-SH3 domain available through the protein databank facilitated the chemical shift mapping of the interacting site of the SH3 domain upon complexation with a MUC1-peptide.

ii) Residue-shifts of Src-SH3 upon titration with MUC1 peptides

The observed chemical shift changes of the residues (residue-shifts) of 2D ^{1}H-^{15}N HSQC spectra upon addition of i) 23-residue, ii) 69-residue and iii) 48-residue-dimer MUC1 peptides were almost the same, signifying that all 3 peptides interact with the same binding site of the Src-SH3 domain. Since these peptides differ in sequence length and modifications (i.e. 48-residue-dimer) it was possible that each peptide would interact differently with the Src-SH3 domain, *i.e., i)* if the 23-residue peptide did not have the full consensus motif or flanking residues required for actual interaction, it would interact differently with Src-SH3 than the 69-residue peptide and would show different residue-shifts; similarly*, ii)* if the 48-residue peptide, due to dimerization, would have selected a different site to bind, it would also show different residue-shifts compared to that of the 23-residue monomer. However, there were no marked differences in the ^{1}H-^{15}N HSQC spectra of Src-SH3 domain obtained after titrating these peptides, which clearly indicates that the shorter MUC1 peptides (23-residue monomer and 48-residue-dimer) seem to include the intact SH3-binding-motif that is present in the 69-residue full length MUC1-CD.

However, the NMR data acquired in this study do not provide information on potential differences in the structural behavior of the three MUC1 peptides, as they bind to Src-SH3, because the experiments investigated only the chemical shift changes occuring in the SH3 domain (protein) but not the changes occuring in MUC1-peptide (ligand), which is unlabelled and invisible in protein-based NMR spectra.

Based on the SH3 domain-observed NMR data, it is apparent that the

structural differences of MUC1 peptides did not seem to play a huge role in making contacts with the Src-SH3 domain. The small chemical shift changes and weak binding affinity imply that none of the MUC1 peptides contact the Src-SH3 domain as closely as typical PXXP ligands that make hydrophobic contacts with SH3, as shown by similar studies (278), (279). The best possible explanation for such weak interactions, which mainly involve charged residues, are electrostatic attraction forces. It has been well established that the electrostatic interactions between charged and polar groups in proteins and ligands play a major role in protein–ligand binding specificity (280). Among the residues of Src-SH3 domain that demonstrated the highest chemical shift changes upon addition of MUC1 peptides were charged residues (E100, R98, H125) (Fig. 2.7, 2.8), two of which reside in the highly flexible RT loop. One study that addressed an ultra-weak interaction between Nck-SH3-3 domain and PINCH-LIM4 domain (that has a non-PXXP motif) reported a ^{1}H-^{15}N HSQC spectra of NcK-SH3-3 domain that shows very small residue shifts upon binding to PINCH-LIM4 domain similar to the HSQC spectra obtained in this study (242). The NMR solution structure of the complex of Nck-SH3-3 and PINCH-LIM4 domains shows that the binding interface was exclusively dependent on the salt bridges formed between charged residues.

Molecular dynamics simulations show that ligands contact SH3 domains via long-range electrostatic interactions to initiate the formation of transient encounter complexes prior to establishing the typical hydrophobic contacts with the canonical ligand binding surface of the SH3 domain (281). The results indicate that all three MUC1-pepides (69-residue, 23-residue and 48-residue) may

form such transient complexes via electrostatic interactions but may not in fact contact the ligand binding surface of Src-SH3. If the MUC1 peptides formed a PPII-helical conformation at the SH3 binding motif, the residues in the conventional binding surface of Src-SH3 would have been perturbed. This lack of binding to the canonical site suggests that MUC1-CD does not seem to form a polyproline type II helix, which could have been anticipated since MUC1-CD does not have either i) a proline rich motif that extends beyond RYVPP that would provide two consecutive XP dipeptide moieties to make hydrophobic contacts with the two SH3 binding grooves or ii) a patch of hydrophobic residues that would specifically make hydrophobic contacts with the ligand binding surface of Src-SH3.

The comparison of binding data of native and $R^{34}A$ mutant MUC1 peptides (23-residue) showed a marked increase in K_D for the titration of the $R^{34}A$ mutant (Fig. 2.10 vs. 2.13 and 2.15). Although this data is not sufficient to confirm that MUC1-peptides specifically use this N terminal motif to contact the Src-SH3 domain, the results imply that the N terminal arginine in the $R^{34}YVPP$ motif is an integral part of the SH3-binding-motif. This observation can be related to the fact that the electrostatic field significantly changes when residues are mutated in a binding partner (280). However, there was no marked difference in the movement of D102 (D99 as in (229)) in the HSQC spectra obtained for the titration of $R^{34}A$ mutant, compared to 23-residue peptide (Fig. 2.9), indicating that D102 may not form a salt bridge with R^{34} of the ligand as expected. Since the chemical shift changes are small it can be assumed that D102 is just moving due to the changes in the vicinity. Since E118 is not moving upon titration of the

$R^{34}A$ mutant MUC1 peptide (compared to 69-residue full length MUC1-CD), it is reasonable to assume that E118 might be interacting with R^{34} (at the RYVPP site) but the chemical shift change of this residue is very small for the titrations of native peptides. The residues Y134, S97 and N138 were not perturbed upon the titration of mutant MUC1 peptide suggesting that the N terminal portion of MUC1-CD, relative to R^{34}, is interacting with these residues of SH3 domain.

Although MUC1 has the proper spacing of a typical class I motif, RXXP (RYVP in MUC1-CD), the VP dipeptide moiety may not contact the binding grooves of Src-SH3 because none of the residues that form this pocket (Y95 and W121) show any change in chemical shift. As a matter of fact, none of the residues in the ligand-binding site (W121, Y134, Y95, Y93, Y139 and D99) except D99 and Y134 show chemical shift perturbations suggesting that MUC1 peptides may not make hydrophobic contact with the SH3 ligand-binding site, unlike the classic SH3 ligands, as discussed above. Out of all the aromatic residues in the ligand binding surface of Src-SH3, Y134 is the only residue that shows a relatively significant chemical shift change, especially in the 69-residue full-length MUC1 peptide titration (Fig. 2.14). It is known that Y134 is involved only in class-I peptide binding (233) and the residues Y134 and D102 (D99 as in (233)) form the third specificity pocket of Src-SH3 domain (Fig. 1.11).

Even if the MUC1 peptide does not seem to contact the canonical binding surface, it seems to interact with the charged residues in the RT loop such as R98 and E100 (Fig. 16). The residue, R98 of Src-SH3 shows the highest total chemical shift change (~ 0.09ppm) in all MUC1 peptide titrations (Fig(s). 2.7, 2.8, 2.9) but did not seem to saturate with the increasing amounts (used for this study) of any

of the MUC1 peptides (23-residue, 69-residue, 48-residue and $R^{34}A$ mutant). The plots of molar ratio of [MUC1]/[SH3] vs. chemical shift (Fig.(s) 2.14, 2.15) shows a weaker binding constant compared to that of the other shift-perturbed residues, even at the highest MUC1 peptide concentrations.

It is possible that D42 and R43 of MUC1-CD are making electrostatic interactions with R98 and E100 of Src-SH3, respectively. Out of these two residues, E100 shows a tighter binding and seems to saturate with higher peptide concentrations used in the experiments (K_D= 1.12 mM) whereas R98 shows a much weaker interaction (K_D= 3.22 mM). Since D42 and R43 are adjoining residues (Fig. 2.2), there may be electrostatic attraction as well as repulsion forces (e.g. between R98 of Src-SH3 and R43 of MUC1 peptide) involved while these two residues attempt to establish contacts with R98 and E100 of the Src-SH3 domain (Fig. 2.16). Since E100 is closer to the peptide–binding–path than R98, it may have a higher chance to interact with an appropriate basic residue (presumably R43) than R98, which may weakly bind to the neighboring acidic residue (presumably D42).

The $R^{34}A$ mutant MUC1 peptide titration also revealed that R98 significantly shifts; the estimated K_D based on all residues was > 5.00 mM and it was 4.63 for R98, indicating that it contacts the peptide in a similar manner as the native peptides. The mutant binding data also suggest that the point mutation of R^{34} (to Alanine) did not affect other potential residue-specific interactions (presumably D42 and R43) of MUC1 peptides. The R43A mutant MUC1 peptide (Fig. 2.2) and a $D^{42}A$ mutant would help to identify whether these residues, in fact, are responsible for interacting with E100 and R98, and would provide

important clues about the binding orientation of MUC1 peptides and specificity. The X-ray crystal structure of the repressed form of intact Src molecule (pdb accession code 2Src), which lacks the D20-ligand salt bridge, is stabilized by several other electrostatic interactions including R98 (275). The chemical shift mapping also shows that R98 is not a surface exposed residue, like E100, in the inactive Src molecule (Fig. 17) suggesting that it may only be available for binding when the Src is partially or fully active, in which the SH3 domain is free for ligand binding. The mutation of R98 directly affected the interaction between SH3 domain and SH2-kinase linker region (282).

In summary, the charged residues interspersed in the vicinity of the putative SH3 binding motif of MUC1-CD (R34, R43, D42) could be mainly responsible for mediating the MUC1-CD/Src-SH3 interaction. Mutation of R34 to Alanine revealed the potential residues of Src-SH3 (E118, Y134, S97, N138) that may contact the N terminal portion of MUC1 peptide relative to the putative binding site, and confirmed that D102 may not be responsible for salt bridge formation. The $R^{34}A$ mutant MUC1 peptide also provided invaluable insights into the residues of MUC1-CD that may be involved in the binding of E100 and R98 of the Src-SH3 domain. The mutant also revealed a possible MUC1-CD-binding pocket that partially surrounds the canonical specificity pocket but extends towards the distal loop via the RT loop (Fig. 2.17). The residue-specific data such as similar NMR binding studies of point mutants, alanine scanning mutagenesis (or peptide walking arrays) or NMR structural information of the ligand binding interface, will be required to confirm the other MUC1-CD residues that are involved in binding.

iii) The similarities and differences in chemical shift perturbations of Src-SH3 domain upon titrating 23-residue, 69-residue and 48-residue peptides

Apart from minor differences, 2D ^{1}H-^{15}N HSQC spectra obtained by titrating 69-residue, 23-residue and 48-residue peptides appeared almost identical. Since the SH3 domain-based NMR spectra do not provide evidence that would help to differentiate the structural dynamics of different peptides as binding to Src-SH3, it is not possible to define any reason other than the pre-determined differences among peptides such as length/sequence differences and modifications (in 48-residue-dimer peptide).

The direction of chemical shift change in the residue, D102, was different in the HSQC spectrum obtained for the titration of 69-residue peptide compared to that of the 23-residue and 48-residue peptides (Fig.(s) 2.7, 2.8), probably due to the difference in sequence length. Also, the residues, Q112 and T99 were not perturbed in the 69-residue peptide titration. These observations suggest that the 69-residue full length MUC1-CD may contact the Src-SH3 in a slightly different manner compared to 23-residue and 48-residue peptides.

As described above, it is hard to assume different ways that 23-residue monomer and dimer bind with Src-SH3, since the ligand-based data are not available. The non-native 48-residue-dimer may not form a composite binding site, as there was no difference in the way that it contacted the SH3 domain compared to that of the 23-residue monomer. Since MUC1-CD peptides are structurally flexible/disordered (data not shown) the two monomers in the 48-residue-dimer may not aggregate non-covalently (no hydrophobic patch of residues). As a result, each dimer partner (which is a monomer) may interact with

one SH3 domain in the same way that the 23-residue-monomer peptide would interact and therefore, may not show any differences in SH3-observed NMR spectra; if the ligand-based binding data were available, this would be a 1:2 binding event (one 48-residue-dimer: two Src-SH3 domains). The other possibility would be that only the unphosphorylated dimer partner may interact with the SH3 domain while the phosphorylated dimer partner may not contact SH3 probably due a change in electrostatic forces that might have occurred due to phosphorylation affecting binding since Y^{46} is located in the vicinity of putative SH3 binding motif. Theoretically, other NMR studies such as NMR relaxation dispersion studies, filtered-edited NOESY experiments based on doubly labeled proteins (^{15}N and ^{13}C) and/or site specific spin labeling of peptides could be used to characterize the structure and dynamics of the protein-ligand complex. However, acquiring well-resolved NMR spectra seems not to be realistic due to the flexible nature of the structure of MUC1-CD.

Nonetheless, the choice of 69-residue full-length MUC1-CD peptide for this study (vs. short 23- and 48-residue peptides) reflects even the slightest advantage of studying a particular interaction based on the native, intact molecules, whenever possible, to avoid any bias introduced by modifications. In SH3 domain binding studies though, short synthetic peptides are used first and foremost due to the fact that the conformational entropy could unfavorably change with increasing length of an unstructured peptide.

iv) The binding affinity of Src-SH3 and MUC1-CD

Determination of a dissociation constant (K_D) that keep a protein-ligand complex together is an essential step of characterization of the complex, and also provides a means of cross-referencing the binding affinities among similar studies (264). Since the typical SH3-ligand affinity is relatively weak (K_D=5-100 μM) (228), the K_D of MUC1-CD and Src-SH3 interaction was expected to fall within the micromolar range. The results however, revealed that the K_D of MUC1-CD and Src-SH3 domain interaction falls in the millimolar range; viz. the lowest reported K_D based on the residues with total chemical shift >0.04 ppm was 1.85 mM for the full length MUC1-CD. Based on all analyses, K_D ranges roughly from 1.0 - 3.0 mM. Nonetheless, it is apparent from the literature that some SH3 domain-ligand interactions are promiscuous and the affinities can vary from 1.0 μM – 3.0 mM. The weakest interactions that involve only tertiary contacts with the ligands have been proven to be physiologically relevant (242).

The ultra-weak (millimolar) binding affinity of MUC1-CD/Src-SH3 interaction and binding to a non-canonical site suggests that MUC1-CD may not fall into the category of a high affinity activating ligand of Src-SH3 domain (the ligands that can competitively bind to the canonical binding site of SH3 displacing its intramolecular interaction with SH2-kinase linker). The *in vitro* structural studies of HCK show that the SH2-kinase linker does not bind to SH3 domain in the same manner as *in vivo*, due to the absence of the kinase domain (283), implying that the mechanism responsible for the intramolecular binding operates only in the intact Src molecule, which is stabilized by both hydrophobic and electrostatic interactions with the SH3 domain as well as the small lobe of the

kinase domain (284). The physical and dynamic differences between the cytosol of a living cell (*in vivo*) vs. a buffer solution used in a laboratory (*in vitro*) may show differences in binding affinity of MUC1-CD and Src-SH3 due to similar reasons. The membrane bound MUC1-C subunit (which has MUC1-CD at the C terminal end), may be more stable than the free full-length MUC1-CD peptide in solution and thus may bind to Src more firmly. Also, the binding affinity of MUC1-CD and the SH3 module in the intact Src molecule may be slightly different compared to the results of the current study that addressed MUC1-CD binding to the free SH3 domain in solution.

The current study used a high salt buffer solution (100 mM NaCl) to provide the most appropriate buffering conditions for the SH3 domain (229). The ionic strength of a solution is increased with salt concentration, which weakens the electrostatic interactions. It has been shown that the electrostatic network, which is salt dependent, has a significant influence in intramolecular interactions of HCK (285). Thus, the differences in ionic strength *in vitro* vs. *in vivo* (cytosol) may also account for MUC1-CD and Src-SH3 interaction, which seems to be dependent on electrostatic interactions, as revealed by the current study.

Cellular signaling pathways are often dynamic and must be activated and inactivated quickly. This is especially true for SH3 domains which participate in assemblies of molecules that operate as transient but specific switching between multiple interaction partners with fast on and off rates (286). MUC1-CD may also show faster on/off rates as a scaffold protein that has evolved to bind multiple molecules. As confirmed by the current study, the interaction between MUC1-CD and Src-SH3 was anticipated to be much weaker than the regular SH3 domain

activator molecules such as FAK (224).

v) The binding specificity of Src-SH3 and MUC1-CD

The current study reports a unique binding specificity of Src-SH3 domain that involved the residues, H125, T132 and G130 whereas the other perturbed-residues (R98, E100, E118, Y134, S97, L103, N138) are commonly involved in determining the specificity of several other SH3-ligand interactions (245), (233). The chemical shift mapping showed that MUC1-CD binding site is mainly located on a side of the SH3 domain (Fig. 2.16), which forms a binding site that partially surrounds the specificity pocket of the canonical binding site (N138, E118, Y134) => RT loop (E100/R98) = > H125 on β-sheet-c and T132 on β-sheet-d => distal loop (G130). This binding site is oriented at the top of the inactive Src molecule (Fig. 2.17), based on the relative location of the SH3 domain in the intact Src molecule. Although the non-PXXP ligands choose a different path than the PPII ligands that directly interact with the canonical binding surface of SH3 domain (241), the binding site revealed by this study is significantly different compared to those studies. Some differences are also shown to be due to the differences in amino acid composition among SH3 domains, specifically those of n-Src and/or RT loops, of different SH3 domains (287). Given that the estimated binding affinity of MUC1-CD and Src-SH3 interaction is ultraweak, such specificity could serve as a distinguishing mechanism for MUC1-CD/Src-SH3 interaction that might have been evolved in order for MUC1-CD to contact Src in a specific cellular context.

3.3. Contribution to the advancement of knowledge and future directions

The current study was completely novel and thus creates a solid basis for further structural studies that would completely characterize the interaction of the MUC1 and Src-SH3 domain. The experimental data based on different short synthetic native/mutant peptides vs. the full length MUC1-CD, reflects the uniqueness of this interaction. Since the interacting residues of Src-SH3 were identified in this study, the binding affinity and specificity can be further investigated by mutating some of the key residues of the SH3 domain as well as those of the putative binding site of MUC1-CD. The $R^{34}A$ and $P^{37}A$ mutants have already been obtained for binding studies. Once these residues are confirmed to be involved in binding, the SH3 domain and MUC1-CD mutants can be tested *in vivo* using the breast cancer cells lines such as MCF7 and T47D, which have already been used to identify MUC1-ICAM-1 interaction along with the MUC1-transfected vs. non-transfected 293T cells, to characterize the downstream physiological effects of SH3 domain binding to MUC1-CD.

The structure of the binding interface of the MUC1-Src-SH3 complex can theoretically be calculated using half-filtered NOESY, 3D (three dimensional) NMR experiments based on the doubly labeled (^{15}N and ^{13}C) Src-SH3 domain bound to unlabelled MUC1 peptide. This would yield the ligand based intermolecular NOEs (^{12}C and ^{14}N) that are attached to ^{15}N and ^{13}C nuclei of labeled Src-SH3 domain (262). Since the MUC1-CD is unstructured and the interaction is ultraweak, however, it may not be feasible to differentiate the peaks of an NMR spectrum due to absence of and/or overlapping resonances.

The other biophysical methods such as isothermal titration calorimetry

(ITC) can be used to complement the binding data and to determine the change in Gibbs free energy, Enthalpy and Entropy of MUC1-CD/Src-SH3 interaction that would assess whether this reaction is favorable. Such data can be used to compare with similar studies in the literature that would expand the body of knowledge about the Src-SH3 domain as well as the MUC1-CD that can be directly applicable in planning other studies.

The chemical shift mapping of the current study indicates that the SH3 and potential SH2 domain binding events of MUC1-CD could either be spatially (two MUC1-CD molecules are required for binding) or temporally (SH3 domain binding is short-lived and preceded by SH2 domain binding) separated. If the cysteine-linked dimers of MUC1 are involved in binding, the fixed relative orientation of Src-SH3 and SH2 domains may strengthen the interaction of MUC1 and Src. Specifically, one dimer partner may bind to SH3 while the other binds to p-Y^{46}EKV motif of MUC1-CD. The chemical shift mapping of individual SH3 and SH2 domains vs. combined (SH3-linker-SH2) domain of Src upon binding of phosphorylated and unphosphorylated MUC1-CD would provide a basis for unraveling the possibilities of monomer or dimer binding to combined (SH3+SH2) domain and help to differentiate the binding events.

3.5. Conclusions

The current study was the first to investigate the binding affinity and specificity of the interaction of the MUC1-CD with the Src-SH3 domain. It revealed completely novel details of a unique type of interaction. The chemical shift mapping of Src-SH3 domain suggested that MUC1-CD does not bind to the canonical binding site indicating the lack of polyproline type-II helical structure in the MUC1 binding site. However, as anticipated, the residue, R^{34}, in the putative binding site was crucial for binding while the other charged residues downstream of RYVPPSS motif were also seen to be important in establishing contacts with the charged residues of the Src-SH3 domain.

The mapping of the binding site of MUC1-CD on the SH3 domain of the inactive Src molecule provided important insights into a possible mechanism of this interaction and suggested that MUC1-CD may interact with partially activated Src molecules through electrostatic attraction forces. The SH3 domain interactions are temporally and spatially separated to increase the specificity (288). As confirmed by the previous *in vivo* studies, MUC1 is involved in Src mediated cancer cell migration and MUC1 may potentiate Src signaling in breast cancer. Taken together, it can be speculated that in the presence of both MUC1 and Src in high amounts, in a cancer cell, the membrane bound C terminal subunit of MUC1 (that has MUC1-CD on the cytoplasmic side) may attract membrane bound or partially activated Src molecules via electrostatic interactions to bring Src closer to MUC1; then Src may i) interact with its binding partners such as integrins, located at the cell membrane, to get activated and then ii) may phosphorylate Y^{46} and bind to p-YEKV motif of MUC1-CD. In normal cells

MUC1-CD is membrane-bound but Src is not and the two molecules are expressed at the normal levels that may spatially and temporally control contact of each other at the right orientation. Therefore, it is possible that MUC1-CD/Src-SH3 interaction is physiologically relevant in breast cancer cells, despite its low affinity, although *in vivo* studies are crucial to complement these findings and characterize the interaction.

Src plays a decisive role in pathways implicated in cell growth/cell-cycle control, differentiation, proliferation, survival, and motility in a variety of cells and tissues (216). Activation of Src has been correlated with the chemo-resistance of cancer cells (289) and Src has become a prime target for selective small molecule inhibitors (290), (194), (291). However, it has been shown that the transfection of Src alone does not have transforming ability of cancer cells (292) and early clinical studies with the Src inhibitors show that they are well tolerated but have minimal tumor response suggesting that the combination therapy (with the inhibitors of epidermal growth factor receptor family) might be more effective (293)

If the MUC1-CD and Src-SH3 domain interaction is proven specific to the MUC1-ICAM-1 induced cell migration pathway, it can be a definite target for anti-metastatic therapy. This is in contrast to the interaction between MUC1-CD and the Src-SH2 domain, which is involved in different signaling pathways in a variety of cancer cells. A protein-protein interaction between one structured and one unstructured partner are thought to have druggable features (294), such as a small molecule that would show tighter binding to the structured partner (Src) than the weak interaction by the disordered molecule (MUC1-CD). Since the

binding specificity of MUC1-CD and Src-SH3 domain is unique, as revealed by this study, it can be a potential target of higher-affinity small molecules that would uniquely inhibit MUC1-Src-SH3 interaction.

Chapter 4
BIBLIOGRAPHY

(1) Jemal A, Siegel R, Xu J, Ward E. Cancer statistics 2010. Cancer J Clin 2010 Sep-Oct; 60 (5): 277-300.

(2) Fidler IJ. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nat Rev Cancer 2003 Jun; 3 (6): 453-458.

(3) Weigelt B, Peterse JL, van 't Veer LJ. Breast cancer metastasis: markers and models. Nat Rev Cancer 2005 Aug; 5 (8): 591-602.

(4) Hovey RC, Trott JF, Vonderhaar BK. Establishing a framework for the functional mammary gland: from endocrinology to morphology. J Mammary Gland Biol Neoplasia 2002 Jan; 7 (1): 17-38.

(5) Howard BA, Gusterson BA. Human breast development. J Mammary Gland Biol Neoplasia 2000 Apr; 5 (2): 119-137.

(6) Lawrence RA, Lawrence RM. Anatomy of the human breast. Breastfeeding (Sixth Edition) Philadelphia: Mosby; 2005. p. 39-63.

(7) Williams JM, Daniel CW. Mammary ductal elongation: Differentiation of myoepithelium and basal lamina during branching morphogenesis. Dev Biol 1983 6; 97(2): 274-290.

(8) Wiseman BS, Werb Z. Stromal effects on mammary gland development and breast cancer. Science 2002; 296: 1046-1049.

(9) Lowery A, Miller N, Devaney A, McNeill R, Davoren P, Lemetre C, et al. MicroRNA signatures predict oestrogen receptor, progesterone receptor and HER2/neu receptor status in breast cancer. Breast Cancer Res 2009; 11(3): R27.

(10) Petersen OW, Polyak K. Stem Cells in the Human Breast. Cold Spring Harbor Perspectives in Biology 2010 May 01; 2(5).

(11) Weigelt B, Peterse JL, van 't Veer LJ. Breast cancer metastasis: markers and models. Nat Rev Cancer 2005 Aug; 5(8):591-602.

(12) van't Veer L, Dal H, Van DV, He YD, Hart AA, Mao M, et al. Gene

expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530-536.

(13) Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumors. Nature 2000 Aug 17;406(6797):747-752.

(14) Shiang C, Pusztai L. Molecular profiling contributes more than routine histology and immonohistochemistry to breast cancer diagnostics. Breast Cancer Res 2010 Dec 20;12 Suppl 4:S6.

(15) Gruvberger S, Ringner M, Chen Y, Panavally S, Saal LH, Borg A, et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression pattern. Cancer Res 2001; 61: 5979-5984.

(16) Sotiriou C, Neo S, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. P Natl Acad Sci USA 2003 September 02; 100 (18): 10393-10398.

(17) Hergueta-Redondo M, Palacios J, Cano A, Moreno-Bueno G. "New" molecular taxonomy in breast cancer. Clin Transl Oncol 2008 Dec; 10 (12): 777-785.

(18) Ibrahim E, Al-Gahmi AM, Zeenelin AA, Zekri JM, Elkhodary TR, Gaballa HE, et al. Basal vs. luminal A breast cancer subtypes: a matched case-control study using estrogen receptor, progesterone receptor, and HER-2 as surrogate markers. Med Oncol 2009; 26 (3): 372-378.

(19) Jumppanen M, Gruvberger-Saal S, Kauraniemi P, Tanner M, Bendahl P, Lundin M, et al. Basal-like phenotype is not associated with patient survival in estrogen-receptor-negative breast cancers. Breast Cancer Res 2007; 9 (1): R16.

(20) Pusztai L, Mazouni C, Anderson K, Wu Y, Symmans WF. Molecular Classification of Breast Cancer: Limitations and Potential. Oncologist 2006 September 1; 11(8): 868-877.

(21) Rugo HS. The importance of distant metastases in hormone-sensitive breast cancer. The Breast 2008 1; 17(Supplement 1): S3-S8.

(22) de Roos MA, van der Vegt B, Peterse JL, Patriarca C, de Vries J, de Bock GH, et al. The expression pattern of MUC1 (EMA) is related to tumour characteristics and clinical outcome in 'pure' ductal carcinoma in situ of the breast. Histopathology 2007 Aug; 51(2): 227-238.

(23) Mukhopadhyay P, Chakraborty S, Ponnusamy MP, Lakshmanan I, Jain M, Batra SK. Mucins in the Pathogenesis of Breast Cancer: Implications in Diagnosis, Prognosis and Therapy. Biochim Biophys Acta 2011 Jan 25.

(24) Schroeder JA, Masri AA, Adriance MC, Tessier JC, Kotlarczyk KL, Thompson MC, et al. MUC1 overexpression results in mammary gland tumorigenesis and prolonged alveolar differentiation. Oncogene 2004 Jul 29; 23(34): 5739-5747.

(25) Lee EYHP, Muller WJ. Oncogenes and Tumor Suppressor Genes. Cold Spring Harbor Perspectives in Biology 2010 October 01; 2(10).

(26) Copeland NG, Jenkins NA. Deciphering the genetic landscape of cancer – from genes to pathways. Trends in Genetics 2009 10; 25(10): 455-462.

(27) Osborne C, Wilson P, Tripathy D. Oncogenes and Tumor Suppressor Genes in Breast Cancer: Potential Diagnostic and Therapeutic Applications. Oncologist 2004 July 1;9(4):361-377.

(28) Hanahan D, Weinberg RA. The Hallmarks of Cancer. Cell 2000 1/7; 100 (1): 57-70.

(29) Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, et al. Genetic Heterogeneity and Penetrance Analysis of the BRCA1 and BRCA2 Genes in Breast Cancer Families. Am J Hum Genet 1998 3; 62(3): 676-689.

(30) Ingvarsson S. Molecular genetics of breast cancer progression. Semin Cancer Biol 1999 8; 9(4): 277-288.

(31) Ellisen MP, Haber MP. Hereditary Breast Cancer. Annu Rev Med 1998 02/01; 2011/01; 49(1): 425-436.

(32) Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, Higgins J, et al. Spectrum of Mutations in BRCA1, BRCA2, CHEK2, and TP53 in Families at High Risk of Breast Cancer. J Amer Med Assoc. March 22/29, 2006 March 22/29, 2006;295(12):1379-1388.

(33) Kenemans P, Verstraeten RA, Verheijen RHM. Oncogenic pathways in hereditary and sporadic breast cancer. Maturitas 2004 9/24; 49(1):34-43.

(34) Holli K. Hereditary breast cancer. Atlas Genet Cytogenet Oncol Haematol. November 2000.

(35) Pavelić K, Gall-Trošelj K. Recent advances in molecular genetics of breast cancer. J Mol Med 2001; 79 (10): 566-573.

(36) Venkitaraman AR. Cancer Susceptibility and the Functions of BRCA1 and BRCA2. Cell 2002 1/25; 108 (2): 171-182.

(37) Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, et al. Low-penetrance susceptibility to breast cancer due to CHEK2 (*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet 2002 May; 31(1):55-59.

(38) Nunes RA, Harris LN. The HER2 extracellular domain as a prognostic and predictive factor in breast cancer. Clin Breast Cancer 2002 Jun; 3 (2): 125-35; discussion 136-7.

(39) Liao D, Dickson R. c-Myc in breast cancer. Endocr Relat Cancer 2000 September 1; 7 (3):143-164.

(40) Sutherland RL, Watts CK, Musgrove EA. Cyclin gene expression and growth control in normal and neoplastic human breast epithelium. J Steroid Biochem Mol Biol 1993 Dec; 47 (1-6):99-106.

(41) Normanno N, Ciardiello F, Brandt R, Salomon DS. Epidermal growth factor-

related peptides in the pathogenesis of human breast cancer. Breast Cancer Res Treat 1994 Jan; 29 (1): 11-27.

(42) Lippman ME, Dickson RB, Gelmann EP, Rosen N, Knabbe C, Bates S, et al. Growth regulation of human breast carcinoma occurs through regulated growth factor secretion. J Cell Biochem 1987 Sep; 35 (1):1-16.

(43) Murphy LC, Dotzlaw H, Wong MS, Miller T, Mrockowski B, Gong Y, et al. Epidermal growth factor: receptor and ligand expression in human breast cancer. Semin Cancer Biol 1990 Oct; 1 (5): 305-315.

(44) Barcellos-Hoff MH, Akhurst RJ. Transforming growth factor-beta in breast cancer: too much, too late. Breast Cancer Res 2009; 11(1): 202.

(45) Sachdev D, Yee D. The IGF system and breast cancer. Endocr Relat Cancer 2001 September 1; 8(3): 197-209.

(46) Tuck AB, Park M, Sterns EE, Boag A, Elliott BE. Coexpression of hepatocyte growth factor and receptor (Met) in human breast carcinoma. Am J Pathol 1996 Jan; 148(1):225-232.

(47) Massarweh S, Schiff R. Resistance to endocrine therapy in breast cancer: exploiting estrogen receptor/growth factor signaling crosstalk. Endocr Relat Cancer 2006 December 1;13(Supplement_1):S15-24.

(48) Zaretsky JZ, Barnea I, Aylon Y, Gorivodsky M, Wreschner DH, Keydar I. MUC1 gene overexpressed in breast cancer: structure and transcriptional activity of the MUC1 promoter and role of estrogen receptor alpha (ERalpha) in regulation of the MUC1 gene expression. Mol Cancer 2006 Nov 5;5:57.

(49) Mego M, Mani SA, Cristofanilli M. Molecular mechanisms of metastasis in breast cancer-clinical applications. Nat Rev Clin Oncol 2010 Dec;7(12):693-701.

(50) Wijnhoven BP, Dinjens WN, Pignatelli M. E-cadherin-catenin cell-cell adhesion complex and human cancer. Br J Surg 2000 Aug;87(8):992-1005.

(51) Chabottaux V, Noel A. Breast cancer progression: insights into multifaceted

matrix metalloproteinases. Clin Exp Metastasis 2007;24(8):647-656.

(52) Andreasen PA, Kjoller L, Christensen L, Duffy MJ. The urokinase-type plasminogen activator system in cancer metastasis: a review. Int J Cancer 1997 Jul 3;72(1):1-22.

(53) Tomaskovic-Crook E, Thompson E, Thiery J. Epithelial to mesenchymal transition and breast cancer. Breast Cancer Research 2009;11(6):213.

(54) Sarrió D, Rodriguez-Pinilla SM, Hardisson D, Cano A, Moreno-Bueno G, Palacios J. Epithelial-Mesenchymal Transition in Breast Cancer Relates to the Basal-like Phenotype. Cancer Research 2008 February 15;68(4):989-997.

(55) Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. Nat Rev Cancer 2009 Apr;9(4):239-252.

(56) Friedl P, Gilmour D. Collective cell migration in morphogenesis, regeneration and cancer. Nat Rev Mol Cell Biol 2009 Jul;10(7):445-457.

(57) Kim LC, Song L, Haura EB. Src kinases as therapeutic targets for cancer. Nat Rev Clin Oncol 2009 Oct;6(10):587-595.

(58) Parsons SJ, Parsons JT. Src family kinases, key regulators of signal transduction. Oncogene 2004 Oct 18;23(48):7906-7909.

(59) Frame MC. Src in cancer: deregulation and consequences for cell behaviour. Biochim Biophys Acta 2002 Jun 21;1602(2):114-130.

(60) Machesky LM. Lamellipodia and filopodia in metastasis and invasion. FEBS Lett 2008 6/18;582(14):2102-2111.

(61) Mocellin S, Keilholz U, Rossi CR, Nitti D. Circulating tumor cells: the 'leukemic phase' of solid cancers. Trends Mol Med 2006 Mar;12(3):130-139.

(62) Paget S. The distribution of secondary growths in cancer of the breast. The Lancet 1889 3/23;133(3421):571-573.

(63) Kostenuik PJ. Revisiting the seed and soil theory of bone metastasis: new

tools, same answer. J Musculoskelet Neuronal Interact 2004 Dec;4(4):375-376.

(64) Moore MA. The role of chemoattraction in cancer metastases. Bioessays 2001 Aug;23(8):674-676.

(65) Muller A, Homey B, Soto H, Ge N, Catron D, Buchanan ME, et al. Involvement of chemokine receptors in breast cancer metastasis. Nature 2001 Mar 1;410(6824):50-56.

(66) Honn KV, Tang DG. Adhesion molecules and tumor cell interaction with endothelium and subendothelial matrix. Cancer Metastasis Rev 1992 Nov;11(3-4):353-375.

(67) Strell C, Entschladen F. Extravasation of leukocytes in comparison to tumor cells. Cell Communication and Signaling 2008;6(1):10.

(68) Osborn L. Leukocyte adhesion to endothelium in inflammation. Cell 1990 7/13;62(1):3-6.

(69) Rosette C, Roth RB, Oeth P, Braun A, Kammerer S, Ekblom J, et al. Role of ICAM1 in invasion of human breast cancer cells. Carcinogenesis 2005 May;26(5):943-950.

(70) Vogetseder W, Feichtinger H, Schulz TF, Schwaeble W, Tabaczewski P, Mitterer M, et al. Expression of 7F7-antigen, a human adhesion molecule identical to intercellular adhesion molecule-1 (ICAM-1) in human carcinomas and their stromal fibroblasts. Int J Cancer 1989 May 15;43(5):768-773.

(71) Regimbald LH, Pilarski LM, Longenecker BM, Reddish MA, Zimmermann G, Hugh JC. The breast mucin MUCI as a novel adhesion ligand for endothelial intercellular adhesion molecule 1 in breast cancer. Cancer Res 1996 Sep 15;56(18):4244-4249.

(72) Rahn JJ, Chow JW, Horne GJ, Mah BK, Emerman JT, Hoffman P, et al. MUC1 mediates transendothelial migration in vitro by ligating endothelial cell ICAM-1. Clin Exp Metastasis 2005;22(6):475-483.

(73) Senapati S, Das S, Batra SK. Mucin-interacting proteins: from function to therapeutics. Trends Biochem Sci 2010 Apr;35(4):236-245.

(74) Kufe DW. Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer 2009 Dec;9(12):874-885.

(75) Chaturvedi P, Singh AP, Batra SK. Structure, evolution, and biology of the MUC4 mucin. FASEB J 2008 Apr;22(4):966-981.

(76) Bafna S, Kaur S, Batra SK. Membrane-bound mucins: the mechanistic basis for alterations in the growth and survival of cancer cells. Oncogene 2010 May 20;29(20):2893-2904.

(77) Bafna S, Kaur S, Batra SK. Membrane-bound mucins: the mechanistic basis for alterations in the growth and survival of cancer cells. Oncogene 2010 May 20;29(20):2893-2904.

(78) Hollingsworth MA, Swanson BJ. Mucins in cancer: protection and control of the cell surface. Nat Rev Cancer 2004 Jan;4(1):45-60.

(79) Desseyn J, Aubert J, Porchet N, Laine A. Evolution of the Large Secreted Gel-Forming Mucins. Molecular Biology and Evolution 2000 August 01;17(8):1175-1184.

(80) Hattrup CL, Gendler SJ. Structure and function of the cell surface (tethered) mucins. Annu Rev Physiol 2008;70:431-457.

(81) Cullen PJ. Signaling mucins: the new kids on the MAPK block. Crit Rev Eukaryot Gene Expr 2007;17(3):241-257.

(82) Voynow JA, Rubin BK. Mucins, mucus, and sputum. Chest 2009 Feb;135(2):505-512.

(83) Gendler SJ, Burchell JM, Duhig T, Lamport D, White R, Parker M, et al. Cloning of partial cDNA encoding differentiation and tumor-associated mucin glycoproteins expressed by human mammary epithelium. Proc Natl Acad Sci U S A 1987 Sep;84(17):6060-6064.

(84) Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, Duhig T, Peat N, Burchell J, et al. Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. J Biol Chem 1990 Sep 5;265(25):15286-15293.

(85) Lan MS, Batra SK, Qi WN, Metzgar RS, Hollingsworth MA. Cloning and sequencing of a human pancreatic tumor mucin cDNA. J Biol Chem 1990 Sep 5;265(25):15294-15299.

(86) Patton S, Gendler SJ, Spicer AP. The epithelial mucin, MUC1, of milk, mammary gland and other tissues. Biochim Biophys Acta 1995 Dec 20;1241(3):407-423.

(87) Bieche I, Lidereau R. A gene dosage effect is responsible for high overexpression of the MUC1 gene observed in human breast tumors. Cancer Genetics & Cytogenetics 1997 Oct 1;98(1):75-80.

(88) Lacunza E, Baudis M, Colussi AG, Segal-Eiras A, Croce MV, Abba MC. MUC1 oncogene amplification correlates with protein overexpression in invasive breast carcinoma cells. Cancer Genet Cytogenet 2010 Sep;201(2):102-110.

(89) Zaretsky JZ, Barnea I, Aylon Y, Gorivodsky M, Wreschner DH, Keydar I. MUC1 gene overexpressed in breast cancer: structure and transcriptional activity of the MUC1 promoter and role of estrogen receptor alpha (ERalpha) in regulation of the MUC1 gene expression. Molecular Cancer 2006;5:57.

(90) Abba MC, Nunez MI, Colussi AG, Croce MV, Segal-Eiras A, Aldaz CM. GATA3 protein as a MUC1 transcriptional regulator in breast cancer cells. Breast Cancer Res 2006;8(6):R64.

(91) Lagow EL, Carson DD. Synergistic stimulation of MUC1 expression in normal breast epithelia and breast cancer cells by interferon-gamma and tumor necrosis factor-alpha. J Cell Biochem 2002;86(4):759-772.

(92) Rahn JJ, Dabbagh L, Pasdar M, Hugh JC. The importance of MUC1 cellular localization in patients with breast carcinoma: an immunohistologic study of 71 patients and review of the literature. Cancer 2001 Jun 1;91(11):1973-1982.

(93) Walsh MD, Luckie SM, Cummings MC, Antalis TM, McGuckin MA. Heterogeneity of MUC1 expression by human breast carcinoma cell lines in vivo and in vitro. Breast Cancer Research & Treatment 1999 Dec;58(3):255-266.

(94) McGuckin MA, Walsh MD, Hohn BG, Ward BG, Wright RG. Prognostic significance of MUC1 epithelial mucin expression in breast cancer. Hum Pathol 1995 Apr;26(4):432-439.

(95) McGuckin MA, Wright RG, McKenzie IF, Ward BG. Demonstration of two ovarian tumor-associated antigens in primary and metastatic breast, gastric, and colonic tumors. Am J Clin Pathol 1990 Aug;94(2):137-141.

(96) Dong Y, Walsh MD, Cummings MC, Wright RG, Khoo SK, Parsons PG, et al. Expression of MUC1 and MUC2 mucins in epithelial ovarian tumours. J Pathol 1997 Nov;183(3):311-317.

(97) Johansson DGA, Macao B, Sandberg A, Härd T. SEA Domain Autoproteolysis Accelerated by Conformational Strain: Mechanistic Aspects. J Mol Biol 2008 4/4;377(4):1130-1143.

(98) Macao B, Johansson DG, Hansson GC, Hard T. Autoproteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. Nature Structural & Molecular Biology 2006 Jan;13(1):71-76.

(99) Levitin F, Stern O, Weiss M, Gil-Henn C, Ziv R, Prokocimer Z, et al. The MUC1 SEA module is a self-cleaving domain. J Biol Chem 2005 Sep 30;280(39):33374-33386.

(100) Palmai-Pallag T, Khodabukus N, Kinarsky L, Leir SH, Sherman S, Hollingsworth MA, et al. The role of the SEA (sea urchin sperm protein, enterokinase and agrin) module in cleavage of membrane-tethered mucins. FEBS J 2005 Jun;272(11):2901-2911.

(101) Baldus SE, Engelmann K, Hanisch FG. MUC1 and the MUCs: a family of human mucins with impact in cancer biology. Crit Rev Clin Lab Sci 2004;41(2):189-231.

(102) Shen Q, University of Alberta. Dept. of Medical Sciences, University of Alberta. Dept. of Laboratory Medicine and Pathology. The role of MUC1ICAM-1 interaction in breast cancer cell motility. ; 2007.

(103) Gendler S, Taylor-Papadimitriou J, Duhig T, Rothbard J, Burchell J. A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up of tandem repeats. J Biol Chem 1988 Sep 15;263(26):12820-12823.

(104) Matsushima N, Creutz CE, Kretsinger RH. Polyproline, beta-turn helices. Novel secondary structures proposed for the tandem repeats within rhodopsin, synaptophysin, synexin, gliadin, RNA polymerase II, hordein, and gluten. Proteins 1990;7(2):125-155.

(105) Fontenot JD, Tjandra N, Bu D, Ho C, Montelaro RC, Finn OJ. Biophysical characterization of one-, two-, and three-tandem repeats of human mucin (muc-1) protein core. Cancer Res 1993 Nov 15;53(22):5386-5394.

(106) Schuman JT, Grinstead JS, Apostolopoulos V, Campbell AP. Structural and dynamic consequences of increasing repeats in a MUC1 peptide tumor antigen. Biopolymers 2005 Feb 5;77(2):107-120.

(107) Schuman J, Campbell AP, Koganty RR, Longenecker BM. Probing the conformational and dynamical effects of O-glycosylation within the immunodominant region of a MUC1 peptide tumor antigen. Journal of Peptide Research 2003 Mar;61(3):91-108.

(108) Karsten U, von Mensdorff-Pouilly S, Goletz S. What makes MUC1 a tumor antigen? Tumour Biology 2005 Jul-Aug;26(4):217-220.

(109) Croce MV, Isla-Larrain MT, Demichelis SO, Gori JR, Price MR, Segal-Eiras A. Tissue and serum MUC1 mucin detection in breast cancer patients. Breast Cancer Research & Treatment 2003 Oct;81(3):195-207.

(110) Parry S, Hanisch FG, Leir SH, Sutton-Smith M, Morris HR, Dell A, et al. N-Glycosylation of the MUC1 mucin in epithelial cells and secretions.

Glycobiology 2006 Jul;16(7):623-634.

(111) Kinarsky L, Suryanarayanan G, Prakash O, Paulsen H, Clausen H, Hanisch FG, et al. Conformational studies on the MUC1 tandem repeat glycopeptides: implication for the enzymatic O-glycosylation of the mucin protein core. Glycobiology 2003 Dec;13(12):929-939.

(112) Kirnarsky L, Prakash O, Vogen SM, Nomoto M, Hollingsworth MA, Sherman S. Structural effects of O-glycosylation on a 15-residue peptide from the mucin (MUC1) core protein. Biochemistry 2000 Oct 3;39(39):12076-12082.

(113) Brockhausen I, Yang JM, Burchell J, Whitehouse C, Taylor-Papadimitriou J. Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. European Journal of Biochemistry 1995 Oct 15;233(2):607-617.

(114) van de Wiel-van Kemenade E, Ligtenberg MJ, de Boer AJ, Buijs F, Vos HL, Melief CJ, et al. Episialin (MUC1) inhibits cytotoxic lymphocyte-target cell interaction. Journal of Immunology 1993 Jul 15;151(2):767-776.

(115) Altschuler Y, Kinlough CL, Poland PA, Bruns JB, Apodaca G, Weisz OA, et al. Clathrin-mediated endocytosis of MUC1 is modulated by its glycosylation state. Mol Biol Cell 2000 Mar;11(3):819-831.

(116) McDermott KM, Crocker PR, Harris A, Burdick MD, Hinoda Y, Hayashi T, et al. Overexpression of MUC1 reconfigures the binding properties of tumor cells. International Journal of Cancer 2001 Dec 15;94(6):783-791.

(117) Yuan Z, Wong S, Borrelli A, Chung MA. Down-regulation of MUC1 in cancer cells inhibits cell migration by promoting E-cadherin/catenin complex formation. Biochemical & Biophysical Research Communications 2007 Oct 26;362(3):740-746.

(118) Wesseling J, van der Valk SW, Hilkens J. A mechanism for inhibition of E-cadherin-mediated cell-cell adhesion by the membrane-associated mucin episialin/MUC1. Mol Biol Cell 1996 Apr;7(4):565-577.

(119) Alpaugh ML, Tomlinson JS, Kasraeian S, Barsky SH. Cooperative role of E-cadherin and sialyl-Lewis X/A-deficient MUC1 in the passive dissemination of tumor emboli in inflammatory breast carcinoma. Oncogene 2002 May 16;21(22):3631-3643.

(120) Kondo K, Kohno N, Yokoyama A, Hiwada K. Decreased MUC1 expression induces E-cadherin-mediated cell adhesion of breast cancer cell lines. Cancer Res 1998 May 1;58(9):2014-2019.

(121) Bork P, Patthy L. The SEA module: a new extracellular domain associated with O-glycosylation. Protein Sci 1995 Jul;4(7):1421-1425.

(122) Lillehoj EP, Han F, Kim KC. Mutagenesis of a Gly-Ser cleavage site in MUC1 inhibits ectodomain shedding. Biochemical & Biophysical Research Communications 2003 Aug 1;307(3):743-749.

(123) Thathiah A, Blobel CP, Carson DD. Tumor necrosis factor-alpha converting enzyme/ADAM 17 mediates MUC1 shedding. J Biol Chem 2003 Jan 31;278(5):3386-3394.

(124) Thathiah A, Carson DD. MT1-MMP mediates MUC1 shedding independent of TACE/ADAM17. Biochem J 2004 Aug 15;382(Pt 1):363-373.

(125) Williams CJ, Wreschner DH, Tanaka A, Tsarfaty I, Keydar I, Dion AS. Multiple protein forms of the human breast tumor-associated epithelial membrane antigen (EMA) are generated by alternative splicing and induced by hormonal stimulation. Biochemical & Biophysical Research Communications 1990 Aug 16;170(3):1331-1338.

(126) Smorodinsky N, Weiss M, Hartmann ML, Baruch A, Harness E, Yaakobovitz M, et al. Detection of a secreted MUC1/SEC protein by MUC1 isoform specific monoclonal antibodies. Biochemical & Biophysical Research Communications 1996 Nov 1;228(1):115-121.

(127) Rubinstein DB, Karmely M, Ziv R, Benhar I, Leitner O, Baron S, et al. MUC1/X protein immunization enhances cDNA immunization in generating anti-

MUC1 alpha/beta junction antibodies that target malignant cells. Cancer Res 2006 Dec 1;66(23):11247-11253.

(128) Baruch A, Hartmann M, Yoeli M, Adereth Y, Greenstein S, Stadler Y, et al. The breast cancer-associated MUC1 gene generates both a receptor and its cognate binding protein. Cancer Res 1999 Apr 1;59(7):1552-1561.

(129) Lodish HF. Molecular cell biology. 6th ed. New York: W.H. Freeman; 2008.

(130) Handa K, Jacobs F, Longenecker BM, Hakomori SI. Association of MUC-1 and SPGL-1 with low-density microdomain in T-lymphocytes: a preliminary note. Biochemical & Biophysical Research Communications 2001 Jul 20;285(3):788-794.

(131) Kinlough CL, McMahan RJ, Poland PA, Bruns JB, Harkleroad KL, Stremple RJ, et al. Recycling of MUC1 is dependent on its palmitoylation. J Biol Chem 2006 Apr 28;281(17):12112-12122.

(132) Pemberton LF, Rughetti A, Taylor-Papadimitriou J, Gendler SJ. The epithelial mucin MUC1 contains at least two discrete signals specifying membrane localization in cells. J Biol Chem 1996 Jan 26;271(4):2332-2340.

(133) Raina D, Ahmad R, Joshi MD, Yin L, Wu Z, Kawano T, et al. Direct targeting of the mucin 1 oncoprotein blocks survival and tumorigenicity of human breast carcinoma cells. Cancer Res 2009 Jun 15;69(12):5133-5141.

(134) Leng Y, Cao C, Ren J, Huang L, Chen D, Ito M, et al. Nuclear import of the MUC1-C oncoprotein is mediated by nucleoporin Nup62. J Biol Chem 2007 Jul 6;282(27):19321-19330.

(135) Carson DD. The cytoplasmic tail of MUC1: a very busy place. Science Signaling [Electronic Resource] 2008;1(27):e35.

(136) Zrihan-Licht S, Baruch A, Elroy-Stein O, Keydar I, Wreschner DH. Tyrosine phosphorylation of the MUC1 breast cancer membrane proteins.

Cytokine receptor-like molecules. FEBS Lett 1994 Dec 12;356(1):130-136.

(137) Carraway KL, Ramsauer VP, Haq B, Carothers Carraway CA. Cell signaling through membrane mucins. Bioessays 2003 Jan;25(1):66-71.

(138) Spicer AP, Duhig T, Chilton BS, Gendler SJ. Analysis of mammalian MUC1 genes reveals potential functionally important domains. Mamm Genome 1995 Dec;6(12):885-888.

(139) Li Y, Kuwahara H, Ren J, Wen G, Kufe D. The c-Src tyrosine kinase regulates signaling of the human DF3/MUC1 carcinoma-associated antigen with GSK3 beta and beta-catenin. J Biol Chem 2001 Mar 2;276(9):6061-6064.

(140) Li Y, Bharti A, Chen D, Gong J, Kufe D. Interaction of glycogen synthase kinase 3beta with the DF3/MUC1 carcinoma-associated antigen and beta-catenin. Molecular & Cellular Biology 1998 Dec;18(12):7216-7224.

(141) Ren J, Li Y, Kufe D. Protein kinase C delta regulates function of the DF3/MUC1 carcinoma antigen in beta-catenin signaling. J Biol Chem 2002 May 17;277(20):17616-17622.

(142) Li Y, Ren J, Yu W, Li Q, Kuwahara H, Yin L, et al. The epidermal growth factor receptor regulates interaction of the human DF3/MUC1 carcinoma antigen with c-Src and beta-catenin. J Biol Chem 2001 Sep 21;276(38):35239-35242.

(143) Schroeder JA, Thompson MC, Gardner MM, Gendler SJ. Transgenic MUC1 interacts with epidermal growth factor receptor and correlates with mitogen-activated protein kinase activation in the mouse mammary gland. J Biol Chem 2001 Apr 20;276(16):13057-13064.

(144) Ren J, Raina D, Chen W, Li G, Huang L, Kufe D. MUC1 oncoprotein functions in activation of fibroblast growth factor receptor signaling. Molecular Cancer Research: MCR 2006 Nov;4(11):873-883.

(145) Pandey P, Kharbanda S, Kufe D. Association of the DF3/MUC1 breast cancer antigen with Grb2 and the Sos/Ras exchange protein. Cancer Res 1995 Sep

15;55(18):4000-4003.

(146) Yamamoto M, Bharti A, Li Y, Kufe D. Interaction of the DF3/MUC1 breast carcinoma-associated antigen and beta-catenin in cell adhesion. J Biol Chem 1997 May 9;272(19):12492-12494.

(147) Li Y, Kufe D. The Human DF3/MUC1 carcinoma-associated antigen signals nuclear localization of the catenin p120(ctn). Biochemical & Biophysical Research Communications 2001 Feb 23;281(2):440-443.

(148) Li Y, Yu WH, Ren J, Chen W, Huang L, Kharbanda S, et al. Heregulin targets gamma-catenin to the nucleolus by a mechanism dependent on the DF3/MUC1 oncoprotein. Molecular Cancer Research: MCR 2003 Aug;1(10):765-775.

(149) Hattrup CL, Fernandez-Rodriguez J, Schroeder JA, Hansson GC, Gendler SJ. MUC1 can interact with adenomatous polyposis coli in breast cancer. Biochemical & Biophysical Research Communications 2004 Apr 2;316(2):364-369.

(150) Shen Q, Rahn JJ, Zhang J, Gunasekera N, Sun X, Shaw AR, et al. MUC1 initiates Src-CrkL-Rac1/Cdc42-mediated actin cytoskeletal protrusive motility after ligating intercellular adhesion molecule-1. Molecular Cancer Research: MCR 2008 Apr;6(4):555-567.

(151) Hattrup CL, Gendler SJ. MUC1 alters oncogenic events and transcription in human breast cancer cells. Breast Cancer Research 2006;8(4):R37.

(152) Li Y, Chen W, Ren J, Yu WH, Li Q, Yoshida K, et al. DF3/MUC1 signaling in multiple myeloma cells is regulated by interleukin-7. Cancer Biology & Therapy 2003 Mar-Apr;2(2):187-193.

(153) Mukherjee P, Tinder TL, Basu GD, Gendler SJ. MUC1 (CD227) interacts with lck tyrosine kinase in Jurkat lymphoma cells and normal T cells. J Leukoc Biol 2005 Jan;77(1):90-99.

(154) Raina D, Kharbanda S, Kufe D. The MUC1 oncoprotein activates the anti-apoptotic phosphoinositide 3-kinase/Akt and Bcl-xL pathways in rat 3Y1 fibroblasts. J Biol Chem 2004 May 14;279(20):20607-20612.

(155) Raina D, Kosugi M, Ahmad R, Panchamoorthy G, Rajabi H, Alam M, et al. Dependence on the MUC1-C Oncoprotein in Non–Small Cell Lung Cancer Cells. Molecular Cancer Therapeutics 2011 May 01;10(5):806-816.

(156) Singh PK, Hollingsworth MA. Cell surface-associated mucins in signal transduction. Trends Cell Biol 2006 Sep;16(9):467-476.

(157) Burack WR, Shaw AS. Signal transduction: hanging on a scaffold. Curr Opin Cell Biol 2000 4/1;12(2):211-216.

(158) Cortese MS, Uversky VN, Keith Dunker A. Intrinsic disorder in scaffold proteins: Getting more from less. Prog Biophys Mol Biol 2008 9;98(1):85-106.

(159) Dunker AK, Oldfield CJ, Meng J, Romero P, Yang JY, Chen JW, et al. The unfoldomics decade: an update on intrinsically disordered proteins. BMC Genomics 2008 Sep 16;9 Suppl 2:S1.

(160) Kufe DW. Functional targeting of the MUC1 oncogene in human cancers. Cancer Biol Ther 2009 Jul;8(13):1197-1203.

(161) Gendler SJ. MUC1, the renaissance molecule. Journal of Mammary Gland Biology & Neoplasia 2001 Jul;6(3):339-353.

(162) Schroeder JA, Adriance MC, Thompson MC, Camenisch TD, Gendler SJ. MUC1 alters beta-catenin-dependent tumor formation and promotes cellular invasion. Oncogene 2003 Mar 6;22(9):1324-1332.

(163) Pochampalli MR, Bitler BG, Schroeder JA. Transforming growth factor alpha dependent cancer progression is modulated by Muc1. Cancer Res 2007 Jul 15;67(14):6591-6598.

(164) Al Masri A, Gendler SJ. Muc1 affects c-Src signaling in PyV MT-induced mammary tumorigenesis. Oncogene 2005 Sep 1;24(38):5799-5808.

(165) Hattrup CL, Bradley JM, Kotlarczyk KL, Madsen CS, Hentz JG, Marler RJ, et al. The MUC1 Cytoplasmic Tail and Tandem Repeat Domains Contribute to Mammary Oncogenesis in FVB Mice. Breast Cancer (Auckl) 2008;1:57-63.

(166) Ren J, Bharti A, Raina D, Chen W, Ahmad R, Kufe D. MUC1 oncoprotein is targeted to mitochondria by heregulin-induced activation of c-Src and the molecular chaperone HSP90. Oncogene 2006 Jan 5;25(1):20-31.

(167) Lillehoj EP, Hyun SW, Kim BT, Zhang XG, Lee DI, Rowland S, et al. Muc1 mucins on the cell surface are adhesion sites for Pseudomonas aeruginosa. American Journal of Physiology - Lung Cellular & Molecular Physiology 2001 Jan;280(1):181-187.

(168) Lillehoj EP, Kim H, Chun EY, Kim KC. Pseudomonas aeruginosa stimulates phosphorylation of the airway epithelial membrane glycoprotein Muc1 and activates MAP kinase. American Journal of Physiology - Lung Cellular & Molecular Physiology 2004 Oct;287(4):809-815.

(169) Meerzaman D, Xing PX, Kim KC. Construction and characterization of a chimeric receptor containing the cytoplasmic domain of MUC1 mucin. American Journal of Physiology - Lung Cellular & Molecular Physiology 2000 Mar;278(3):625-629.

(170) Thompson EJ, Shanmugam K, Hattrup CL, Kotlarczyk KL, Gutierrez A, Bradley JM, et al. Tyrosines in the MUC1 cytoplasmic tail modulate transcription via the extracellular signal-regulated kinase 1/2 and nuclear factor-kappaB pathways. Molecular Cancer Research: MCR 2006 Jul;4(7):489-497.

(171) Wang H, Lillehoj EP, Kim KC. Identification of four sites of stimulated tyrosine phosphorylation in the MUC1 cytoplasmic tail. Biochemical & Biophysical Research Communications 2003 Oct 17;310(2):341-346.

(172) Verbeek BS, Adriaansen-Slot SS, Vroom TM, Beckers T, Rijksen G. Overexpression of EGFR and c-erbB2 causes enhanced cell migration in human breast cancer cells and NIH3T3 fibroblasts. FEBS Lett 1998 Mar 20;425(1):145-

150.

(173) Hiscox S, Morgan L, Green TP, Barrow D, Gee J, Nicholson RI. Elevated Src activity promotes cellular invasion and motility in tamoxifen resistant breast cancer cells. Breast Cancer Res Treat 2006 Jun;97(3):263-274.

(174) Li Y, Liu D, Chen D, Kharbanda S, Kufe D. Human DF3/MUC1 carcinoma-associated protein functions as an oncogene. Oncogene 2003 Sep 4;22(38):6107-6110.

(175) Rahn JJ, Shen Q, Mah BK, Hugh JC. MUC1 initiates a calcium signal after ligation by intercellular adhesion molecule-1. J Biol Chem 2004 Jul 9;279(28):29386-29390.

(176) Wesseling J, van der Valk SW, Vos HL, Sonnenberg A, Hilkens J. Episialin (MUC1) overexpression inhibits integrin-mediated cell adhesion to extracellular matrix components. J Cell Biol 1995 Apr;129(1):255-265.

(177) Ligtenberg MJ, Buijs F, Vos HL, Hilkens J. Suppression of cellular aggregation by high levels of episialin. Cancer Res 1992 Apr 15;52(8):2318-2324.

(178) Nelson WJ. Regulation of cell-cell adhesion by the cadherin-catenin complex. Biochem Soc Trans 2008 Apr;36(Pt 2):149-155.

(179) Hilkens J, Vos HL, Wesseling J, Boer M, Storm J, van der Valk S, et al. Is episialin/MUC1 involved in breast cancer progression? Cancer Lett 1995 Mar 23;90(1):27-33.

(180) Ciborowski P, Finn OJ. Non-glycosylated tandem repeats of MUC1 facilitate attachment of breast tumor cells to normal human lung tissue and immobilized extracellular matrix proteins (ECM) in vitro: potential role in metastasis. Clin Exp Metastasis 2002;19(4):339-345.

(181) Hilkens J, Ligtenberg MJ, Vos HL, Litvinov SV. Cell membrane-associated mucins and their adhesion-modulating property. Trends Biochem Sci 1992 Sep;17(9):359-363.

(182) Staunton DE, Marlin SD, Stratowa C, Dustin ML, Springer TA. Primary structure of ICAM-1 demonstrates interaction between members of the immunoglobulin and integrin supergene families. Cell 1988 3/25;52(6):925-933.

(183) Reilly PL, Woska JR,Jr, Jeanfavre DD, McNally E, Rothlein R, Bormann BJ. The native structure of intercellular adhesion molecule-1 (ICAM-1) is a dimer. Correlation with binding to LFA-1. J Immunol 1995 Jul 15;155(2):529-532.

(184) Yang Y, Jun CD, Liu JH, Zhang R, Joachimiak A, Springer TA, et al. Structural basis for dimerization of ICAM-1 on the cell surface. Mol Cell 2004 Apr 23;14(2):269-276.

(185) Casasnovas JM, Stehle T, Liu JH, Wang JH, Springer TA. A dimeric crystal structure for the N-terminal two domains of intercellular adhesion molecule-1. Proc Natl Acad Sci U S A 1998 Apr 14;95(8):4134-4139.

(186) Jun CD, Shimaoka M, Carman CV, Takagi J, Springer TA. Dimerization and the effectiveness of ICAM-1 in mediating LFA-1-dependent adhesion. Proc Natl Acad Sci U S A 2001 Jun 5;98(12):6830-6835.

(187) Bella J, Kolatkar PR, Marlor CW, Greve JM, Rossmann MG. The structure of the two amino-terminal domains of human ICAM-1 suggests how it functions as a rhinovirus receptor and as an LFA-1 integrin ligand. Proc Natl Acad Sci U S A 1998 Apr 14;95(8):4140-4145.

(188) Lawson C, Wolf S. ICAM-1 signaling in endothelial cells. Pharmacol Rep 2009 Jan-Feb;61(1):22-32.

(189) Hayashi T, Takahashi T, Motoya S, Ishida T, Itoh F, Adachi M, et al. MUC1 mucin core protein binds to the domain 1 of ICAM-1. Digestion 2001;63(Suppl 1):87-92.

(190) Kam JL, Regimbald LH, Hilgers JH, Hoffman P, Krantz MJ, Longenecker BM, et al. MUC1 synthetic peptide inhibition of intercellular adhesion molecule-1 and MUC1 binding requires six tandem repeats. Cancer Res 1998 Dec

1;58(23):5577-5581.

(191) Feller SM, Posern G, Voss J, Kardinal C, Sakkab D, Zheng J, et al. Physiological signals and oncogenesis mediated through Crk family adapter proteins. J Cell Physiol 1998 Dec;177(4):535-552.

(192) Anderson SK, Gibbs CP, Tanaka A, Kung HJ, Fujita DJ. Human cellular src gene: nucleotide sequence and derived amino acid sequence of the region coding for the carboxy-terminal two-thirds of pp60c-src. Mol Cell Biol 1985 May 1;5(5):1122-1129.

(193) Tanaka A, Gibbs CP, Arthur RR, Anderson SK, Kung HJ, Fujita DJ. DNA sequence encoding the amino-terminal region of the human c-src protein: implications of sequence divergence among src-type kinase oncogenes. Mol Cell Biol 1987 May;7(5):1978-1983.

(194) Finn RS. Targeting Src in breast cancer. Ann Oncol 2008 Aug;19(8):1379-1386.

(195) Parsons SJ, Parsons JT. Src family kinases, key regulators of signal transduction. Oncogene 2004 Oct 18;23(48):7906-7909.

(196) Summy JM, Gallick GE. Src family kinases in tumor progression and metastasis. Cancer Metastasis Rev 2003 Dec;22(4):337-358.

(197) Roskoski R,Jr. Src protein-tyrosine kinase structure and regulation. Biochem Biophys Res Commun 2004 Nov 26;324(4):1155-1164.

(198) Xu W, Harrison SC, Eck MJ. Three-dimensional structure of the tyrosine kinase c-Src. Nature 1997 Feb 13;385(6617):595-602.

(199) Cartwright CA, Kamps MP, Meisler AI, Pipas JM, Eckhart W. Pp60c-Src Activation in Human Colon Carcinoma. J Clin Invest 1989 Jun;83(6):2025-2033.

(200) Muthuswamy SK, Siegel PM, Dankort DL, Webster MA, Muller WJ. Mammary tumors expressing the neu proto-oncogene possess elevated c-Src tyrosine kinase activity. Mol Cell Biol 1994 Jan;14(1):735-743.

(201) Budde RJ, Ke S, Levin VA. Activity of pp60c-src in 60 different cell lines derived from human tumors. Cancer Biochem Biophys 1994 Oct;14(3):171-175.

(202) Masaki T, Okada M, Shiratori Y, Rengifo W, Matsumoto K, Maeda S, et al. pp60c-src activation in hepatocellular carcinoma of humans and LEC rats. Hepatology 1998 May;27(5):1257-1264.

(203) Masaki T, Shiratori Y, Okada H, Nishioka M, Taniguchi K, Hatanaka Y, et al. Pp60c-Src Activation in Gastric Carcinoma: a Preliminary Study. Am J Gastroenterol 2000 Mar;95(3):837-838.

(204) Rosen N, Bolen JB, Schwartz AM, Cohen P, DeSeau V, Israel MA. Analysis of pp60c-src protein kinase activity in human tumor cell lines and tissues. J Biol Chem 1986 Oct 15;261(29):13754-13759.

(205) Verbeek BS, Vroom TM, Adriaansen-Slot SS, Ottenhoff-Kalff AE, Geertzema JG, Hennipman A, et al. c-Src protein expression is increased in human breast cancer. An immunohistochemical and biochemical analysis. J Pathol 1996 Dec;180(4):383-388.

(206) Williams JC, Weijland A, Gonfloni S, Thompson A, Courtneidge SA, Superti-Furga G, et al. The 2.35 A crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. J Mol Biol 1997 Dec 19;274(5):757-775.

(207) Williams JC, Wierenga RK, Saraste M. Insights into Src kinase functions: structural comparisons. Trends Biochem Sci 1998 May;23(5):179-184.

(208) Cowan-Jacob SW, Fendrich G, Manley PW, Jahnke W, Fabbro D, Liebetanz J, et al. The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. Structure 2005 Jun;13(6):861-871.

(209) Lerner EC, Smithgall TE. SH3-dependent stimulation of Src-family kinase autophosphorylation without tail release from the SH2 domain in vivo. Nat Struct Biol 2002 May;9(5):365-369.

(210) Boonyaratanakornkit V, Scott MP, Ribon V, Sherman L, Anderson SM, Maller JL, et al. Progesterone receptor contains a proline-rich motif that directly interacts with SH3 domains and activates c-Src family tyrosine kinases. Mol Cell 2001 Aug;8(2):269-280.

(211) Masaki T, Okada M, Tokuda M, Shiratori Y, Hatase O, Shirai M, et al. Reduced C-terminal Src kinase (Csk) activities in hepatocellular carcinoma. Hepatology 1999 Feb;29(2):379-384.

(212) Cam WR, Masaki T, Shiratori Y, Kato N, Ikenoue T, Okamoto M, et al. Reduced C-terminal Src kinase activity is correlated inversely with pp60(c-src) activity in colorectal carcinoma. Cancer 2001 Jul 1;92(1):61-70.

(213) Irby RB, Mao W, Coppola D, Kang J, Loubeau JM, Trudeau W, et al. Activating SRC mutation in a subset of advanced human colon cancers. Nat Genet 1999 Feb;21(2):187-190.

(214) Dehm SM, Bonham K. SRC gene expression in human cancer: the role of transcriptional activation. Biochem Cell Biol 2004 Apr;82(2):263-274.

(215) Playford MP, Schaller MD. The interplay between Src and integrins in normal and tumor biology. Oncogene 2004 Oct 18;23(48):7928-7946.

(216) Thomas SM, Brugge JS. Cellular functions regulated by Src family kinases. Annu Rev Cell Dev Biol 1997;13:513-609.

(217) Boonyaratanakornkit V, Scott MP, Ribon V, Sherman L, Anderson SM, Maller JL, et al. Progesterone receptor contains a proline-rich motif that directly interacts with SH3 domains and activates c-Src family tyrosine kinases. Mol Cell 2001 Aug;8(2):269-280.

(218) Patwardhan P, Resh MD. Myristoylation and membrane binding regulate c-Src stability and kinase activity. Mol Cell Biol 2010 Sep;30(17):4094-4107.

(219) Agrawal V, Kishan KV. Promiscuous binding nature of SH3 domains to their target proteins. Protein Pept Lett 2002 Jun;9(3):185-193.

(220) Ayrapetov MK, Wang YH, Lin X, Gu X, Parang K, Sun G. Conformational basis for SH2-Tyr(P)527 binding in Src inactivation. J Biol Chem 2006 Aug 18;281(33):23776-23784.

(221) Arias-Salgado EG, Lizano S, Sarkar S, Brugge JS, Ginsberg MH, Shattil SJ. Src kinase activation by direct interaction with the integrin beta cytoplasmic domain. Proc Natl Acad Sci U S A 2003 Nov 11;100(23):13298-13302.

(222) Superti-Furga G, Fumagalli S, Koegl M, Courtneidge SA, Draetta G. Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. EMBO J 1993 Jul;12(7):2625-2634.

(223) Tessari M, Gentile LN, Taylor SJ, Shalloway DI, Nicholson LK, Vuister GW. Heteronuclear NMR studies of the combined Src homology domains 2 and 3 of pp60 c-Src: effects of phosphopeptide binding. Biochemistry 1997 Nov 25;36(47):14561-14571.

(224) Thomas JW, Ellis B, Boerner RJ, Knight WB, White GC,2nd, Schaller MD. SH2- and SH3-mediated interactions between focal adhesion kinase and Src. J Biol Chem 1998 Jan 2;273(1):577-583.

(225) Alexandropoulos K, Baltimore D. Coordinate activation of c-Src by SH3- and SH2-binding sites on a novel p130Cas-related protein, Sin. Genes Dev 1996 Jun 1;10(11):1341-1355.

(226) Yadav SS, Miller WT. Cooperative activation of Src family kinases by SH3 and SH2 ligands. Cancer Lett 2007 Nov 8;257(1):116-123.

(227) Pawson T. Protein modules and signalling networks. Nature 1995 Feb 16;373(6515):573-580.

(228) Dalgarno DC, Botfield MC, Rickles RJ. SH3 domains and drug design: ligands, structure, and biological function. Biopolymers 1997;43(5):383-400.

(229) Feng S, Chen JK, Yu H, Simon JA, Schreiber SL. Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-

ligand interactions. Science 1994 Nov 18;266(5188):1241-1247.

(230) Pawson T, Nash P. Assembly of cell regulatory systems through protein interaction domains. Science 2003 Apr 18;300(5618):445-452.

(231) Musacchio A, Noble M, Pauptit R, Wierenga R, Saraste M. Crystal structure of a Src-homology 3 (SH3) domain. Nature 1992 Oct 29;359(6398):851-855.

(232) Yu H, Rosen MK, Shin TB, Seidel-Dugan C, Brugge JS, Schreiber SL. Solution structure of the SH3 domain of Src and identification of its ligand-binding site. Science 1992;258(5088):1665-1668.

(233) Feng S, Kasahara C, Rickles RJ, Schreiber SL. Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. Proc Natl Acad Sci U S A 1995 Dec 19;92(26):12408-12415.

(234) Waksman G, Shoelson SE, Pant N, Cowburn D, Kuriyan J. Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. Cell 1993 Mar 12;72(5):779-790.

(235) Mayer BJ, Eck MJ. SH3 domains. Minding your p's and q's. Curr Biol 1995 Apr 1;5(4):364-367.

(236) Yu H, Chen JK, Feng S, Dalgarno DC, Brauer AW, Schreiber SL. Structural basis for the binding of proline-rich peptides to SH3 domains. Cell 1994 Mar 11;76(5):933-945.

(237) Li SS. Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. Biochem J 2005 Sep 15;390(Pt 3):641-653.

(238) Rickles RJ, Botfield MC, Weng Z, Taylor JA, Green OM, Brugge JS, et al. Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. EMBO J 1994 Dec 1;13(23):5598-5604.

(239) Arold S, O'Brien R, Franken P, Strub MP, Hoh F, Dumas C, et al. RT loop

flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef. Biochemistry 1998 Oct 20;37(42):14683-14691.

(240) Lewitzky M, Harkiolaki M, Domart MC, Jones EY, Feller SM. Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif, R/KXXK, with a classical PXXP motif embedded in a polyproline type II (PPII) helix. J Biol Chem 2004 Jul 2;279(27):28724-28732.

(241) Kami K, Takeya R, Sumimoto H, Kohda D. Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p. EMBO J 2002 Aug 15;21(16):4268-4276.

(242) Vaynberg J, Fukuda T, Chen K, Vinogradova O, Velyvis A, Tu Y, et al. Structure of an ultraweak protein-protein complex and its crucial role in regulation of cell morphology and motility. Mol Cell 2005 Feb 18;17(4):513-523.

(243) Lee CH, Leung B, Lemmon MA, Zheng J, Cowburn D, Kuriyan J, et al. A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein. EMBO J 1995 Oct 16;14(20):5006-5015.

(244) Sparks AB, Rider JE, Hoffman NG, Fowlkes DM, Quillam LA, Kay BK. Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. Proc Natl Acad Sci U S A 1996 Feb 20;93(4):1540-1544.

(245) Kang H, Freund C, Duke-Cohan JS, Musacchio A, Wagner G, Rudd CE. SH3 domain recognition of a proline-independent tyrosine-based RKxxYxxY motif in immune cell adaptor SKAP55. EMBO J 2000 Jun 15;19(12):2889-2899.

(246) Mongiovi AM, Romano PR, Panni S, Mendoza M, Wong WT, Musacchio A, et al. A novel peptide-SH3 interaction. EMBO J 1999 Oct 1;18(19):5300-5309.

(247) Freund C, Kuhne R, Yang H, Park S, Reinherz EL, Wagner G. Dynamic interaction of CD2 with the GYF and the SH3 domain of compartmentalized effector molecules. EMBO J 2002 Nov 15;21(22):5985-5995.

(248) Kojima C, Hashimoto A, Yabuta I, Hirose M, Hashimoto S, Kanaho Y, et al. Regulation of Bin1 SH3 domain binding by phosphoinositides. EMBO J 2004 Nov 10;23(22):4413-4422.

(249) Douangamath A, Filipp FV, Klein AT, Barnett P, Zou P, Voorn-Brouwer T, et al. Topography for independent binding of alpha-helical and PPII-helical ligands to a peroxisomal SH3 domain. Mol Cell 2002 Nov;10(5):1007-1017.

(250) Barnett P, Bottger G, Klein AT, Tabak HF, Distel B. The peroxisomal membrane protein Pex13p shows a novel mode of SH3 interaction. EMBO J 2000 Dec 1;19(23):6382-6391.

(251) Hiroaki H, Ago T, Ito T, Sumimoto H, Kohda D. Solution structure of the PX domain, a target of the SH3 domain. Nat Struct Biol 2001 Jun;8(6):526-530.

(252) Chan B, Lanyi A, Song HK, Griesbach J, Simarro-Grande M, Poy F, et al. SAP couples Fyn to SLAM immune receptors. Nat Cell Biol 2003 Feb;5(2):155-160.

(253) Pawson T. Organization of cell-regulatory systems through modular-protein-interaction domains. Philos Transact A Math Phys Eng Sci 2003 Jun 15;361(1807):1251-1262.

(254) Pawson T, Gish GD. SH2 and SH3 domains: From structure to function. Cell 1992 10/30;71(3):359-362.

(255) Liu X, Pawson T. Biochemistry of the Src protein-tyrosine kinase: regulation by SH2 and SH3 domains. Recent Prog Horm Res 1994;49:149-160.

(256) Pawson T, Olivier P, Rozakis-Adcock M, McGlade J, Henkemeyer M. Proteins with SH2 and SH3 domains couple receptor tyrosine kinases to intracellular signalling pathways. Philos Trans R Soc Lond B Biol Sci 1993 Jun 29;340(1293):279-285.

(257) Yaffe MB. Phosphotyrosine-binding domains in signal transduction. Nat Rev Mol Cell Biol 2002 Mar;3(3):177-186.

(258) Bibbins KB, Boeuf H, Varmus HE. Binding of the Src SH2 domain to phosphopeptides is determined by residues in both the SH2 domain and the phosphopeptides. Mol Cell Biol 1993 Dec;13(12):7278-7287.

(259) González L, Agulló-Ortuño MT, García-Martínez JM, Calcabrini A, Gamallo C, Palacios J, et al. Role of c-Src in Human MCF7 Breast Cancer Cell Tumorigenesis. Journal of Biological Chemistry 2006 July 28;281(30):20851-20864.

(260) Jallal H, Valentino ML, Chen G, Boschelli F, Ali S, Rabbani SA. A Src/Abl kinase inhibitor, SKI-606, blocks breast cancer invasion, growth, and metastasis in vitro and in vivo. Cancer Res 2007 Feb 15;67(4):1580-1588.

(261) Clegg RA. Protein targeting protocols. Totowa, N.J.: Humana Press; 1998.

(262) Vaynberg J, Qin J. Weak protein-protein interactions as probed by NMR spectroscopy. Trends Biotechnol 2006 Jan;24(1):22-27.

(263) Meyer B, Peters T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. Angew Chem Int Ed Engl 2003 Feb 24;42(8):864-890.

(264) Fielding L. NMR methods for the determination of protein-ligand dissociation constants. Curr Top Med Chem 2003;3(1):39-53.

(265) Clarkson J, Campbell ID. Studies of protein-ligand interactions by NMR. Biochem Soc Trans 2003 Oct;31(Pt 5):1006-1009.

(266) Harding SE, Chowdhry BZ. Protein-ligand interactions, structure and spectroscopy: a practical approach. Oxford ; New York: Oxford University Press; 2001.

(267) Lindfors HE, de Koning PE, Drijfhout JW, Venezia B, Ubbink M. Mobility of TOAC spin-labelled peptides binding to the Src SH3 domain studied by paramagnetic NMR. J Biomol NMR 2008 Jul;41(3):157-167.

(268) Jacobsen NE. NMR spectroscopy explained: simplified theory, applications

and examples for organic chemistry and structural biology. Hoboken, N.J.: Wiley-Interscience; 2007.

(269) Lepre CA, Moore JM, Peng JW. Theory and applications of NMR-based screening in pharmaceutical research. Chem Rev 2004 Aug;104(8):3641-3676.

(270) Hill JM. NMR screening for rapid protein characterization in structural proteomics. Methods Mol Biol 2008;426:437-446.

(271) Nietlispach D, Mott HR, Stott KM, Nielsen PR, Thiru A, Laue ED. Structure determination of protein complexes by NMR. Methods Mol Biol 2004;278:255-288.

(272) W©othrich K. NMR of proteins and nucleic acids. New York: Wiley; 1986.

(273) Smith DB, Corcoran LM. Expression and purification of glutathione-S-transferase fusion proteins. Curr Protoc Mol Biol 2001 May;Chapter 16:Unit16.7.

(274) Harper S, Speicher DW. Expression and purification of GST fusion proteins. Curr Protoc Protein Sci 2001 May;Chapter 6:Unit 6.6.

(275) Wang C, Pawley NH, Nicholson LK. The role of backbone motions in ligand binding to the c-Src SH3 domain. J Mol Biol 2001 Nov 2;313(4):873-887.

(276) Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995 Nov;6(3):277-293.

(277) Boyko R. and Sykes B.D. xcrvfit: A graphical X-windows program for binding curve studies and NMR spectroscopic analysis (University of Alberta, Canada) 1994:http://www.bionmr.ualberta.ca/bds/software/xcrvfit.

(278) Morton CJ, Pugh DJ, Brown EL, Kahmann JD, Renzoni DA, Campbell ID. Solution structure and peptide binding of the SH3 domain from human Fyn. Structure 1996 Jun 15;4(6):705-714.

(279) Karatan E, Merguerian M, Han Z, Scholle MD, Koide S, Kay BK.

Molecular recognition properties of FN3 monobodies that bind the Src SH3 domain. Chem Biol 2004 Jun;11(6):835-844.

(280) Kukic P, Nielsen JE. Electrostatics in proteins and protein-ligand complexes. Future Med Chem 2010 Apr;2(4):647-666.

(281) Ahmad M, Gu W, Helms V. Mechanism of fast peptide recognition by SH3 domains. Angew Chem Int Ed Engl 2008;47(40):7626-7630.

(282) Miyazaki K, Senga T, Matsuda S, Tanaka M, Machida K, Takenouchi Y, et al. Critical amino acid substitutions in the Src SH3 domain that convert c-Src to be oncogenic. Biochem Biophys Res Commun 1999 Oct 5;263(3):759-764.

(283) Hochrein JM, Lerner EC, Schiavone AP, Smithgall TE, Engen JR. An examination of dynamics crosstalk between SH2 and SH3 domains by hydrogen/deuterium exchange and mass spectrometry. Protein Sci 2006 Jan;15(1):65-73.

(284) Gonfloni S, Williams JC, Hattula K, Weijland A, Wierenga RK, Superti-Furga G. The role of the linker between the SH2 domain and catalytic domain in the regulation and function of Src. EMBO J 1997 Dec 15;16(24):7261-7271.

(285) Ozkirimli E, Yadav SS, Miller WT, Post CB. An electrostatic network and long-range regulation of Src kinases. Protein Sci 2008 Nov;17(11):1871-1880.

(286) Mayer BJ. SH3 domains: complexity in moderation. J Cell Sci 2001 Apr;114(Pt 7):1253-1263.

(287) Urquhart AJ, Kennedy D, Gould SJ, Crane DI. Interaction of Pex5p, the type 1 peroxisome targeting signal receptor, with the peroxisomal membrane proteins Pex14p and Pex13p. J Biol Chem 2000 Feb 11;275(6):4127-4136.

(288) Serrano L. Modular Protein Domains Edited by G. Cesareni, M. Gimona, M. Sudol and M. Yaffe. Nature cell biology 2005;7(7):651-652.

(289) Shah AN, Gallick GE. Src, chemoresistance and epithelial to mesenchymal transition: are they related? Anticancer Drugs 2007 Apr;18(4):371-375.

(290) Araujo J, Logothetis C. Dasatinib: a potent SRC inhibitor in clinical development for the treatment of solid tumors. Cancer Treat Rev 2010 Oct;36(6):492-500.

(291) Mayer EL, Krop IE. Advances in targeting SRC in the treatment of breast cancer and other solid malignancies. Clin Cancer Res 2010 Jul 15;16(14):3526-3532.

(292) Shalloway D, Coussens PM, Yaciuk P. Overexpression of the c-src protein does not induce transformation of NIH 3T3 cells. Proc Natl Acad Sci U S A 1984 Nov;81(22):7071-7075.

(293) Aleshin A, Finn RS. SRC: a century of science brought to the clinic. Neoplasia 2010 Aug;12(8):599-607.

(294) Cheng Y, LeGall T, Oldfield CJ, Mueller JP, Van YY, Romero P, et al. Rational drug design via intrinsically disordered protein. Trends Biotechnol 2006 Oct;24(10):435-442.

(295) Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. J Mol Biol 1993 Jul 20;232(2):584-599.

Chapter 5
APPENDICES

Appendix 1

Uniformly ^{15}N labeled, Protein Expression Protocol

(GST tagged-Src-SH3/SH2 and SH3+SH2)

Materials:

- PGEX-2T-SH3/SH2 Transformants
- Sterile Tooth picks
- Sterile 10 ml Tubes
- Sterile 2L Baffled Flasks

Solutions:

1) *M9 Salts (10X)*
   - Weigh M9 salts (120 g Na2HPO4; 60 g KH2PO4 in 1L H2O)
   - Dissolve in 500 – 700 ml H2O. Warm the solution while constantly stirring since Na2HPO4 does not dissolve well at room temperature.
   - Adjust pH to 7.3
   - Adjust volume to 1L while checking for pH (readjust pH to 7.3 if necessary)
   - Autoclave
2) *1M MgSo4 (100 ml):*
   - Dissolve 12.06 g in 100 ml of H2O. Autoclave
3) *0.1 M CaCl2 (100 ml)*
   - Dissolve 0.11 g in 100 ml of H2O. Autoclave
4) *1mM FeSO4*
   - Dissolve 0.152 g in 100 ml H2O. Autoclave
5) *1g $^{15}NH4Cl$*
   - Dissolve in 10 ml of autoclaved 1X M9 media and filter with 0.22 uM filter into the same (1L) M9 media (**DO NOT Autoclave**).
6) *10 g Glucose*
   - Dissolve in 20 ml of autoclaved 1X M9 media and filter with 0.22 uM filter into the same (1L) M9 media (**DO NOT Autoclave**)

7) *Thiamin HCL (10mg/ml)*
   - Dissolve in ddH2O and/or sterile filter. Store at 4 C. (**DO NOT Autoclave**)
8) *Biotin (10mg/ml)*
   - Dissolve in sterile H2O. Do not filter sterilize as this concentration is above the level of solubility of Biotin. Store at 4 C. (**DO NOT Autoclave**)
9) *Autoclaved MQ H2O*
10) *100ug/ml Ampicillin (-20 C stock) – filter sterilize using a 0.22 uM filter*
11) *25ug/ml Chloramphenicol (- 20C stock)*
12) *500 mM IPTG (-20C Stock) – filter sterilize using a 0.22 uM filter*
13) *LB media (on Shelf)* - Make by dissolving an appropriate amount of Luria Broth in MQ-H2O. Autoclave and store at RT.

Methods

A) Make 1L M9 Minimal Media

1) Dilute 10X minimal salts into 1X. (Add 900 ml autoclaved H2O to 100 ml 10X M9 salts)
2) Add 4 ml of 1M MgSO4
3) Add 1.8ml of 1mM FeSO4
4) Add 1g of 15NH4Cl. Dissolve 1g in 10 ml of 1X M9 media (from step 1) and filter with 0.22. Filter into the same bottle (with 1L M9 media).
5) Add 10g Glucose. Dissolve 10g in 20 ml of 1X M9 media (from step 1) and filter with 0.22. Filter into the same bottle (with 1L M9 media).
6) Add 1ml of 0.1M CaCl2
7) Add 1ml of 10mg/ml Biotin (vortex well before adding this)
8) Add 1ml of 10mg/ml Thiamin-HCL
9) Add 1 ml of 100ug/ml Ampicillin
10) Add 1 ml of 25ug/ml Chloramphenicol

11) Stir at RT for about 10 minutes before use.
12) Divide into 2 (500 ml each) and put into 2-litre baffled flasks and inoculate with 5 ml O/N (overnight) cultures.

B) Expression Scale-up Procedure

1) Inoculate a recombinant, PGEX-2T-SH3 (or SH3+SH2), *E.coli* colony into 6 ml LB media with 100 μg/ml Ampicillin and 25 μg/ml Chloramphenicol; grow at 37^0 C overnight at 250 rpm.
2) Inoculate 1 Liter (in 2 flasks, 50 ml each) M9 media + 100 μg/ml Ampicillin + 25 μg/ml Chloramphenicol. Add 1ml of O/N culture for every 100 ml of M9 media.
3) Grow until OD_{600} reaches 0.65 (6 - 8 hours) at 37^0 C - 250 rpm.
4) Add 500mM IPTG to 1 mM (final) concentration (1ml for 500 ml) to induce protein expression.
5) Incubate the culture for 3 - 4 hours.
6) At the time of harvest, pour cells into 500 ml Beckman centrifuge vials and balance roughly.
7) Pellet cells at 4000 rpm for 15 min at 4^0 C in a Beckman centrifuge.
8) Discard media and store the cell pellets at -80 C for later use.

Appendix 2

Cell Lysis Protocol

Materials:

- GST binding buffer (1XPBS pH 7.3)
- 100 X protease inhibitor cocktail (Calbiochem)
- 100 mM PMSF
- DNAse
- 20% Triton-X100

Solutions:

i) GST binding Buffer (1X PBS)

- 8.18g NaCl (140mM)
- 0.2 g KCl (2.7 mM)
- 1.419 g Na2HPO4 (10 mM)
- 0.25 g KH2PO4 (1.8 mM)
- Dissolve in 500 – 700 ml autoclaved H2O
- Adjust pH to 7.3
- Adjust volume to 1L

ii) 100 X protease inhibitor cocktail: - -20 C freezer stock

- Dissolve lyophilized powder in 1 ml sterile H2O for 100X protease inhibitor cocktail

iii) 100 mM PMSF (MW = 174.2)

- Dissolve 0.174 g in 10 ml Isopropanol and store at 4 C

iv) DNAse (Invitrogen) - 20,000 units (50 -375 U/ul)

v) 20% (v/v) Triton-X100

- Dissolve 20 ml of Triton X100 in 80 ml of sterile H2O
  Prepare this solution well in advance (Triton X-100 is very viscous and

takes time to dissolve completely. Mix by vigorous shaking and warm to 37° C briefly) and store in a dark place (cover with an Aluminum foil) to prevent photo-oxidation.

vi) Autoclaved MQ-H2O (ddH2O)

Method

1) Thaw cell pellets on ice (1L culture will yield 4 pellets, one in each 500 ml Beckman centrifuge vial stored at -80 C) and resuspend each pellet in 5mL of GST binding buffer.
2) Vortex and collect cell pellets into a 50 ml beaker.
3) Add 100 X protease inhibitor cocktail (Calbiochem) + 1mM PMSF (final concentration) + 25ul DNAse while the cells are still on ice.
4) Crush the cells using a spatula and leave at least one hour on ice (while stirring time to time) to aid cell lyses and homogenization.
5) Use a French Press or an Emulsiflex to lyse the cells. Collect 100ul of lysate for SDSPAGE.
6) Add 20% Triton X-100 to a final concentration of 1% v/v. Mix gently for 1 hour to aid in solubilization of the fusion protein. Collect 100ul of lysate for SDSPAGE.
7) Centrifuge at 12,000 × g (e.g. 10 000 rpm in a Beckman JA20 rotor) for 10 min at 4° C.
8) Transfer the supernatant to a fresh 50 ml sterile tube.
9) Resuspend the pellet in 5.00 ml of GST binding buffer.
10) Save 100ul aliquots of the supernatant and the pellet to check the solubility of recombinant protein by SDS-PAGE.
11) Freeze the supernatant and pellet as quickly as possible in liquid nitrogen. Store at -80° C

Appendix 3

Tris-Tricine SDSPAGE Protocol

Materials

- Tris Tricine gel Buffer
- 1X Cathode Buffer
- 1X Anode Buffer
- Glycerol
- 29:1 acrylamide/bisacrylamide
- TEMED
- 25% (w/v) Amonium Persulphate (APS)
- MQ H2O
- 4X sample buffer
- BME (β-Mercaptoethanol)

Solutions

10 X Cathode Buffer

- 121.1 g Tris base
- 179.2 g Tricin
- 10 g SDS
- Dissolve in 1L ddH2O ***Do not adjust pH
- Store at 40 C
- Final concentrations (1X): - 0.1M Tris, 0.1M Tricine, 0.1% SDS

5 X Anode Buffer

- Dissolve 121.1 g Tris base in 500 ml H2O
- Adjust to pH 8.9 with concentrated HCl
- Dilute to 1 liter with ddH2O
- Store at 40 C

- Final concentration (1X): - 0.2M Tris-Cl, pH 8.9

Tris Tricine gel Buffer (3M Tris-Cl, 0.3% SDS, pH8.45)

- Dissolve 182 g Tris base in 300ml ddH2O
- Adjust to pH 8.45
- Add H2O to 500ml total volume
- Add 1.5 g SDS Store at 40 C

Method:

Recipe for making 5 (1mm x 8 cm x 10 cm) gels

Stock Solutions	**Separating Gel**	**Stacking Gel**
1) 29:1 Acrylamide/bisacrylamide….....	10.86 ml	2.42 ml
2) Tris-Tricine gel Buffer …….....	10.00 ml	6.2 ml
3) H2O …….....	5.97 ml	16.38 ml
4) Glycerol ……...…….....	3.17 ml	---------
5) 25%(w/v) Ammonium Persulfate.....	50 ul (fresh)	50 ul (fresh)
6) TEMED …….	15 ul	30 ul

Recipe was adapted from www.aci.uzh.ch/MT/pdf/Tris-Tricine.pdf

Casting Gels

Choose the long glass plate that match with the comb (of thickness 0.75 mm, 1.0mm or 1.5 mm combs) and a short (BioRad) glass plate.

1) Thoroughly clean glass plates prior to casting the gel (wipe with 70% EtOH until no debris or precipitates are visible).
2) Set up glass plate sandwich in gel casting stand. (Bio-rad Mini PROTEAN Electrophoresis System).
3) In a 50 ml tube mix all components for separating gels except for 25% APS and TEMED.
4) Add 30% APS. Mix by inverting a few times. Add TEMED to. Mix by inverting a few times and working quickly pour the gel in between the glass

plates until solution is about ¾ of the way to the top. Gel will begin to polymerize now.

5) Overlay with water saturated isobutanol (just enough to cover the top).
6) Allow polymerizing (leave the remaining gel solution in the tube and use it as a guide to determine when it's polymerized. It usually takes about 30-45 min.)
7) After gel has polymerized, dump off water and isobutanol. Wash with distilled water and remove all water by placing Kimwipes placed in the corner of the sandwich.
8) In a 50 ml tube mix all components for the stacking gel
9) Add 25% APS to stacking gel. Mix by inverting. Add TEMED and mix by inverting.
10) Quickly pour the stacking gel until gel reaches the very top of the glass plate.
11) Insert combs and allow the stacking gel to polymerize. (It takes 30 - 45 min.).
12) Proceed to running the gel or wrap the gel in damp paper towels and saran wrap and store in the 4° C fridge for up to a week.

Preparing Cell Lysates

1) Add 27 ul of sample buffer and 7 ul of BME into 100 ul of each fraction (cell lysate).
2) Boil for 5 minutes
3) Load the gels right away or store at -20 C until used.

Assembling, Loading and Running Gels

1) Place 2 gels (or 1 gel and 1 buffer dam) into clamping frame and electrode assembly.
2) Place the assembly in the electrophoresis tank.
3) Fill the inner chamber with 1X Cathode Buffer. Make sure that this buffer is filled up above the top of the lower glass plate.
4) Fill the outer chamber with 1X Anode Buffer.

5) After boiling the samples, load them in the wells. Use 5-8 µL of Bio-Rad Precision Plus All Blue Standard (or an equivalent based on the expected size of the target recombinant protein). Load equivalent amounts of protein/well for samples. (20 µL will fit into each well of a 15-well comb. 30-35 µL will fit into each well of a 10-well comb).
6) Run at 70V constant until samples have completely run through the stacking gel and are in the separating gel (about 1 hour).
7) Run at 100V constant until ion front reaches the bottom of the gel (about 2 hours).

Coomassie Stain Protocol

Adapted from - (http://www.ccic.ohiostate.edu/MS/Files/Coomassie%20Stain%20Protocol.doc)

Reagents

1. ***Gel-fixing solution:***
   - Add 500 mL of USP-grade 95% (v/v) ethanol to 300 mL of water.
   - Add 100 mL of acetic acid and adjust the total volume to 1000 mL with water.
   - The final concentrations are 50% (v/v) ethanol in water with 10% (v/v) acetic acid.

2. ***Gel-washing solution:***
   - Add 500mL of HPLC-grade methanol to 300 mL of water.
   - Add 100mL of acetic acid and adjust the total volume to 1000 mL with water.
   - The final concentrations are 50% (v/v) methanol in water with 10% (v/v) acetic acid.

3. ***Stain solution***
   - Dissolve 0.4g of Coomassie blue R350 in 200 mL of 40% (v/v) methanol in water with stirring as needed.

- Filter the solution to remove any insoluble material.
- Add 200mL of 20% (v/v) acetic acid in water.
- The final concentration is 0.1% (w/v) Coomassie blue R350, 20% (v/v) methanol, and 10% (v/v) acetic acid.

4. Destain Solution:

- Add 500mL of HPLC-grade methanol to 300 mL of water.
- Add 100 mL of acetic acid
- After mixing, adjust the total volume to 1000mL with water.
- The final concentrations are 50% (v/v) methanol in water with 10% (v/v) acetic acid.

5. Storage solution:

- Add 25mL of acetic acid to 400mL of water.
- After mixing, adjust the final volume to 500mL with water.
- The final concentration of acetic acid is 5% (v/v).

Procedure

1) After electrophoresis, the apparatus is disassembled and the gel is washed off the glass plates with 500 ml of the gel-fixing solution and soaked in that solution for 1hr. The purpose of this step is to gently remove the gel from the plate and begin washing the SDS-containing gel buffers out of the gel. At the end of this time, remove the solution by aspiration.
2) Cover the gel with 500ml of the gel-washing solution, and continue to fix the proteins in the gel by incubating overnight at room temperature with gentle agitation. The gel should be covered during this process to avoid contamination and to prevent the evaporation of the solution. At the end of this time, remove the solution by aspiration.
3) Cover the gel with 400ml of the Coomassie stain. Stain the gel at room temperature for 3 to 4 hr with gentle agitation. The Coomassie stain is removed by aspiration after staining.

4) Cover the gel with ~250ml of the destain solution and allow the gel to destain with gentle agitation. The destain solution should be changed several times, removing it at each change by aspiration. Continue the destaining until the protein bands are seen without background staining of the gel.
5) Equilibrate the gel in the 500ml of the storage solution for at least 1 hr. The gel should return to its original dimensions during this process.
6) Store the gel in the storage solution as needed. It may be convenient to carefully transfer the gel to a heat-sealable bag for longer-term storage.

Appendix 4

4.1.) Purification of GST tagged Src-SH3/SH2 domains

Materials:

- GST binding buffer (1XPBS pH 7.3)
- GST elution buffer
- dd H2O
- 20% Ethanol
- 10ml Glutathione Sepharose 4B (GS4B) column
- Peristaltic Pump
- TYGON or TEFLON tubing and connectors
- 50 ml sterile collection tubes

GST Binding Buffer (1X PBS) – 1L

- 8.18g NaCl (140mM)
- 0.2 g KCl (2.7 mM)
- 1.419 g Na2HPO4 (10 mM)
- 0.25 g KH2PO4 (1.8 mM)
- Dissolve in 500 – 700 ml autoclaved H2O
- Adjust pH to 7.3
- Adjust volume to 1L

GST Elution Buffer – 100ml (make fresh every time)

- Make 200ml of 50 mM Tris HCl
- Dissolve 0.307 g of reduced glutathione in 80 ml of 50mM Tris (makes 10mM reduced glutathione)
- Adjust pH to 8.0
- Adjust volume to 100 ml

Method

1) Make Binding and Elution buffers sufficient for the purification (refer to the volumes below). The binding buffer can be stored at +4 °C but make the elution buffer fresh (on the same day).
2) Connect the 10 ml column to the pump using the TYGON tubing and connect the outlet valve to a collection tube or a flask. Use a peristaltic pump to fill the tubing with binding buffer. Allow the pump tubing to fill slowly, "drop to drop" to avoid introducing air into the column.
3) Equilibrate the column with 5 column volumes of binding buffer (50ml per 10ml column).
4) Apply the sample by pumping it onto the column. For best results, keep a flow rate of 0.2 ml/min during sample application. Collect the flow through into a 50 ml sterile tube. Take a sample for SDSPAGE.
5) Wash with 5–10 column volumes of binding buffer (50 -100 ml) or until no material appears in the effluent. A flow rate of 1ml/min is recommended for washing. Collect the wash into a flask. Take a sample for SDSPAGE.
6) Elute with 50 ml x2 (2 consecutive elutions, each 50 ml) of elution buffer. A flow rate of 1–2 ml/min is recommended for elution. Collect 2 or 3 (10 ml each) fractions. Take a sample for SDSPAGE.
7) Wash the GST column with 20 column volumes of binding buffer (200 ml). A flow rate of 1–2 ml/min is recommended for wash. Store the column at +4 °C in 20% Ethanol for further use. (This column can be used again just for the first purification of the same 15N labeled protein But do not use this for removal of GST after cleavage. Pack a new 10 ml column for the removal of GST after cleavage, if necessary).
8) Analyze the fractions by SDSPAGE, followed by Coomassie Staining.

4.2.) Thrombin Cleavage of Pure GST-tagged Protein

Materials

- Eluted protein solutions (GST-SH3 and GST-SH3+SH2)
- Thrombin (MW 37KDa) (500-Unit vial)
- PMSF (Phenyl-Methyl-Sulfonyl-Fluoride)
- BME (β-Mercaptoethanol)
- 4X Sample buffer
  (0.5 M Tris pH 6.8; 405 Glycerol; 8% SDS; 0.1% Bromophenol Blue;
  ddH2O)

Method

1) Prepare thrombin solution:
    a. Dissolve 500 Units of thrombin in cold 500 µl PBS (1 U/µl).
    b. Swirl gently to dissolve.
    c. Freeze as 80 µl aliquots and keep at –80 °C.
2) Add the appropriate amount of thrombin (10µl per 10 units of thrombin per each milligram of the tagged protein). Mix gently and incubate at room temperature (+22 to 25 °C) for 16 hours (not more than 16 hours).
3) Stop the cleavage reaction after 16 hrs by adding 1mM PMSF (to a final concentration).
4) Take a 100 ul aliquot for SDS-PAGE analyses. Add 34 ul of final sample buffer, boil for 5 minutes and load on a Tris-Tricine Gel along with a sample of uncleaved protein.

4.3.) Isolation of Pure Protein

Materials

- Cleaved protein solution (SH3 domain and cleaved GST tag)
- G25-Sephadex Desalting Column
- Superdex-75 Size Exclusion Column (SEC)
- 200ml – 500 ml Flasks to collect fractions from G25 column
- 10ml/60 ml Syringes to add protein samples to the columns
- Collection trays and tubes to collect protein samples from SEC
- A Recorder (measures the absorbance and monitor the protein sample)
- Lyophilizing jars
- Liquid N2

Solutions

Size Exclusion Buffer (make 2L)

- 150 mM NaCl
- 50 mM Tris
- adjust pH to 7.5

Desalting Buffer (10 mM NH_4HCO_3) Make 5 L – do not adjust the pH

Method

1) Desalt the cleaved protein solution using the Sephadex G25 column.
    a. First connect the buffer, column, recorder and the collection flasks with tubing
    b. Wash/equilibrate the G25 column with 500 ml – 1L (depending on the size of the column) of fresh desalting buffer.
    c. Let the buffer on the top of the column to absorb to the matrix while making sure not to let the G25 matrix dry out.
    d. Slowly add the protein solution using a syringe and let it be absorbed by the matrix (watch the column every 5 minutes).

e. Wash the tube to collect the last bit of protein and add that to the column slowly and let it absorb again.
f. After adding the protein, let the fresh desalting buffer to run through the column and collect each fraction as monitored by the recorder (check the recorder every 5-10 minutes and change the flask at the beginning of each peak to separate salt and protein fractions)
g. Save 100 µl from each fraction for SDSPAGE and check for the pure protein.

2) Lyophilize the desalted protein.
3) Do Size Exclusion Chromatography to separate the cleaved GST tag from SH3 domain.
   a. Dissolve all the protein from step 2 in Size Exclusion Buffer and check for solubility.
   b. Adjust the pH a bit to make the protein solution completely soluble before adding it to the SEC.
   c. Meanwhile connect the buffer, column, recorder and the collection tubes with tubing and let the size exclusion (SE) column equilibrated wit fresh SEC buffer.
   d. Set up the recorder with the shift of 0.1mm per minute (set it up according to the volume that you are collecting in each tube – e.g. 30 minutes per tube will show as 3 mm in the recorder) and start recording.
   e. Load the protein solution as indicated in the step 1 (desalting). Let it run overnight.
   f. Based on the chart recorded during the run determine the fractions corresponding to each peak and collect 100 ul fractions for SDSPAGE
   g. Check the gel for the fractions that contain only the pure target protein, SH3 domain (Some fractions will have only SH3, some have only GST and some will have both).
   h. Pool the fractions with pure protein (only SH3) and go to the next step.
4) Desalt the pure protein (Same as in step 1).
5) Lyophilize the final protein solution, the pure SH3 domain.

Appendix 5

Calculation of Dissociation Constants (K_D) based on NMR titration data

For a single site-binding model with 1:1 stoichiometry, binding of a small molecule to a protein, is given by equation #1. [P], [L] and [PL] are the concentrations of protein, ligand and the complex at equilibrium.

(#1) $[P] + [L] \Leftrightarrow [PL]$

The rate of formation of complex ([PL]) is k_{on} and the rate of dissociation of complex is k_{off}

(#2) $K_D = [P][L]/[PL] = k_{off}/k_{on}$

The dissociation constant (K_D) cannot be directly determined by a single NMR experiment since the bound molar fractions of protein and ligand at equilibrium cannot be directly measured. However, the total protein concentration, $[P]_0$ and total ligand concentration $[L]_0$ in the solution are known. The $[P]_0$ is the concentration of protein in the NMR tube which is measured before addition of the ligand. The $[L]_0$ is monitored throughout the titration.

(#3) $[P]_0 = [P] + [PL]$ and

(#4) $[L]_0 = [L] + [PL]$

Under the assumption of the concentration of bound ligand equals that of the protein, The K_D from equation #2 can be rewritten using the relationships given in the equations, #3 and #4 ;

(#5) $K_D = \{([P]_0 - [PL])\,([L]_0 - [PL])\}/[PL$

Linking the above relationships with the NMR-observable parameter

The overlay of ^{1}H-^{15}N HSQC NMR spectra of a ^{15}N labeled SH3 domain, obtained by titrating MUC1 peptide, provides an NMR-observable parameter, the

change in chemical shift (Δδ), in response to the amount of peptide that binds to the protein at each consecutive titration point. The NMR-observable parameter depends on the rate of exchange process. Assuming that the MUC1/Src-SH3 interaction is relatively weak (based on the K_D values of a majority of SH3 domain interactions that are reported to date), the nuclei between the free and bound forms of the protein/ligand are assumed to be in fast exchange.

For a system in fast exchange, the observed NMR response to a ligand is the mole fraction weighted average of the NMR parameters of the free and bound states (264).

For fast exchange (NMR timescale), rate of dissociation, k_{off}, is faster than Δδ

k_{off} >> observed chemical shift change, Δδ

For the protein-observed chemical shift changes (Δδ), The following equation gives the relationship among molar fractions of free and bound protein and the observed chemical shift change, Δδ

(#6) $\Delta\delta\ (\delta_{obs}) = f_{P(free)} \cdot \delta_{P(free)} + f_{P(bound)} \cdot \delta_{PL(bound)}$

The molar fractions of free and bound protein add up to 1

(#7) $f_{P(free)} + f_{P(bound)} = 1$

Although equilibrium concentrations of protein and ligand and hence the K_D cannot be directly determined, the known concentrations ($[P]_0$ and $[L]_0$) can be used to relate the equilibrium concentrations of ligand [L] and protein [P] by monitoring the change of an observable NMR parameter (Chemical shift change=Δδ) with varying (and known) concentration of ligand.

The parameter observed (Δ_{obs}) at equilibrium conditions relative to the free (non-bound) state is given by,

(#8) $\Delta_{obs} = \Delta\delta_{obs} - \Delta\delta_{free}$

The change in chemical shift between the fully bound (saturated complex) form

and the non-bound (free protein) form given by;

(#9) $\Delta_{max} = \Delta\delta_{bound} - \Delta\delta_{free}$

The law of mass action predicts the fractional occupancy of protein (fraction of all protein molecules bound to the ligand) at equilibrium as a function of ligand concentration (that is known throughout the titration).

Therefore, from equations, #5, #8 and #9,

(#10) $\Delta_{obs} = [L]_0 \Delta_{max} / K_D + [L]_0$

By fitting a non-linear regression model for 1:1 binding to the experimental data, the Δ_{obs}, Δ_{max}, $[P]_0$ and $[L]_0$ can be used to estimate K_D

(#11) $\Delta_{obs} = \Delta_{max} \{K_D + [P]_0 + [L]_0\} - \{(K_D + [P]_0 + [L]_0)^2 - (4[P]_0 [L]_0)\}^{1/2} / 2[P]_0$

Assigning limits to Δ_{max} and K_D, in equation 11 will generate any number of simulated binding curves to compute the residual sum-of-squares to assess how well the above function fits the observed data. This will provide the means to find the values of Δ_{max} and K_D that gives the smallest possible sums-of-squares based on Δ_{obs}

**The XCRVFIT software* (277) *uses equation #11 to calculate K_D based on the observed chemicals shift values that may follow a 1:1 binding model.*