PART I: STEROID RECEPTOR CO-ACTIVATOR 3 (SRC-3) EXPRESSION IN LUNG CANCER AND ITS ROLE IN REGULATING CANCER CELL SURVIVAL AND PROLIFERATION
PART II: DEVELOPMENT OF PHOSPHOSPECIFIC PEPTOID LIGAND

APPROVED BY SUPERVISORY COMMITTEE

Thomas Kodadek, Ph.D.

John Minna, M.D.

Xiaodong Wang, Ph.D.

Steven Kliewer, Ph.D.

DEDICATED TO MY PARENTS

PART I: STEROID RECEPTOR CO-ACTIVATOR 3 (SRC-3) EXPRESSION IN LUNG CANCER AND ITS ROLE IN REGULATING CANCER CELL SURVIVAL AND PROLIFERATION
PART II: DEVELOPMENT OF PHOSPHOSPECIFIC PEPTOID LIGAND

by

DI CAI

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2009

PART I: STEROID RECEPTOR CO-ACTIVATOR 3 (SRC-3) EXPRESSION IN LUNG CANCER AND ITS ROLE IN REGULATING CANCER CELL SURVIVAL AND PROLIFERATION

Di Cai, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2009

Thomas J. Kodadek, Ph.D., John D. Minna, M.D.

Steroid receptor coactivator-3 (SRC-3) is a histone acetyltransferase and nuclear hormone receptor (NHR) coactivator, located on 20q12, which is amplified in several epithelial cancers and well studied in breast cancer, however, its role in lung tumorigenesis is unknown. We found that SRC-3 is over-expressed in 27% of the NSCLC patients, and SRC-3 high expression correlates with poor disease-free survival and overall survival. We also studied DNA copy number, mRNA and protein expression of SRC-3 in a large panel of lung (55 non-small cell lung cancers and 23 small cell lung cancers) and breast cancers (N=31) and also evaluated the functional consequences of altering its expression in lung cancer cell lines. There are significant alterations in lung

cancers in SRC-3 gene copy number, including examples of both gene amplification and deletion. SRC-3 mRNA and protein expression varied dramatically among lung cancer cell lines. On average, lung cancer cell lines express higher levels of SRC-3 than immortalized human bronchial epithelial cells, which themselves express higher level of SRC-3 than cultured primary human bronchial epithelial cells. We found that ~27% of NSCLCs exhibited SRC-3 gene amplification and expressed SRC-3 mRNA at very high levels, suggesting that the expression of SRC-3 played a role in the malignant phenotype of these cancers. siRNA-mediated down-regulation of SRC-3 in high-expressing tumor cells significantly inhibited tumor cell growth and induced apoptosis. The effect of SRC-3 down-regulation on cell phenotypes correlated with a cell line's endogenous expression level of the gene. Finally, we show that SRC-3 knockdown is "synthetically lethal" to EGFR-TKI-resistant cells. Together these data indicate that SRC-3 is an important new oncogene and therapeutic target for lung cancer.

PART II: DEVELOPMENT OF PHOSPHOSPECIFIC PEPTOID LIGAND

Di Cai, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2009

Thomas J. Kodadek, Ph.D., John D. Minna, M.D.

Most proteins can exist in a variety of post-translationally modified forms. Chemical methods that would allow one to specifically purify or pharmacologically target a particular form of the protein would be of great interest. Here, we report the first peptidomimetic compounds that bind specifically to a serine-phosphorylated PDID domain of Brd4 protein, identified by screening a library of 40,000 peptoids for PDID binders. The isolated hit peptoids are only specific to phos-PDID, but not the non-phosphorylated form of the protein, or other phosphoserine- or phosphothreonine-containing proteins. . Phos-PDID-binding peptoids can specifically capture a recombinant phos-PDID from a crude insect cell extract, without binding to the unmodified PDID in bacteria lysate. Moreover, the phosphospecific peptoid ligand engineered with a Biotin

tag and a DOPA crosslinker can specifically detect phos-PDID from whole cell lysate, demonstrating its potential as a “Western blotting”-like reagent. Furthermore, GST pull-down assay and reporter gene assay reveal that the peptoid ligand can specifically disrupt the interaction between phos-PDID and high-risk HPV 18E2, and hence inhibits the Brd4-dependent transcription activation in human cervical cancer cells. Taken together, these data showed that our phosphospecific peptoid ligand is able to substitute phosphospecific antibodies for the detection and isolation of phosphoproteins; it can also perhaps be developed as a drug-like compound targeting the active form of protein in cells

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Eckstein, N., Servan, K., Girard, L., **Cai, D**., Jonquieres, G., Jaehde U., Kassack, M.U., Gazdar, A.F., Minna, J.D. and Royer, H.D. (2008) "Epidermal growth factor receptor pathway analysis identifies amphiregulin as a key factor for cisplatin resistance of human breast cancer cells" *J. Biol. Chem. 283, 739-750*

Cai, D. (2007) Mini Review: "A semi-synthetic approach to identify kinase substrates" *Molecular BioSystems* 3, 516

Lim, H.-S., **Cai, D**., Archer C.T. and Kodadek, T. (2007) "Periodate-triggered cross-linking reveals Sug2/Rpt4 as the molecular target of a peptoid inhibitor of the 19S proteasome regulatory particle" *J. Amer. Chem. Soc. 129, 12936-12937*

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

NR – nuclear receptor

NHR – nuclear hormone receptor

SRC – steroid receptor co-activator

CBP – CREB-binding protein

ER – estrogen receptor

PR – progesterone receptor

AR – androgen receptor

NFkB – nuclear factor-kB

HAT – histone acetyltransferase

PTM – post-translational modification

HER2/neu – human epidermal growth factor receptor 2

IGF – insulin-like growth factor

MAPK – mitogen-activated protein kinase

EGFR – epidermal growth factor receptor

NSCLC – non-small cell lung cancer

SCLC – small cell lung cancer

VEGF – vascular endothelial growth factor

aCGH – array-based comparative genomic hybridization

HBEC – human bronchial epithelial cells

HBEC-KT – human bronchial epithelial cells immortalized by CDK4 and telomerase

HBEC-UI – un-immortalized

RT-Q-PCR – reverse transcriptase quantitative polymer chain reaction

RPPA – reverse phase protein array

TKI – tyrosine kinase inhibitor

CF – chromatofocusing

IEX – ion exchange

IMAC – immobilized metal ion affinity chromatography

IDA – iminodiacetic acid

NTA – nitrilotriacetic acid

PSSA – phosphorylation state-specific antibodies

POI – the protein of interest

MS/MS – tandem mass spectrometry

PEG – poly(ethylene glycol)

CT – cholera toxin

DOPA – 3,4-dihydroxylphenylalanine

RIP-1 – regulatory particle inhibitor peptoid-1

Brd4 – bromodomain-containing 4

HPV – human papillomavirus

CTM – C-terminal motif

PDID – phosphorylation-dependent interaction domain

GST – glutathione s-transferase

pI – isoelectric point

PB buffer – physiological buffer

CIP – calf intestinal alkaline phosphatase

CK2 – casein kinase 2

NA-HRP – neutraavidin-horse radish peroxidase

PVDF – Polyvinylidene fluoride

BD-DC-pPDID-1 – biotin- and DOPA-conjugated DC-pPDID-1 peptoid

SAR – structure-activity relationship

CHAPTER ONE

INTRODUCTION OF SRC-3 AND LUNG CANCER

Nuclear Receptor and Nuclear Receptor Coactivators

Definitions of Nuclear Receptor and Nuclear Receptor Coactivators

Nuclear receptors (NRs) comprise a superfamily of ligand-regulated and orphan (ligand unidentified) transcription factors that are activated by their steroid hormone ligands and play a central role in diverse biological processes [1]. Gene-specific NR-mediated transcription proceeds through recognizing specific sequences within promoter/enhancer regions of their targeted genes.

Nuclear receptor coactivators are molecules recruited by ligand bound activated NRs (or other DNA binding transcription factors) that enhance gene expression. Coactivators can be classified into two groups: primary coactivators, which contact directly to the NRs, and secondary co-activators, which also contribute to the enhancement of NR-mediated transcription but do not directly contact the NRs [2]. There are now ~200 published NR coactivators that work with ~48 NRs [1].

Physiological Relevance of NR Coactivators

Different NR coactivators contribute to the transcriptional process in very diverse ways. Coactivators may possess multiple enzymatic activities such as acetylation, methylation, ubiquitination, phosphorylation and ATPase activities to regulate gene expression. Coactivators are also predicted to have many activities in addition to

transcription initiation, such as mRNA translation and posttranslational modifications of the synthesized proteins [1]. Thus, coactivators are important for modulating the expression of a wide array of physiologically important genes, which results in particular physiological consequences. For example, knockout mouse study revealed that steroid receptor coactivator SRC-1 and SRC-2/TIF2 play important roles in carbohydrate and lipid metabolism; TIF2-/- mice are protected against obesity and display enhanced adaptive thermogenesis, whereas SRC-1-/- mice are prone to obesity due to reduced energy expenditure [3].

Pathological Relevance of NR Coactivators

Because of their crucial role in regulating NR-mediated metabolic processes, disruptions in coactivator activity can lead to pathological states. Coactivators are involved in many different human diseases. For example, the over- and underexpression of coactivators are largely observed in cancers, both endocrine-related, and many others that do not immediately bring NRs or endocrine relationships to mind, suggesting that coactivators are broadly involved in a large array of cancers [4]. Coactivators are also tightly associated with metabolic syndromes. For examples, defects in the PPARγ coactivator-1 (PGC-1) gene are linked to cholesterol cholelithiasis (gallstones) [5], hypertension [5], and defects in fat metabolism [6]. Germ-line coactivator gene disruptions are the cause of some inherited genetic syndromes, such as mutations in the E6-associated protein (E6-AP) which are responsible for Angelman syndrome [7], and mutations in CBP (CREB-binding protein) or p300 genes that lead to Rubenstein-Taybi syndrome which results in mental retardation [8].

Together, it can be concluded that nuclear receptor coactivators are broadly implicated in human physiological and pathological states. These crucial coregulators will be of growing future interest in clinical medicine [4].

Background of Steroid Receptor Coactivator-3 (SRC-3)

Discovery of SRC-3

Steroid receptor coactivator-3 is a member of the p160 steroid receptor coactivator family. SRC-3, which localizes to a frequently amplified chromosomal region, 20q12, and was initially identified as Amplified in Breast cancer 1 (AIB1) [9]. It is also known as NCoA3, ACTR [10], RAC3 [11], and p/CIP [12].

Molecular Structure of SRC-3

The SRC-3 gene encodes a 160kD coactivator, with 40% sequence identity to other SRC family members, SRC-1 and SRC-2 (TIF2). The molecular structure of SRC-3 includes an N-terminal basic helix-loop-helix-Per/ARNT/Sim (bHLH-PAS) domain which serves as a DNA-binding, protein-protein interaction motif [13], two C-terminal transcriptional activation domains (AD1 and AD2), both of which are responsible for the interactions with CBP and p300 (**Figure 1-1**) [14]. SRC-3 interacts with a broad range of nuclear receptors and transcription factors, including the estrogen receptor (ER), progesterone receptor (PR), E2F1, nuclear factor-kB (NF-kB) and activator protein-1 (AP-1) [15]. The C-terminal domain of SRC-3 possesses a weak histone acetylation (HAT) activity, though the importance of this activity remains unknown [14]. The structure of the SRC-3 protein suggests that it serves as an adaptor protein by recruiting other chromatin modification factors, such as CBP, p300 and PCAF onto the promoter to co-activate transcription [15].

Post-translational Modification of SRC-3

SRC-3's distinct functions are regulated by different posttranslational modifications (PTMs), including phosphorylation, methylation, acetylation, SUMOylation, ubiquitination, and many other modifications [16]. Among them, the phosphorylation of SRC-3 has been most extensively studied. Six Ser/Thr phosphorylation sites have been discovered, which are targeted by different kinases and are important in different signaling pathways. Phosphorylation at distinct sites can affect interactions with nuclear receptors, NF-kB, and CBP selectively, resulting in specific coactivator complex formation, which, in turn, determines differential activation of target genes (**Figure 1-2**) [17]. Moreover, SRC-3 is transformed from an inactive state into a potent transcriptional activator upon phosphorylation, during this process, SRC-3 is also ubiquitinated and ultimately undergoes degradation [2]. Interestingly, the oncogenic potential of SRC-3 was reported to be closely associated with the phosphorylation status of SRC-3 [17].

SRC-3 and Human Cancers

SRC-3 and Hormone-dependent Cancers

SRC-3 was initially discovered in a frequently amplified region, 20q12, in breast and ovarian cancers. In breast cancer biopsies, amplification and over-expression of SRC-3 is detected in 5%-10% and 30%-60% of cases, respectively [9]. SRC-3 over-expression is associated with high levels of HER-2/neu, tamoxifen resistance, and poor disease-free survival, suggesting cross-talk between SRC-3, HER2/neu and ER signaling pathways in the genesis and progression of some breast cancer [18]. Studies in prostate cancer patient samples revealed that SRC-3 expression level was correlated significantly with tumor grade and stage of disease [19], and the overexpression of SRC-3 was correlated with tumor recurrence and poorer survival [20].

SRC-3 and Hormone-independent Cancers

SRC-3 has also been shown to be amplified and over-expressed in many other types of cancer, both hormone-sensitive and hormone-independent ones, including prostate cancer [20], ER, PR-negative breast cancer [21], gastric cancer (GC) and colorectal carcinoma (CC) [15]. SRC-3 over-expression has been correlated with tumor recurrence and metastasis in GC [22] and hepatocellular carcinoma [23]. Increased SRC-3 expression has also been observed during pancreatic cancer [24] and esophageal tumor progression [25]. This clinical evidence implicates SRC-3 as an important player in the progression of many human cancers (**Table 1-1**).

SRC-3 Is an Oncogene in Breast Cancer

In vitro studies give more insight on the role of SRC-3 during tumorigenesis. An MMTV-SRC-3 transgenic mouse model defines SRC-3 as an oncogene, as full-length SRC-3 is sufficient to initiate tumorigenesis [26]. In addition, SRC-3 overexpression is present in many different types of cancers, and is involved in many signaling pathways that control cell proliferation and survival, such as nuclear receptors, HER2/neu, NFkB, and IGF/AKT [15]. These pathways, which will be described below, are frequently deregulated in many tumors, suggesting an important role of SRC-3 in tumorigenesis.

Molecular Mechanisms of SRC-3 Function during Cancer Development

SRC-3 integrates with multiple signal pathways to exert its oncogenic role (**Figure 1-3**). Its involvement in many important signaling cascades controlling cell survival and proliferation is described below.

ER Signal Pathway

Estrogen receptor (ER) mediates the effects of estrogen on the development and progression of breast cancer, and it serves as an important diagnostic marker and as a therapeutic target for both prevention and treatment [18]. ER requires SRC-3's coactivator activity to achieve its complete function [27]. Studies have shown that SRC-3 can enhance the expression of ER target genes, including cyclin D1, through functional interaction of the ER with the cyclin D1 promoter [28]. On the other hand, suppression of SRC-3 leads to a decrease in recruitment of ERα and polymerase II to the promoter region of ERα target gene, resulting in the inhibition of transcription. Moreover, SRC-3 has a higher impact on ER-responsive transcription than other SRC family members [29]. Collectively, SRC-3 is required for the maximal activity of ER and other hormone receptors, and this is one of the major mechanism of how SRC-3 facilitates transformation through ER signaling in breast cancer [15].

Insulin-like Growth Factor-1/AKT Signal Pathway

The insulin-like growth factor-1 (IGF-1)/AKT signaling pathway plays an important role in regulating cell growth, proliferation, survival and migration. Both *in vitro* and *in vivo* studies have shown that the expression of SRC-3 is closely and

positively linked to the IGF-1 expression levels [26, 30]. Moreover, the cognate receptor of IGF-1, IGF1Rα, is induced and highly phosphorylated in SRC-3 transgenic mice, whereas IGF1Rα protein level is reduced in breast cancer cells when SRC-3 is down-regulated [26, 31]. Taken together, these data indicate that the expression level and/or activity of many players in the IGF/AKT pathway are under the tight control of SRC-3.

NF-κB Signal Pathway

Rel/Nuclear factor-κB (NF-κB) is an important transcription factor that controls cell growth, differentiation, and apoptosis. Aberrant NF-κB activation has been implicated in the progression of many human cancers, including breast, prostate, liver and pancreas cancer [32]. SRC-3 can interact and co-activate NFκB. Upon stimulation by tumor necrosis factor (TNF)-α, SRC-3 is phosphorylated by the IκB kinase (IKK), translocates from the cytosol to the nucleus, and induces the expression of it target gene, interleukin (IL)-6, which has an NF-κB binding site in its promoter and has important functions in tumor metastasis and inflammation [33]. This result suggests that SRC-3, especially phospho-SRC-3, is closely associated with NF-κB transcriptional activity.

HER2/neu Signal Pathway

HER-2 receptor is a member of the epidermal growth factor (EGF) receptor family that activates MAPKs. It is also an important oncogene for breast cancer. Several clinical studies on breast cancer patients showed that the expression of SRC-3 and HER2/*neu* are highly correlated with the development of tamoxifen resistance, suggesting that the crosstalk exists between HER2/*neu* and SRC-3 [18]. A mouse model

study has revealed that homozygous deletion of SRC-3 can completely prevent HER2/*neu*-induced tumor formation, indicating that SRC-3 is required for HER2/*neu*'s oncogenic activation, signaling, and mammary tumorigenesis [34].

EGFR Signal Pathway

Amplification and mutation of epidermal growth factor receptor (EGFR) occurs in ~10-20% in non-small cell lung cancer and less frequently in some other cancers [35]. The EGFR signaling pathways can regulate cell proliferation, apoptosis, angiogenesis and tumor invasion [35].Many studies revealed a relationship between SRC-3 expression level and EGFR signaling. Over-expression of SRC-3 in breast cancer cells correlates with increased levels of EGFR and HER2 protein and resistance to tamoxifen therapy [18, 36]. In addition, reduction of SRC-3 expression in lung, breast and pancreatic cancer cells led to decreased EGF-induced proliferation, due to a reduction in EGFR tyrosine phosphorylation level at multiple sites, with overall tyrosine phosphorylation unaffected [37].

Lung Cancer

Lung cancer is the leading cause of cancer death in both men and women in the United States and throughout the world [38]. The two major classes of lung cancer are non-small cell lung cancer (NSCLC, ~85% of all lung cancer) and small-cell lung cancer (SCLC, ~15%). Non-small cell lung carcinoma (NSCLC) includes squamous cell carcinoma, adenocarcinoma and large cell carcinoma.

Important Risk Factors in Lung Cancer

Smoking causes all types of lung cancer but is most strongly linked with SCLC and squamous cell carcinoma; adenocarcimona most commonly occurs in patients who have never smoked [35]. Important oncogenes and tumor suppressors in lung cancer include p53, KRAS, epidermal growth factor receptor (EGFR), c-myc, PIK3CA, etc [35].

Lung Cancer and EGFR

One of the most important oncogenes for NSCLC is EGFR either through amplification and/or mutation. The EGFR is a member of a family of four receptors: EGFR (HER1 or ErbB1), ErbB2 (HER2/neu), ErbB3 (HER3) and ErbB4 (HER4). Upon ligand binding to receptors, homodimers or heterodimers are formed between family members, following by the activation of the intrinsic kinase domain, and phosphorylation of specific tyrosine residues in the cytoplasmic tail of the receptor, which in turn triggers multiple downstream signaling cascades, including the Ras-ERK and PI3K-Akt pathways [39]. These EGFR signaling pathways can regulate cell proliferation, apoptosis, angiogenesis and tumor invasion [35].

EGFR Inhibitors and Drug Resistance

Several anti-EGFR agents, including the monoclonal antibody Cetuximab targeting the ligand-binding domain of the receptor, and small molecules Gefitinib (Iressa) and Erlotinib (Tarceva) targeting the tyrosine kinase domain of EGFR, are currently used in the clinic to treat NSCLC patients. Despite the efficacy of these EGFR inhibitors, the majority of patients receiving these therapies show either de novo resistance or a relapse after initial response(acquired resistance) [40]. Known resistance mechanisms include the common second site mutation (T790M) in EGFR, Kras mutation and MET amplification [35]. Other possible resistance mechanisms include activation of other receptor tyrosine kinases, such as insulin-like growth factor 1 (IGF-1) receptor, which can bypass EGFR to activate critical downstream signaling pathways[41]. Targeting these alternative pathways could be a potential way to achieve synergistic effects with EGFR inhibitors.

Lung Cancer, Nuclear Receptors and SRC-3

Estrogen receptors have been studied for their potential as therapeutic targets to treat lung cancer in combination with EGFR inhibitors [42]. Recently, several reports have suggested a function of nuclear hormone receptors and their coactivators in lung cancer cell survival and proliferation. For example, the estrogen-signaling pathways play an important role in controlling lung cancer growth [43]. This signaling pathway in NSCLC cells may include SRC-3, as it is expressed in several lung cell lines, and both EGF and estradiol elicit serine phosphorylation of SRC-3 *in vitro* [44]. Other researchers have shown that SRC-3 can regulate EGF-induced proliferation and gene expression, through modulation of EGFR phosphorylation at multiple tyrosine residues in the

NSCLC cell line A549 [37]. Furthermore, the candidate lung tumor suppressor gene RBM5/LUCA-15 (locus 3p21.3) decreased the expression of SRC-3 as monitored by cDNA microarray [45].

However, a comprehensive profiling of SRC-3 gene and protein expression levels in primary lung tumors and NSCLC cell lines is lacking in the field. Also, the exact function of SRC-3 in lung cancer cell survival and how SRC-3 expression correlates with gene expression and drug response are still poorly understood.

The Concept of "Oncogene Addiction"

Cancer is generally regarded as a multistep genetic disorder. Though the evolution of the disease is a very complicated process, the growth and proliferation of cancer cells can be impaired by the inactivation of a single oncogene. This concept, called "oncogene addiction", reflects the apparent dependency of some cancers on a single or several important genes to sustain the malignant phenotype[46]. Numerous studies in human cancer cell lines and clinical trials have identified several important oncogenes for different human cancers, including Her-2, Cyclin D1, K-ras^{mut}, EGFR, VEGF and VEGFR. The mechanism of oncogene addiction is not merely an accumulation of the effects of oncogene activation and tumor-suppressor gene inactivation. It is more of the fact that cancer cells rely far more on a specific oncogene than normal cells because of the loss of flexibility due to the inactivation of several other genes [47, 48].

Targeted Therapy (Personalized Medicine)

Identification of the specific state of oncogene addiction in specific types of cancers can provide a justification for molecular targeted therapy. Currently, approaches such as small interfering RNAs (siRNAs) can be used to identify the genes that are important for the proliferation of a particular type of cancer. Specific drugs can be developed targeting these molecules, and presumably the tumors would respond better to treatment [49]. For example, gefitinib and erlotinib, which target the tyrosine kinase domain of EGFR, are used in the clinic to treat NSCLC patients with EGFR mutations.

Because of the heterogeneity and genetic instability of tumors, the use of a single molecule targeted therapy is unlikely to achieve long-lasting effects in tumor regression. Therefore, combination therapy will be needed for more effective treatment, such as the combination use of molecularly targeted drugs with cytotoxic agents, or with a targeted agent in a different pathway to achieve the full potential [46, 50].

In summary, both systematic mechanistic studies of the oncogene addiction and development of various molecular targeted drugs are required to better develop effective treatments for cancer.

Table 1-1 SRC-3 in hormone-independent human cancers.

Type of Change	Types of Cancer
Overexpression	Gastric Cancer (GC)
	Colorectal Carcinoma (CC)
	Hepatocellular Carcinoma (HCC)
	ER, PR-negative Breast Cancer
Increased Gene Copy Number	Pancreatic Adenocarcinomas (37%)
	Esophageal Tumor (ET) (13%)
Tumor Recurrence & Metastasis	Gastric Cancer (GC)
	Colorectal Carcinoma (CC)
	Hepatocellular Carcinoma (HCC)
Increased Expression during Tumor Progression	Pancreatic Cancer
	Esophageal Tumor (ET)

Figure 1-1 Molecular Structure of SRC-3.
Modified from Acta Pharmacologica Sinica, 2006

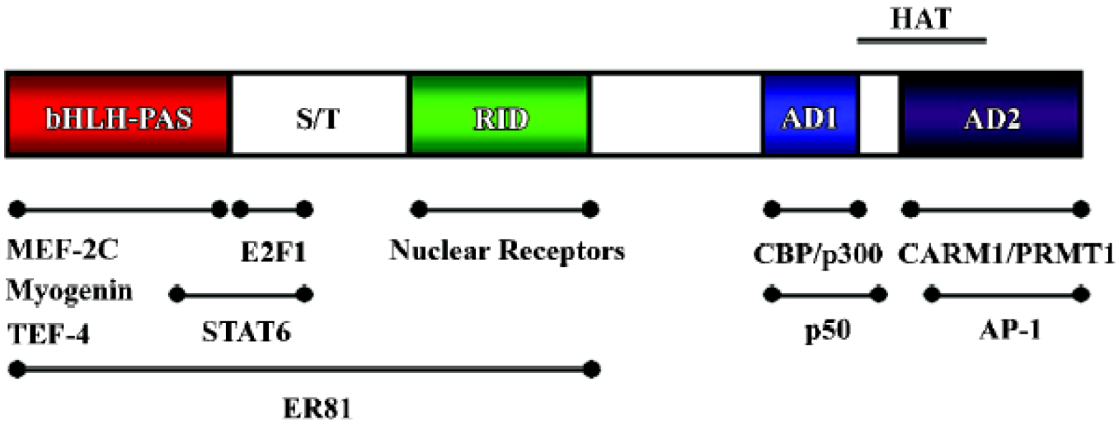


Figure 1-2 Post-translational modification on SRC-3 regulates its signalings. Modified from Molecular Cell 2007.

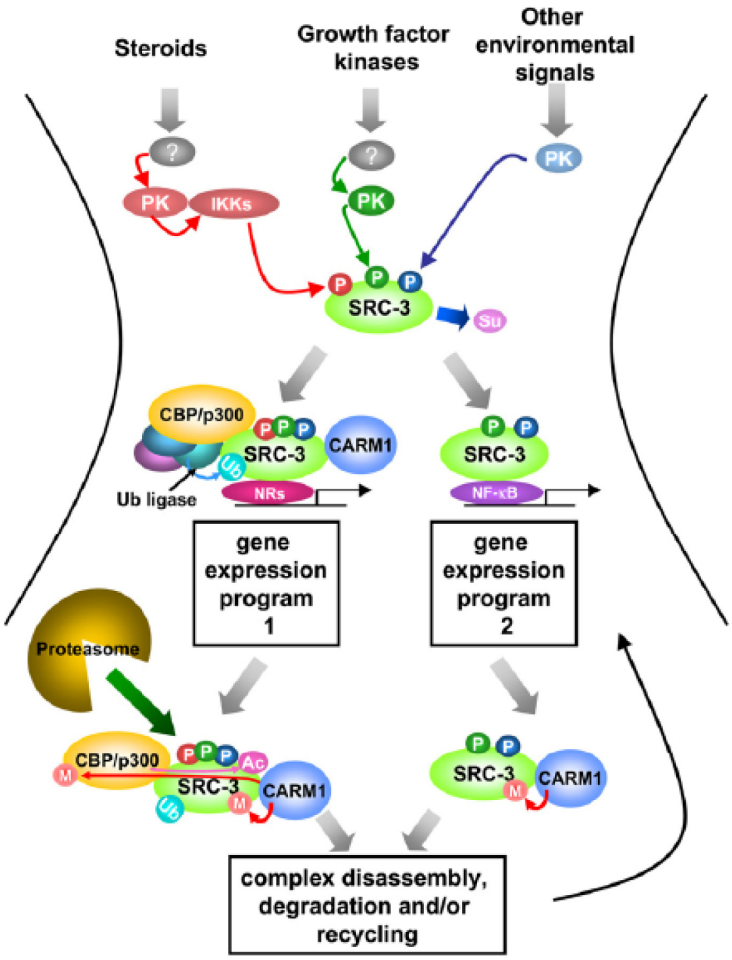
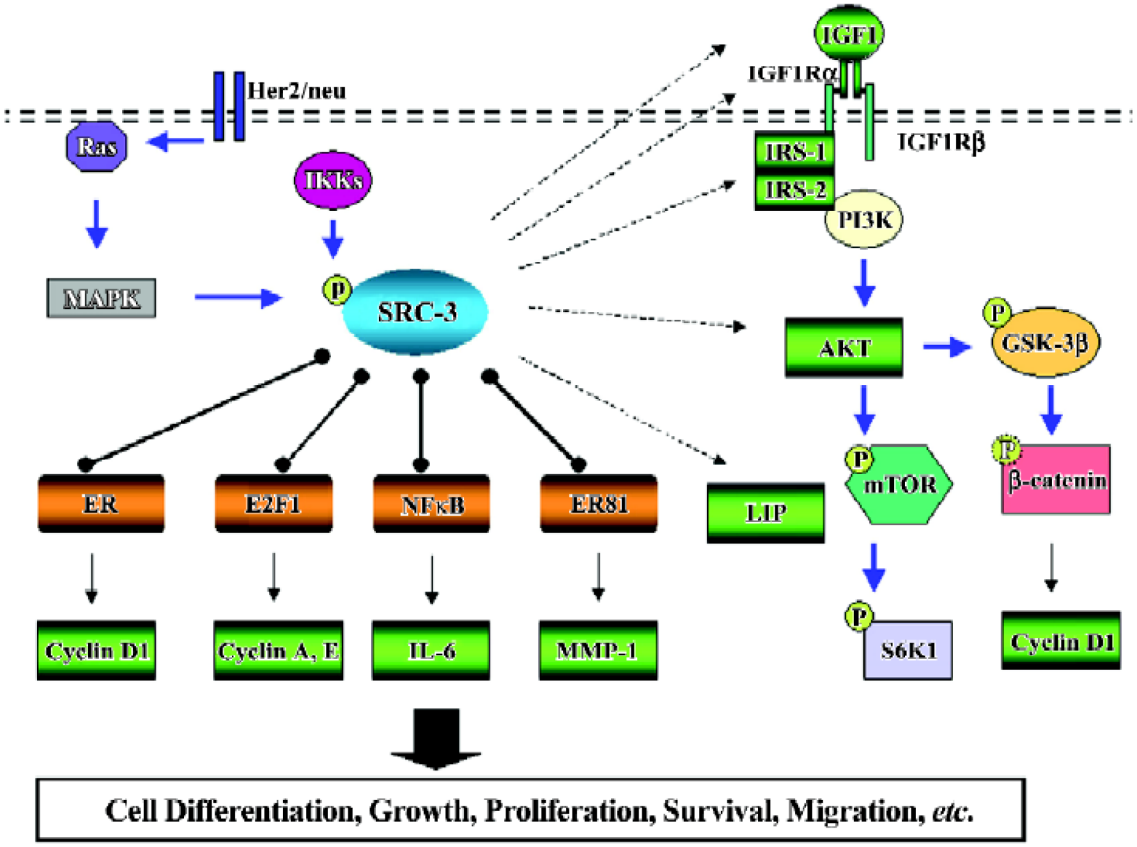


Figure 1-3 SRC-3 integrates with multiple signal pathways to exert its oncogenic role. Modified from Acta Pharmacologica Sinica, 2006.



CHAPTER TWO

STEROID RECEPTOR CO-ACTIVATOR 3 (SRC-3) EXPRESSION IN LUNG CANCER AND ITS ROLE IN REGULATING CANCER CELL SURVIVAL AND PROLIFERATION

Abstract

Steroid receptor coactivator-3 (SRC-3) is a histone acetyltransferase and nuclear hormone receptor (NHR) coactivator, located on 20q12, which is amplified in several epithelial cancers and well studied in breast cancer, however, its role in lung tumorigenesis is unknown. We found that SRC-3 is over-expressed in 27% of the NSCLC patients, and SRC-3 high expression correlates with poor disease-free survival and overall survival. We also studied DNA copy number, mRNA and protein expression of SRC-3 in a large panel of lung (55 non-small cell lung cancers and 23 small cell lung cancers) and breast cancers (N=31) and also evaluated the functional consequences of altering its expression in lung cancer cell lines. There are significant alterations in lung cancers in SRC-3 gene copy number, including examples of both gene amplification and deletion. SRC-3 mRNA and protein expression varied dramatically among lung cancer cell lines. On average, lung cancer cell lines express higher levels of SRC-3 than immortalized human bronchial epithelial cells, which themselves express higher level of SRC-3 than cultured primary human bronchial epithelial cells. We found that ~27% of NSCLCs exhibited SRC-3 gene amplification and expressed SRC-3 mRNA at very high levels, suggesting that the expression of SRC-3 played a role in the malignant phenotype of these cancers. siRNA-mediated down-regulation of SRC-3 in high-expressing tumor cells significantly inhibited tumor cell growth and induced apoptosis. The effect of SRC-

3 down-regulation on cell phenotypes correlated with a cell line's endogenous expression level of the gene. Finally, we show that SRC-3 knockdown is "synthetically lethal" to EGFR-TKI-resistant cells. Together these data indicate that SRC-3 is an important new oncogene and therapeutic target for lung cancer.

Introduction

Lung cancer is the leading cause of cancer death in both men and women in the United States and throughout the world [38]. The two major classes of lung cancer are non-small cell lung cancer (NSCLC, ~85% of all lung cancer) and small-cell lung cancer (SCLC, ~15%). Non-small cell lung carcinoma (NSCLC) includes squamous cell carcinoma, adenocarcinoma and large cell carcinoma. Important prognostic and predictive factors for lung cancer include p53 mutation, KRAS mutations, epidermal growth factor receptor (EGFR) amplification and mutation and other genes [35]. Recently, several reports suggest nuclear hormone receptors and their coactivators promote lung cancer cell survival and proliferation [37, 43, 44].

Steroid receptor coactivator-3 (SRC-3; AIB1/ACTR/RAC3/p/CIP) is a member of the p160 SRC family. SRC-3 has histone acetyltransferase activity and interacts with multiple nuclear receptors and transcription factors, including estrogen receptor (ER), progesterone receptor (PR), E2F1, nuclear factor-kB (NF-kB) and activator protein-1 (AP-1) [15]. Many studies revealed a relationship between SRC-3 expression level and EGFR signaling. Over-expression of SRC-3 in breast cancer cells correlates with increased levels of EGFR and HER2 protein and resistance to tamoxifen therapy [18, 36]. In addition, reduction of SRC-3 expression in lung, breast and pancreatic cancer cells led to decreased EGF-induced proliferation, due to a reduction in EGFR tyrosine phosphorylation at multiple sites, with overall tyrosine phosphorylation of cellular protein unaffected [37].

SRC-3 has been implicated in the development of many human cancers: Amplification and over-expression of SRC-3 is detected in 5%-10% of ovarian cancers

and 30%-60% of breast cancer biopsies [9]. SRC-3 over-expression is associated with high levels of HER-2/neu, tamoxifen resistance, and poor disease-free survival, suggesting that there may be cross-talk between SRC-3, HER2/neu and ER signaling pathways in the genesis and progression of some breast tumors [18]. SRC-3 is also amplified and over-expressed in many other types of cancer, including hormone-sensitive and hormone-independent cancers such as prostate cancer [20], ER, PR-negative breast cancer [21], gastric cancer (GC) and colorectal carcinoma (CC) [15]. Increased SRC-3 expression is also observed during pancreatic cancer [24] and esophageal tumor progression [25]. However, a comprehensive profiling of SRC-3 gene and protein expression level in primary lung tumor and cultured cell lines is lacking. Also, the exact function of SRC-3 in lung cancer cell survival and proliferation is still largely unknown.

One of the most important prognostic and predictive factors for NSCLC is EGFR amplification and mutation. Several anti-EGFR agents, including the monoclonal antibody Cetuximab targeting the ligand-binding domain of the receptor, and small molecule EGFR tyrosine kinase inhibitors (EGFR-TKIs) Gefitinib (Iressa) and Erlotinib (Tarceva) are currently used in the clinic to treat NSCLC patients. Despite the efficacy of these EGFR inhibitors, the majority of patients receiving these therapies show either de novo resistance or a relapse after initial response (acquired resistance) [40]. Known resistance mechanisms include *MET* amplification and the common second site mutation (T790M) in EGFR [35]. Other possible resistance mechanisms include activation of other receptor tyrosine kinases, such as insulin-like growth factor 1 (IGF-1) receptor, which can bypass EGFR to activate critical downstream signaling pathways [41]. SRC-3 is a crucial regulator in these alternative pathways, as it has been shown to mediate IGF-1-

induced phenotypic changes in human breast cancer cells [31] and SRC-3 deficiency affects breast cancer initiation and progression in mice [51]. However, the correlation between SRC-3 expression and the response to EGFR inhibitors remains to be elucidated.

In this study, we show that SRC-3 is overexpressed in a portion of lung cancers directly *ex vivo,* and overexpression of SRC-3 correlates well with poor disease free and overall survival. DNA copy number, mRNA- and protein-level expression of SRC-3 are also profiled in a large panel of NSCLC, SCLC and breast cancer cell lines. Knockdown of SRC-3 in lung cancer cells leads to reduced cell growth, decreased anchorage-independent liquid colony formation ability and increased apoptosis in cell lines with high endogenous levels of SRC-3. In addition, we show that SRC-3 knockdown is "synthetically lethal" to EGFR-TKI-resistant cells. Together these data indicate that SRC-3 is an important new oncogene and therapeutic target for lung cancer.

SRC-3 Expression Is Highly Variable in Lung Cancer Patients

To evaluate the expression of SRC-3 in lung cancer, we performed immuno-histochemistry on a tissue microarray of 330 clinically localized lung cancers using anti-SRC-3 antibody. Stained slides were digitized and expression was quantified using a four-value intensity score (0, 1+, 2+, and 3+) and the percentage (0% to 100%) of reactivity. We defined the intensity categories as follows: 0 = no appreciable staining; 1+ = barely detectable staining in epithelial cells compared with the stromal cells; 2+ = readily appreciable staining; and, 3+ = dark brown staining of cells. Next, an expression score was obtained by multiplying the intensity and reactivity extension values (range, 0–300).

Representative pictures of high and low expression in NSLCLC patients with adenocarcinoma or squamous cell carcinoma are shown respectively (**Figure 2-1A**, i & ii for adenocarcinoma; iii & iv for squamous cell carcinoma). SRC-3 shows highly variable expression among the lung tumors tested: There were 85 samples (27%) that express high level of SRC-3 with staining score ≥10, including 11 samples that express extremely high levels of SRC-3 (score>40). There were 84 samples (27%) with a SRC-3 nuclear staining score between 0 and 10. There were also 144 samples (46%) that were completely negative (score=0) for SRC-3 staining (**Figure 2-1B**). In summary, SRC-3 is very variably expressed among lung cancer patients; with 25% of the samples over-express this protein.

Increased SRC-3 Expression Confers a Worse Prognosis in Lung Cancer Patients

Next, we investigated whether SRC-3 expression is associated with prognosis in lung cancer. The clinical outcomes measured were overall survival (n=215) and progression-free survival (n=210). Progression was defined as either disease recurrence or death; survival time was defined as the time from diagnosis to death (or disease progression), which was censored by the last known number of follow up days. Lung cancer patients were dichotomized into high SRC-3 expression group (>1.67) and low SRC-3 expression group (≤1.67), and the cut-off value is the median SRC-3 expression among all lung cancer samples.

In the high SRC-3 expression group, there were 66 out of 108 patients dead; in the low SRC-3 group, there were 39 out of 107 patients dead within the follow up months. Kaplan-Meier overall survival curves (**Figure 2-1 C,** right panel) shows that patients with high SRC-3 levels have significantly shorter survival time than patients with low SRC-3 levels (p=0.0008, log-rank test); the median survival time was 54.4 months for the high SRC-3 group and 98.0 months for the low SRC-3 group. **Figure 2-1 C** (left panel) shows that SRC-3 was significantly associated with progression free survival time (p=0.0015, log-rank test). **Table 2-1** shows that after adjusting for the effects of histology, gender, race, tobacco and stage in the multivariate proportional hazard survival model, high SRC-3 was still significantly associated with poor survival time (hazard ratio = 2.02, and p=0.0007) and progression free survival time (hazard ratio=1.92, and p=0.0011).

SRC-3 Gene Copy Number Alterations, mRNA- and Protein-levels in a Panel of Non-small Cell Lung Cancer (NSCLC) Cell Lines

After investigating the expression of SRC-3 in primary tissues from lung cancer patients, we extended our study to cultured cell lines. To analyze the DNA copy number alterations of *SRC-3* in tumor cells, array-based comparative genomic hybridization (aCGH) was carried out in a panel of lung and breast cancer cell lines. On this cDNA microarray platform, there were 5 probes targeting different regions of *SRC-3* gene. A gain at 20q12 *SRC-3* locus was present in 14 of 55 (25%) NSCLC lines, 8 of 23 (35%) SCLC lines and 15 of 31 (48%) breast cancer cell lines are amplified at the 20q12 *SRC-3* locus, while 4 of 55 (7%) NSCLC lines, 1 of 23 (4%) of SCLC lines and 1 of 31 (3%) breast cancer cell lines are deleted at this locus (**Table 2-2).**

To determine whether SRC-3 is over-expressed in lung cancer cell lines, we performed quantitative- RT-PCR using 43 NSCLC, 22 SCLC, 16 breast cancer and 11 human bronchial epithelial cell lines immortalized by CDK4 and telomerase (HBEC-KT) and 6 primary, un-immortalized human bronchial epithelial cells (HBEC-UI). Over 30% of NSCLCs (N=43) and SCLCs (N=22) over-express *SRC-3* at a level least 3-4 fold higher than the average for normal HBECs (N=17) (**Figure 2-2 A, Figure 2-6 A & C**). The extent of over-expression in these lung cancer cell lines was comparable to a high-expressing breast cancer cell line (**Figure 2-6 B & D**). Overall, there was a greater than 20 fold variation in the mRNA level of *SRC-3* expression within NSCLCs. When comparing *SRC-3* mRNA levels between different disease groups, the average level of NSCLC group was greater than that of HBEC-KT which was in turn greater than that of HBEC-UI (**Figure 2-2 A**). This suggests a general trend towards increasing *SRC-3*

expression during tumor progression, i.e., from primary normal epithelial cells, to immortalized cells, to malignant tumor, compatible with a pro-oncogenic role of SRC-3 in lung cancer development. In rare cases, some NSCLC cells, such as HCC44, HCC15 and H28, expressed *SRC-3* mRNA at a level that is even lower than the normal epithelial cells: These cells are identified as having deletions at the *SRC-3* gene locus 20q.12 based on the aCGH results.

Western blot analysis revealed an even more dramatic difference of SRC-3 expression in the cell lines at the protein level (**Figure 2-2 B**). Immortalized HBEC-3KT and 13KT cells show either no detectable SRC-3 or low level of the protein. NSCLC cells H1299 and A549 contain very high amounts of SRC-3 protein, while H1819 and H460 show medium levels, whereas H2073 expresses very little, if any detectable SRC-3 protein. SRC-3 protein levels did not always correlate with DNA amplification or mRNA levels, suggesting post-translational modification may be important in NSCLC.

SRC-3 is a phospho-protein that can be regulated by different protein kinases, such as MAPK and IkB. A total of 8 phosphorylation sites have been identified in the SRC-3 protein, and the phosphorylation of these sites mediates different transcription activation and oncogenic transformation [17, 52]. For example, almost all phosphorylation sites are important for interaction with ER and AR, but only Thr24 and Ser867 are required for NF-kB binding. Phosphorylation at Thr24 is also important for the transforming ability of SRC-3 [17]. To determine the level of the phosphorylated form of SRC-3 in lung cancer cells, western blot analysis was performed using a phospho-SRC-3-specific antibody recognizing the Thr24 phosphorylation site (Cell Signaling). The band intensity on the western blot was quantified, and the ratios of phospho-SRC-3 (Thr24)/total SRC-3 were

normalized to the value for H1819, which was defined as 1. As shown in **Figure 2-2 B**, the phosphorylation level of SRC-3 is highly variable among the lung cancer cell lines. Some lines, such as HBEC-13KT and H2073, show no detectable phospho-SRC-3 (Thr24) level though it expresses modest amounts of total protein. On the other hand, H1299 has the highest phospho-SRC-3/total SRC-3 ratio, though its does not express the highest total SRC-3, suggesting SRC-3 pathway is highly active in this cell line.

Down-regulating of SRC-3 Expression in High-expressing NSCLC Cells Significantly Decreases Cell Growth and Induces Apoptosis

To investigate whether the cells expressing high-levels of SRC-3 rely on it for survival and proliferation, a pool of 4 siRNA oligos specifically targeting SRC-3 (Dharmacon) was transfected into H1299, which expresses high-levels of both total and phosphorylated (Thr24) SRC-3. As shown in **Figure 2-3 C,** SRC-3 protein was reduced after 4 days of siRNA treatment, compared to untreated or control siRNA-treated cells. 72 hours and 96 hours after siRNA transfection, the cells were counted by hemocytometer with trypan blue staining. More than a 70% decrease in cell number was observed in SRC-3 down-regulated cells compared to control treatment (**Figure 2-3 B**). This result shows that SRC-3 is important for proliferation or survival of H1299.

To study the role of SRC-3 in *in vitro*clonogenicity, we determined the effect of SRC-3 knockdown on liquid colony formation (**Figure 2-3 D)**. Down-regulation of SRC-3 inhibited liquid colony formation by more than 70% in H1299. **Figure 2-3 E** shows the reduction in colony formation that took place under all 3 different cell densities tested, confirming the importance of SRC-3 even at a very low density of cells.

The decrease in cell growth could be a consequence of either reduced proliferation, increased cell death, or both. Microscopic examination of SRC-3 siRNA treated H1299 cells showed a reduction in cell number and an altered morphology suggestive of cells undergoing apoptosis (**Figure 2-3 A**). To determine whether SRC-3 is required for lung cancer survival, we performed Western blots for one of the executioner caspases, cleaved-caspase 7. We detected cleaved-caspase 7 expression 72 hours after siRNA transfection. Its level was increased and sustained up to 120 hours after siRNA

transfection (**Figure 2-3 B**). These data indicate that the growth reduction effect in SRC-3 knock down H1299 cells is, at least in part due to the activation of the apoptosis pathway.

The Endogenous Expression Level of SRC-3 Correlates with Pro-survival Phenotype

To determine whether the level of SRC-3 expression correlates with the sensitivity of the cells to knockdown of SRC-3, we extended our studies to H1819 and H2073 cells, which have medium and low SRC-3 expression, respectively. As expected, with efficient SRC-3 down-regulation in H1819 , a modest inhibition of cell growth (~30%) and liquid colony formation ability (~30%) was observed consistently compared to the untreated or nonspecific siRNA-treated control groups (**Figure 2-4 A**).

On the other hand, specific down-regulation of SRC-3 in low-expressing H2073 cells did not show a significant effect on either cell proliferation or colony formation ability (**Figure 2-4 B**). These data again suggest that the extent of SRC-3 dependence is reflected by the endogenous level of the protein in the cells.

SRC-3 Knockdown Is "Synthetically Lethal" to EGFR TKI-resistant Cells

Many studies have revealed a relationship between SRC-3 expression level and EGFR signaling [18, 36]. A recent paper also discovered that SRC-3 regulates EGF-induced proliferation through modulating the phosphorylation of EGFR [37]. Therefore, it is possible that SRC-3 expression correlates with the response to the EGFR-targeted chemotherapy. To test this hypothesis, we first examined the expression of 60 oncogenesis- related proteins in 50 NSCLC cell lines by reverse phase protein arrays (RPPAs) using validated antibodies, and correlated protein levels with response to 15 chemotherapy drugs using MTS proliferation assay [53]. Cluster analysis between the two sets of data provided insights into the relationship between drug sensitivity and certain protein expression. Not surprisingly, high EGFR protein expression was correlated with Cetuximab, Erlotinib and Gefitinib sensitivity. Interestingly, high SRC-3 protein expression correlated with the resistance to EGFR-TKIs Erlotinib and Gefitinib (**Figure 2-5 A**). This result led us wonder if SRC-3 knockdown might be "synthetically lethal" with EGFR inhibitors in EGFR TKI-resistant cell lines.

As shown in **Figure 2-2 C** and **Figure 2-4 A**, H1819, a NSCLC cell line which harbors wildtype EGFR, expresses moderate levels of SRC-3, and knocking down SRC-3 alone led to only a modest reduction in cell growth. Phosphorylation of EGFR, HER2, Erb-B3 and downstream effectors such as AKT and P44/42 was detected in H1819 which expresses wildtype EGFR, indicating that the EGFR pathway is active in this cell line (**Figure 2-7**). However, it is resistant to EGFR TKIs, with an IC50= 15uM for Iressa (Gefitnib) [53]. We decided to use this cell line to test whether SRC-3 knockdown is synthetically lethal with a TKI to the resistant cells.

SRC-3 was knocked down in H1819 by specific siRNA, followed by Iressa treatment at 0.1uM, a concentration that is within the "sensitive" definition for the drug response, but is ~100 fold lower than the IC50 for this line. Seventy-two hours after siRNA transfection, SRC-3 protein level was efficiently down-regulated (**Figure 2-5 C**). Cells treated with combination therapy show a "stressed" phenotype under the microscope compared to SRC-3-specific siRNA or Gefitnib treatment alone (**Figure 2-5 B**). Annexin V staining also showed that more cells undergoing early apoptosis in the SRC-3 knockdown and Iressa combination treatment group than in control groups (**Figure 2-5 C**). Therefore, a reduction of SRC-3 expression leads to a dramatic sensitization of these cells to EGFR TKIs.

Taken together, the data reported herein indicate that SRC-3 is an important new oncogene and therapeutic target for lung cancer.

Discussion

Herein, we have provided the first example of a comprehensive investigation of the expression of SRC-3 in lung cancer. By analyzing 330 clinical lung cancer patients by immunohistochemistry, we demonstrated that lung cancer patients express various level of nuclear SRC-3. ~27% of the patients express very high amounts of the protein (**Figure 2-1B**). Clinically, patients with high SRC-3 expression had a worsen prognosis and low overall survival (**Figure 2-1C**). In a parallel cell culture study, DNA copy number alteration, mRNA- and protein-level expression of SRC-3 were determined in 82 lung cancer cell lines, and over-expression of SRC-3 was observed in ~20% of these samples(**Table 2-2** and **Figure 2-2**). Both the clinical observations and *in vitro* cell line studies show a similar frequency of high expression (~20%). This result is comparable to the frequency of SRC-3 over-expression in breast cancer, ovarian cancer and prostate cancer, all of which require SRC-3 for the proliferation of cancer cells [9, 20, 21]. The percentage of SRC-3 amplification and over-expression is also similar to what is observed for other lung cancer oncogenes, such as EGFR (15%) and MET (20%) amplification in adenocarcinoma [35]. These results provide evidence implicating that SRC-3 is an oncogene in lung cancer.

We determined SRC-3 gene copy number, mRNA- and protein-level expressions in lung cancer cell lines, and noticed that cell lines with SRC-3 gene amplification tend to have the highest level of mRNA, but some cells without amplification also have mRNA overexpression (**Figure 2-8**). Furthermore, SRC-3 mRNA levels do not invariably correlate with protein levels, and total SRC-3 protein does not always represent the level of phospho-SRC-3. This inconsistency in SRC-3 DNA, mRNA and protein suggests that

there is regulation of SRC-3 at the transcriptional, translational and post-translational levels. For example, a recent study identified a ubiquitin ligase, CHIP, can directly target SRC-3 for ubiquitinylation and degradation, and inhibits anchorage-independent cell growth and metastatic potential of cancer cells [54].

We have assessed the role of SRC-3 in lung cancer cell growth and proliferation by down-regulating its expression in three cell lines expressing various level of the protein. Knock-down of SRC-3 inhibits cell growth and colony formation, which suggest that depletion of the gene in vivo could prevent or reduce tumor growth. Notably, the siRNA effects correlate with endogenous SRC-3 expression levels. High SRC-3-expressing H1299 cells were most sensitive to siRNA treatment, the medium-expressor H1819 showing modest growth inhibition, while low-expressing cells gave almost no response. This correlation between endogenous expression level of SRC-3 and the pro-survival phenotype fit in very well with the "oncogene addiction" theory, again supporting SRC-3 as an important player for lung cancer progression.

Combination therapy by both down-regulating of SRC-3 via siRNA and a low dose (0.1uM) of EGFR-TKI Iressa can drive the TKI-resistant cell H1819, which harbors wildtype EGFR into apoptosis. As mentioned above, the other majority of the patients are resistant to EGFR TKIs, with a response rate of about 10-20%[55]. Our result of SRC-3 knock down is "synthetical lethal" for EGFR-TKIs suggests that SRC-3 could be a new target for lung cancer treatment. Drugs targeting SRC-3 may synergize with EGFR-TKIs in the subset of cases that express SRC-3 and wildtype EGFR without amplification and do not respond to EGFR-TKIs alone.

We have also tested whether SRC-3 knockdown sensitizes the cells to Cetuximab, a monoclonal antibody targeting the ligand-binding domain of EGFR. However, no synergy effect was observed (data not shown). This is not surprising, as the clinical response to Cetuximab is highly dependent on EGFR copy number, and only ~15% of the adenocarcinoma patients have EGFR amplifications [35, 56].

Taken together, our results indicate that SRC-3 is an important new oncogene and therapeutic target for lung cancer.

Materials and Methods

Case selection and TMA construction. We obtained archived, formalin-fixed, paraffin-embedded (FFPE) tissue from surgically resected (with curative intent) lung cancer specimens (lobectomies and pneumonectomies) containing tumor and adjacent normal epithelium tissues from the Lung Cancer Specialized Program of Research Excellence Tissue Bank at The University of Texas M. D. Anderson Cancer Center (Houston, TX), which has been approved by the institutional review board. The tissue had been collected from 1997 to 2001, and the tissue specimens were histologically examined and classified using the 2004 World Health Organization classification system[57] [58]. We selected 311 NSCLC tissue samples (188 adenocarcinomas and 123 squamous cell carcinomas) for our TMAs. TMAs were constructed using triplicate 1-mm diameter cores per tumor, and each core included central, intermediate, and peripheral tumor tissue. Detailed clinical and pathologic information, including demographics, smoking history (never- and ever-smokers), and smoking status (never, former, and current), clinical and pathologic tumor-node-metastasis (TNM) stage, overall survival (OS) duration, and time to recurrence was available for most cases. Patients who had smoked at least 100 cigarettes in their lifetime were defined as smokers, and smokers who quit smoking at least 12 months before their lung cancer diagnosis were defined as former smokers. Tumors were pathologic TNM stages I–IV according to the revised International System for Staging Lung Cancer [59].

Immunohistochemical staining and evaluation. Using anti-SRC-3/AIB-1 mouse monoclonal antibody from BD Transduction Laboratories, CA, USA (Cat #: 611105), immunohistochemical staining was performed as follows: 5-uM FFPE tissue sections

were deparaffinized, hydrated, heated in a steamer for 10 minutes with 10 mM sodium citrate (pH 6.0) for antigen retrieval, and washed in Tris buffer. Peroxide blocking was done with 3% H_2O_2 in methanol at room temperature for 15 min, followed by 10% fetal bovine serum in tris-buffered saline-t for 30 min. The slides were incubated with primary antibody at 4 °C for 90 minutes , washed with phosphate-buffered saline, and incubated with biotin-labeled secondary antibody (Envision Dual Link +, DAKO, Carpinteria, CA) for 30 min. Staining for the slides was developed with 0.05% 3', 3-diaminobenzidine tetrahydrochloride, which had been freshly prepared in 0.05 mol/L Tris buffer at pH 7.6 containing 0.024% H_2O_2, and then the slides were counterstained with hematoxylin, dehydrated, and mounted. FFPE A549 was used as the positive control. For the negative control, we used the same specimens used for the positive controls but replaced the primary antibody with phosphate-buffered saline. For this antibody, we performed titration experiments using a relatively wide range of antibody concentration (1:50, 1:100, 1:200, and 1:500), including the concentration suggested by the manufacturer. One observer (M.G.R.) quantified the immunohistochemical expression using light microscopy (magnification 20×). Both nuclear and cytoplasmic expressions were quantified using a four-value intensity score (0, 1+, 2+, and 3+) and the percentage (0% to 100%) of reactivity. We defined the intensity categories as follows: 0 = no appreciable staining; 1+ = barely detectable staining in epithelial cells compared with the stromal cells; 2+ = readily appreciable staining; and, 3+ = dark brown staining of cells. Next, an expression score was obtained by multiplying the intensity and reactivity extension values (range, 0–300).

EGFR mutation analysis. Exons 18–21 of EGFR were polymerase chain reaction (PCR)-amplified using intron-based primers as previously described [60, 61]. Approximately 200 microdissected FFPE cells were used for each PCR amplification. All PCR products were directly sequenced using the Applied Biosystems PRISM dye terminator cycle sequencing method. All sequence variants were confirmed by independent PCR amplifications from at least two independent microdissections and DNA extraction, and the variants were sequenced in both directions, as previously reported [60, 61].

Array CGH. CGH on cDNA microarrays was carried out as previously described [62]. Briefly, 4 ug tumor and normal sex-matched reference genomic DNA were random-primer labeled with Cy5 and Cy3 respectively, then hybridized to a cDNA microarray (Stanford microarray core) containing ~39,000 cDNAs representing ~26,000 mapped genes/ESTs. Hybridized arrays were scanned on a GenePix scanner (Axon Instruments), and fluorescence ratios extracted using SpotReader software (Niles). Normalized $\log_2$ ratios were then mapped onto the genome using the NCBI genome assembly (Build 36). A preliminary report of these data (focused on *TITF1* amplification) was recently published[63]. Copy number at the SRC-3 locus was determined with the cghFLasso algorithm using a false discovery rate of 0.001[64].

Cell culture. All cells were maintained in RPMI supplemented with 5% fetal bovine serum (FBS, Invitrogen). All these cell lines were established by John D. Minna and Adi Gazdar at the NCI and the Hamon CancerCenter for Therapeutic Oncology.

Reverse transcription. Total RNAs were isolated using Trizol (Invitrogen) reagent following manufacturer's instruction First-strand cDNA was reverse transcribed with 3ug

of mRNA using Superscript II first-strand synthesis system (Invitrogen). The final volume was 20ul.

Quantitative PCR. TaqMan Gene Expression Assays for SRC-3/NCoA3 gene and internal control GAPDH gene are purchased from ABI. To generate the standard curve, serial dilution of MCF-7 cDNAs was run in tetraplicate for both SRC-3 and GAPDH. Efficiency for both primer probe is almost as high as 2 [E=10(-1/slope) =10(-1/-3.3) ≈2(data not shown)], appropriate to use ΔΔCt Method for data analysis.

For the SRC-3 expression in 83 tested cell lines, all the cDNAs were run for both SRC-3 and GAPDH in triplicate. Threshold and baseline were setup the same for both relative standard curve generation and comparative Ct generation for the whole panel. In each group of cells (i.e., NSCLC vs. SCLC vs. HBEC), the cell line expresses lowest level of SRC-3 in the group is defined as 1, and the other lines are relative levels comparing to the lowest one, using [Relative Expression Level = 100 x 2^ (-ΔΔCt)] for calculation.

Immunoblot analysis. Cell lysates were prepared by 1% SDS containing lysis buffer followed by boiling. Cell extracts with equal amount of proteins were analyzed by immonoblotting. The SRC-3/AIB1 antibody (BD Transduction Laboratories) was used at 1:2000 dilutions. Antibodies for cleaved caspase-7 (1:1000) and phos-SRC-3 (Thr24) (1:1000) were from Cell Signaling Technology and antibody for GAPDH was from Santa Cruz.Western blot band intensity is quantified using ImageJ following the program instruction.

Small interfering RNA transfection. Human SRC-3 Smart pool siRNA and nonspecific control siRNA were obtained from Dharmacon Research, Inc. siRNAs were transfected

with Dharmafect 2 (Dharmacon) into H1299, Dharmafect 4 into H1819, Dharmafect 3 into H2073 and Dharmafect 1 into A549 at 50nmol/L following the reverse transfection protocol.

Liquid colony formation assay. 48h after siRNA treatment, cells are harvested and counted. Suspensions of single cells were seeded at cell density 500-2K in triplicate in 6-well plates. The cells were grown for 11 days, 12 days or 14 days for H1299, H1819 and H2073 respectively. Then the colonies were stained with either methylene blue (0.5% in 70% isopropyl alcohol) or crystal violet for 1 hour at room temperature and colonies were counted.

EGFR inhibitor treatment. Gefitinib and Cetuximab were obtained from UTSW pharmacy. Gefitinib was dissolved in DMSO to prepare 10mM solution as drug stock, stored at 4C. NSCLC cell line H1819 was treated with SRC-3 specific siRNA, non-targeting siRNA pool or transfection reagent only. Twenty-four hours after the transfection, 100nM Gefitinib was added to the medium, then observe the cells and harvest them at 72 hours after drug treatment.

Annexin V Staining. H1819 cells treated with above siRNA and/or Gefitinib are trypsinized and counted. After centrifugation, cell pellets of each treatment were resuspended in 1x binding buffer (BD Pharmingen) at a concentration of $1x10^6$ cells/ml. Transfer 100ul solution from each sample, and add 5ul of Annexin V-FITC and 5ul of Propidium Iodide (BD Pharmingen), incubated in the dark at room temperature for 15 minutes, then quenched the staining by adding 400ul of 1x binding buffer, followed by flow cytometry analysis. Percentages of Annexin V positive cells are normalized to mock (transfection reagent) only.

Table 2-1 Multivariate survival analysis for the clinical correlation with SRC-3 expression in lung cancer patients.

(A) Multivariate analysis in disease-free survival.

Effect	p-value	HR	95% CI of HR	
Histology (Squamous vs. Adenocarcinoma)	0.646	1.10	0.74	1.64
Gender (Male vs. Female)	0.481	1.15	0.78	1.71
Race (Caucasian vs. other)	0.967	0.99	0.48	2.04
Tobacco Use (Yes vs. No)	0.260	1.59	0.71	3.56
Stage (II vs. I)	< 0.001	2.37	1.55	3.64
SRC 3 N (High vs. Low)	0.001	1.92	1.30	2.83

(B) Multivariate analysis in overall survival.

Effect	p-value	HR	95% CI of HR	
Histology (Squamous vs. Adenocarcinoma)	0.437	1.18	0.78	1.78
Gender (Male vs. Female)	0.322	1.23	0.82	1.85
Race (Caucasian vs. other)	0.739	1.14	0.53	2.48
Tobacco Use (Yes vs. No)	0.378	1.48	0.62	3.52
Stage (II vs. I)	< 0.001	2.32	1.51	3.57
SRC 3 N (High vs. Low)	0.001	2.02	1.35	3.04

Table 2-2. SRC-3 copy number analysis in NSCLC (A), SCLC (B) and breast cancer (C) cell lines by arrayCGH. NC, no change.

(A)

NSCLC	
Cell Lines	*SRC-3* copy number
Calu-1	Gain
Calu-3	Gain
H23	Gain
H322	Gain
H1650	Gain
H1666	Gain
H1781	Gain
H1993	Gain
H2052	Gain
H2887	Gain
H3255	Gain
HCC193	Gain
HCC461	Gain
HCC1171	Gain
H28	Loss
HCC15	Loss
HCC44	Loss
HCC4006	Loss
A549	NC
Calu-6	NC
H157	NC
H358	NC
H441	NC
H460	NC
H661	NC
H820	NC
H1155	NC
H1299	NC
H1355	NC
H1395	NC
H1437	NC
H1648	NC
H1770	NC
H1792	NC
H1819	NC
H1975	NC
H2009	NC
H2087	NC
H2122	NC
H2126	NC
H2347	NC
H2882	NC
H3122	NC
HCC78	NC
HCC95	NC
HCC366	NC
HCC515	NC
HCC827	NC
HCC1195	NC
HCC1359	NC
HCC1833	NC
HCC2279	NC
HCC2429	NC
HCC2450	NC
HCC2935	NC

(B)

SCLC	
Cell Lines	*SRC-3* copy number
H82	Gain
H128	Gain
H146	Gain
H187	Gain
H1607	Gain
H1672	Gain
H2171	Gain
HCC970	Gain
H524	Loss
H209	NC
H289	NC
H378	NC
H526	NC
H841	NC
H889	NC
H1184	NC
H1339	NC
H1963	NC
H2107	NC
H2141	NC
H2195	NC
H2227	NC
HCC33	NC

(C)

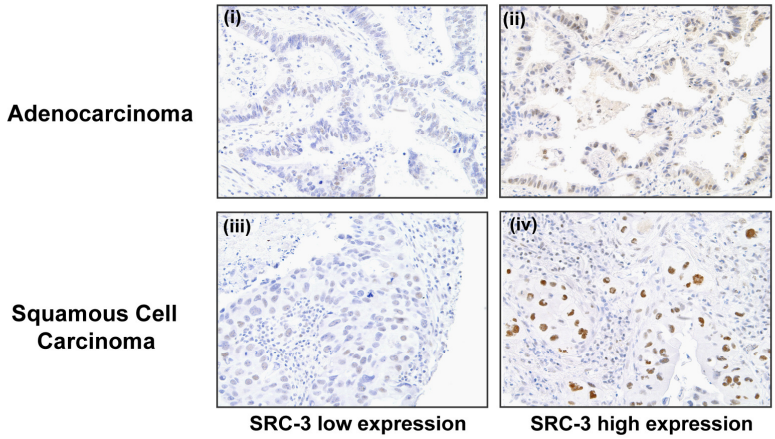
Breast	
Cell Lines	*SRC-3* copy number
HCC712	Gain
HCC1007	Gain
HCC1419	Gain
HCC1500	Gain
HCC1806	Gain
HCC2157	Gain
HCC2185	Gain
HCC2218	Gain
HTB25	Gain
HTB26	Gain
HTB121	Gain
HTB122	Gain
HTB131	Gain
MCF7	Gain
SKBR3	Gain
HCC202	Loss
HCC38	NC
HCC70	NC
HCC1143	NC
HCC1187	NC
HCC1395	NC
HCC1428	NC
HCC1569	NC
HCC1599	NC
HCC1937	NC
HCC1954	NC
HCC2688	NC
HCC3153	NC
HTB23	NC
HTB24	NC
HTB126	NC

(D) Summary of *SRC-3* copy number change in 3 different types of cancer.

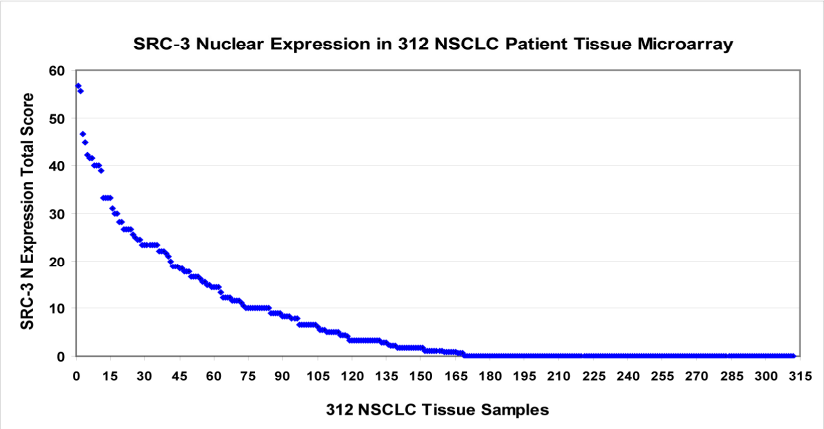
Cancer Type	Cases Tested	SRC-3 copy number		
		Gain (%)	Loss (%)	No Change (%)
NSCLC	55	14 (25%)	4 (7%)	37 (67%)
SCLC	23	8 (35%)	1 (4%)	14 (61%)
Breast Cancer	31	15 (48%)	1 (3%)	15 (48%)

Figure 2-1 Clinical correlation of SRC-3 expression in lung cancer patients.

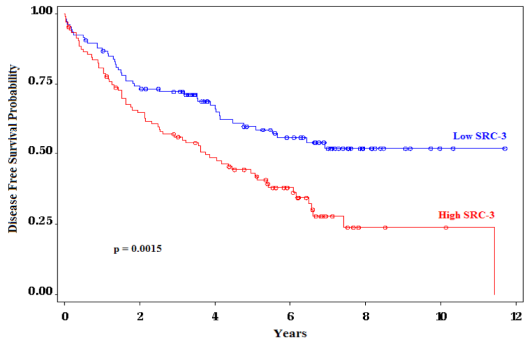
(A) Representative pictures showing SRC-3 low and high expression in two subtypes of NSCLCs.



(B) SRC-3 nuclear expression in 312 clinically localized NSCLCs.



(C) Kaplan-Meier analysis of disease-free survival (left) and overall survival (right) with SRC-3 high expression (red, nuclear staining index>1.6667) or low expression (blue, index ≤ 1.6667).



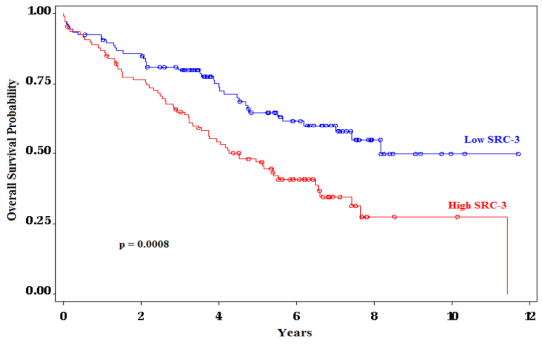
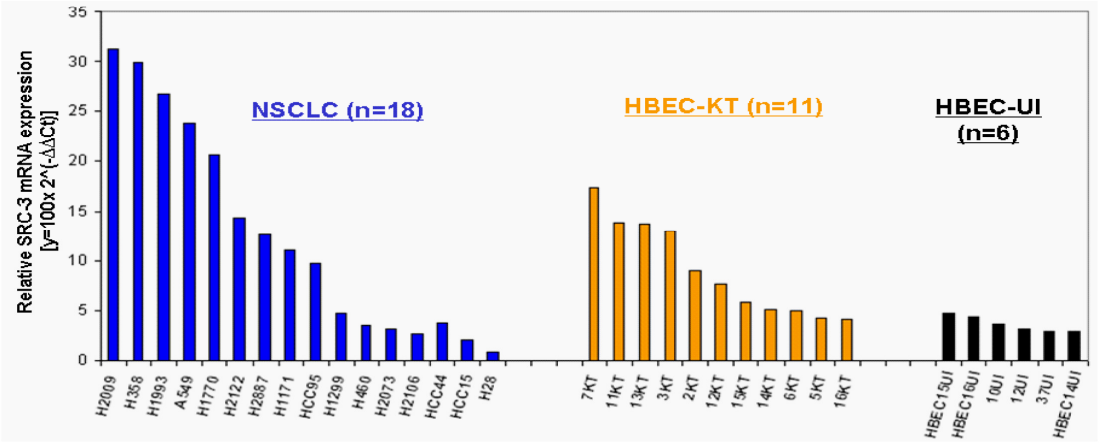


Figure 2-2 Characterization of SRC-3 mRNA- and protein-level expression in a panel of non-small cell lung cancer (NSCLC) cell lines.

(A) SRC-3 mRNA-level expression in NSCLCs and immortalized and primary, un-immortalized human bronchial epithelial cells (HBECs) by Q-RT-PCR.



(B) Total and phosphorylated SRC-3 protein level determined in 8 cell lines by Western blot.

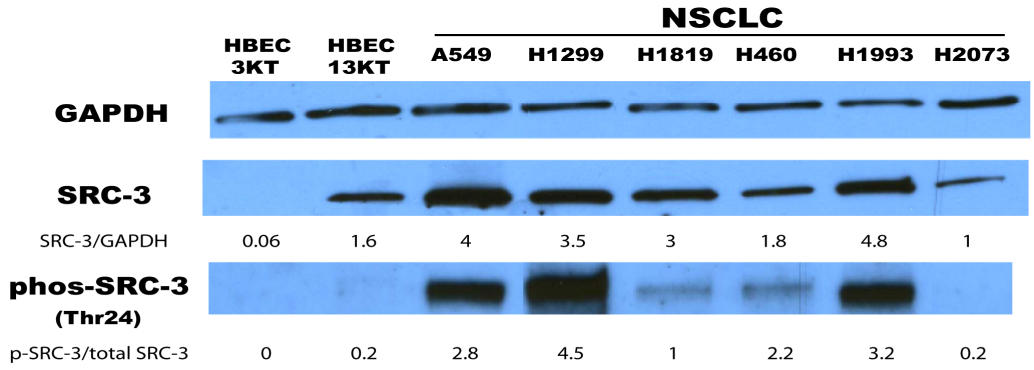
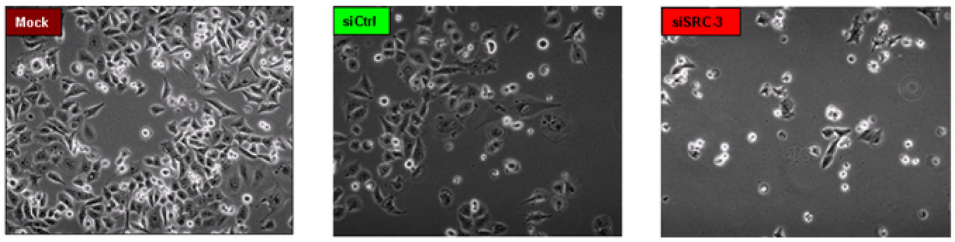
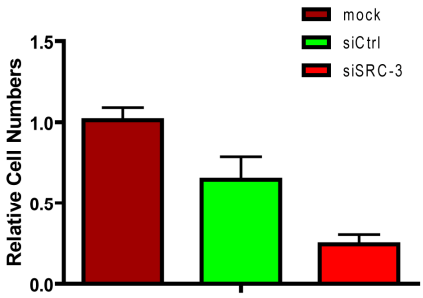


Figure 2-3 Knocking down SRC-3 in NSCLC cell line H1299 which expresses high level SRC-3 decreases cell growth and induced apoptosis.

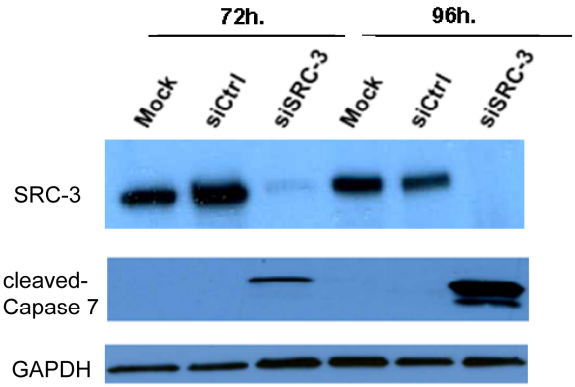
(A) Microscopic examination of cell morphology 96h. after SRC-3-specific siRNA transfection.



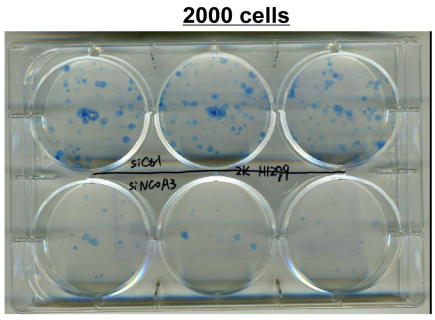
(B) Down-regulation of SRC-3 inhibits cell growth. mock, vehicle control. siCtrl, nonspecific siRNA control; siSRC-3, SRC-3 specific siRNA. **, P<0.001, significant difference with t test.



(C) SRC-3 is effiently depleted with specific siRNA and an upregulated of activated caspase 7 is detected in a time-dependent manner.



(D) Reduction of SRC-3 strongly decreases the colony formation ability of H1299 cells. Representative colony formation plate shown as below.



(E) Quantification of colonies formed in (D). siCtrl, nonspecific siRNA control; siSRC-3, SRC-3 specific siRNA. **, P<0.001, significant difference with Anova test.

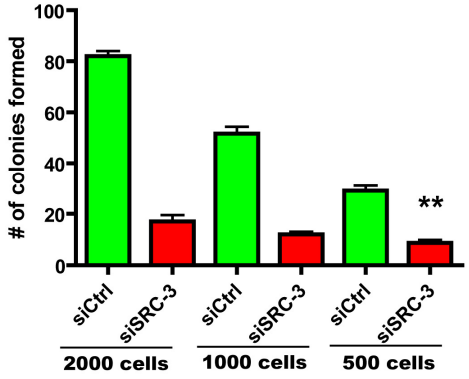
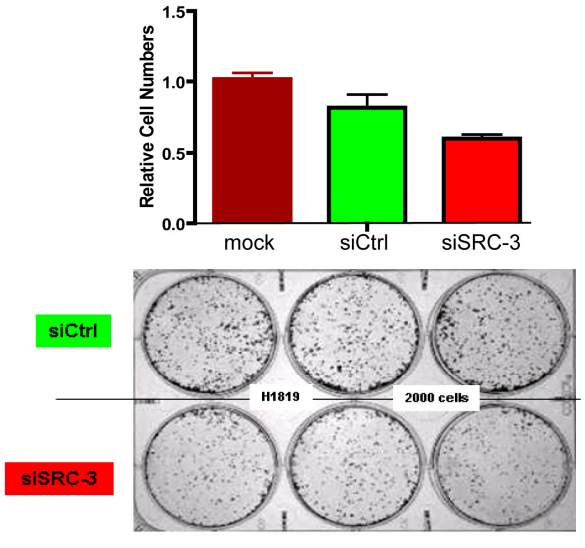
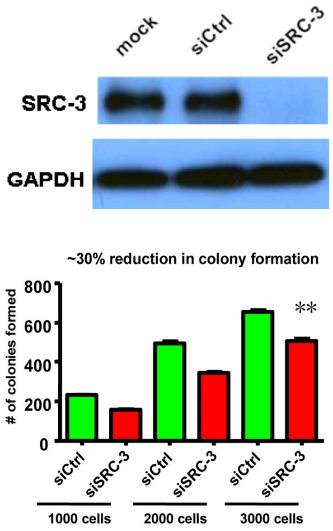


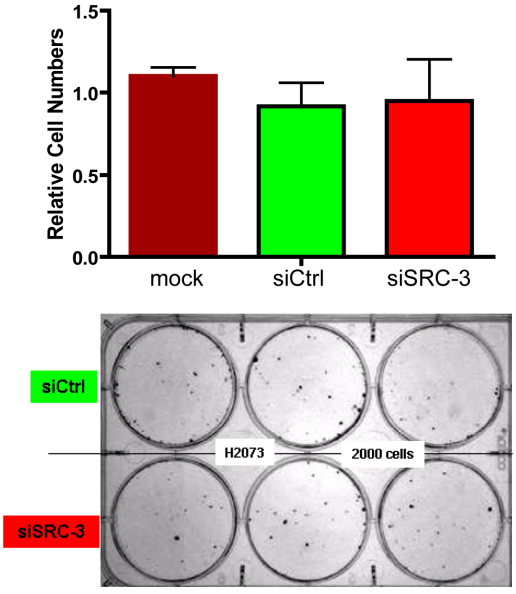
Figure 2-4 The effect of SRC-3 down-regulation on cell phenotype correlates with a cell line's endogenous expression level of SRC-3.

(A) Down-regulation of SRC-3 in H1819 which expresses medium-level of the gene only exerts a mild (~30%) effect in cell growth and colony formation ability. P=0.0047, significant difference with Anova test.





(B) Down-regulation of SRC-3 in H2073 which expresses low-level of the gene shows no detectable change in cell growth and colony formation ability.



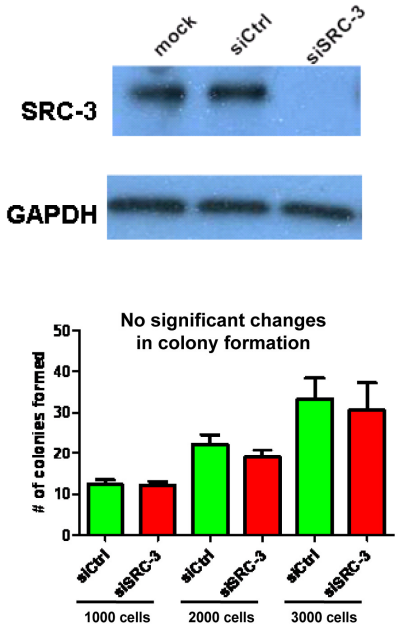
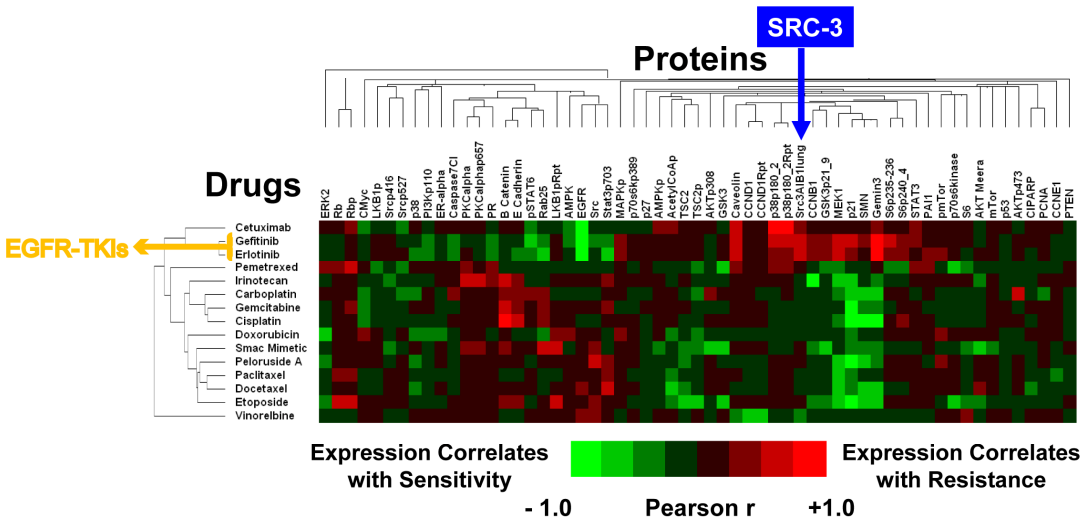
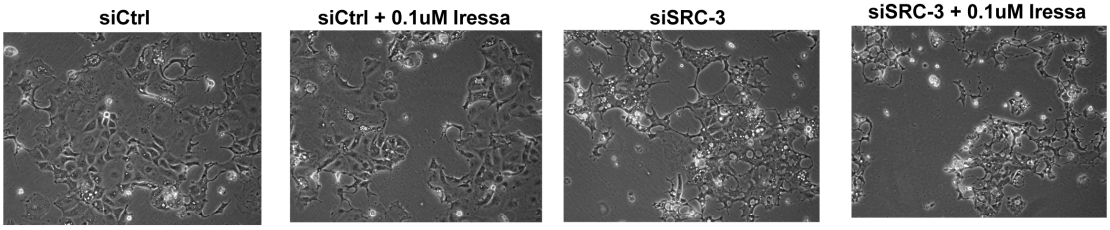


Figure 2-5 SRC-3 knockdown is "synthetically lethal" to EGFR-TKI resistant cells

(A) Reverse Phase Protein Array (RPPA) correlates with drug sensitivity suggests a relationship between SRC-3 high expression with EGFR tyrosine kinase inhibitor (EGFR-TKI) resistance.



(B) Microscopic observation of H1819 cells treated with siRNA oligos alone or together with 0.1uM Iressa.



(C) Down-regulation of SRC-3 combined with Iressa treatment induces apoptosis in H1819 cells.

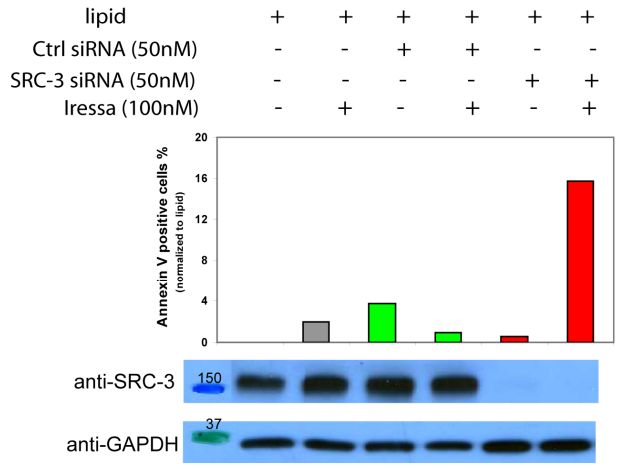
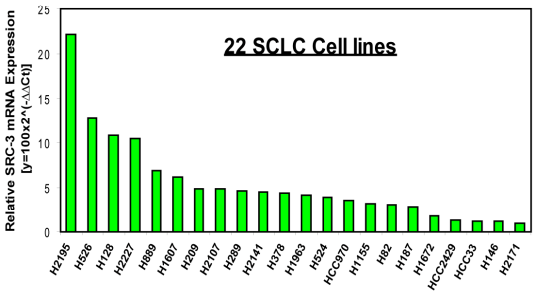
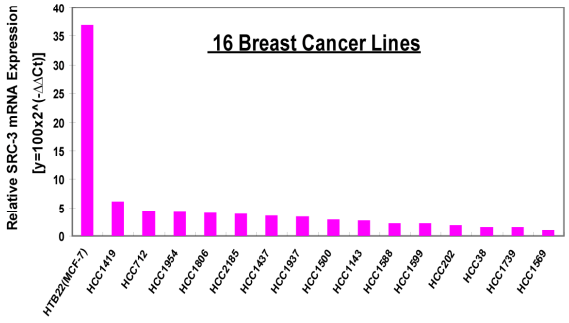


Figure 2-6 mRNA level expression of SRC-3 in a panel of breast cancer, small cell lung cancer and non-small cell lung cancer cell lines.

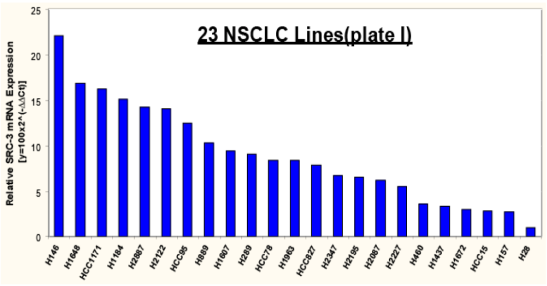
(A) SRC-3 mRNA expression in 22 small cell lung cancer cell lines determined by Q-RT-PCR.

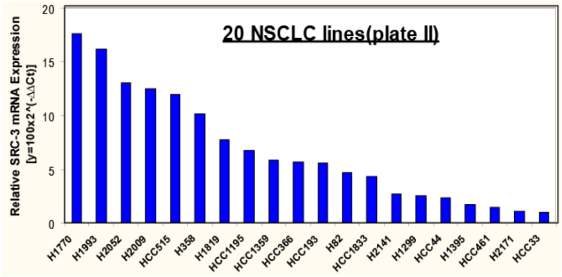


(B) SRC-3 mRNA expression in 16 breast cancer cell lines determined by Q-RT-PCR.



(C) SRC-3 mRNA expression in 43 non-small cell lung cancer cell lines determined by Q-RT-PCR.





(D) mRNA level expression of SRC-3 in selected cell lines from different disease groups.

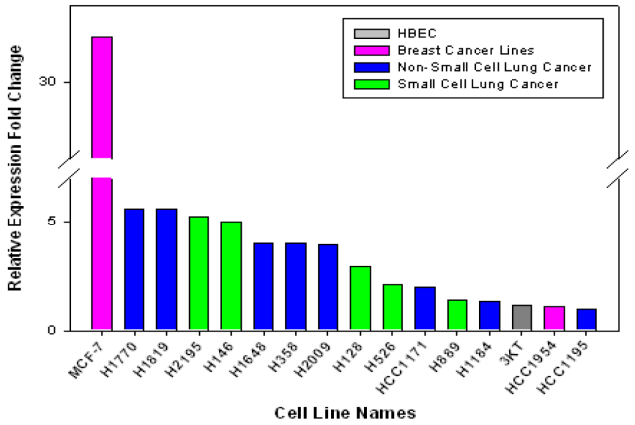


Figure 2-7 Detection of total and phosphorylated level of EGFR, Erb-B2, Erb-B3, AKT and P44/42 expression in human bronchial epithelial cell, and NSCLC cells H358, H1819 and HCC366 all of which harbor wildtype EGFR.

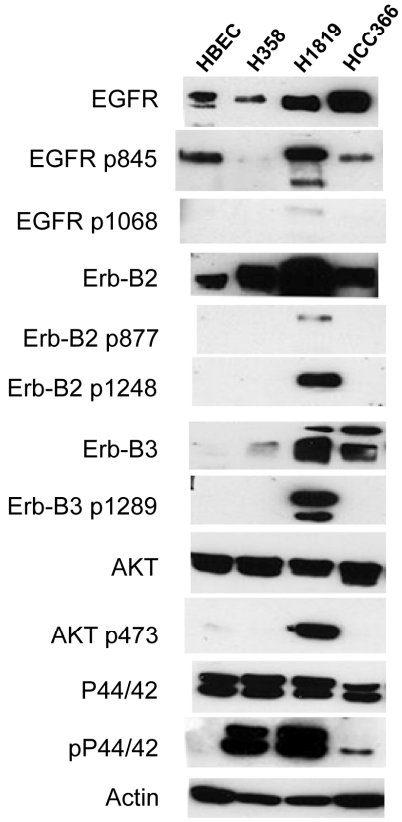
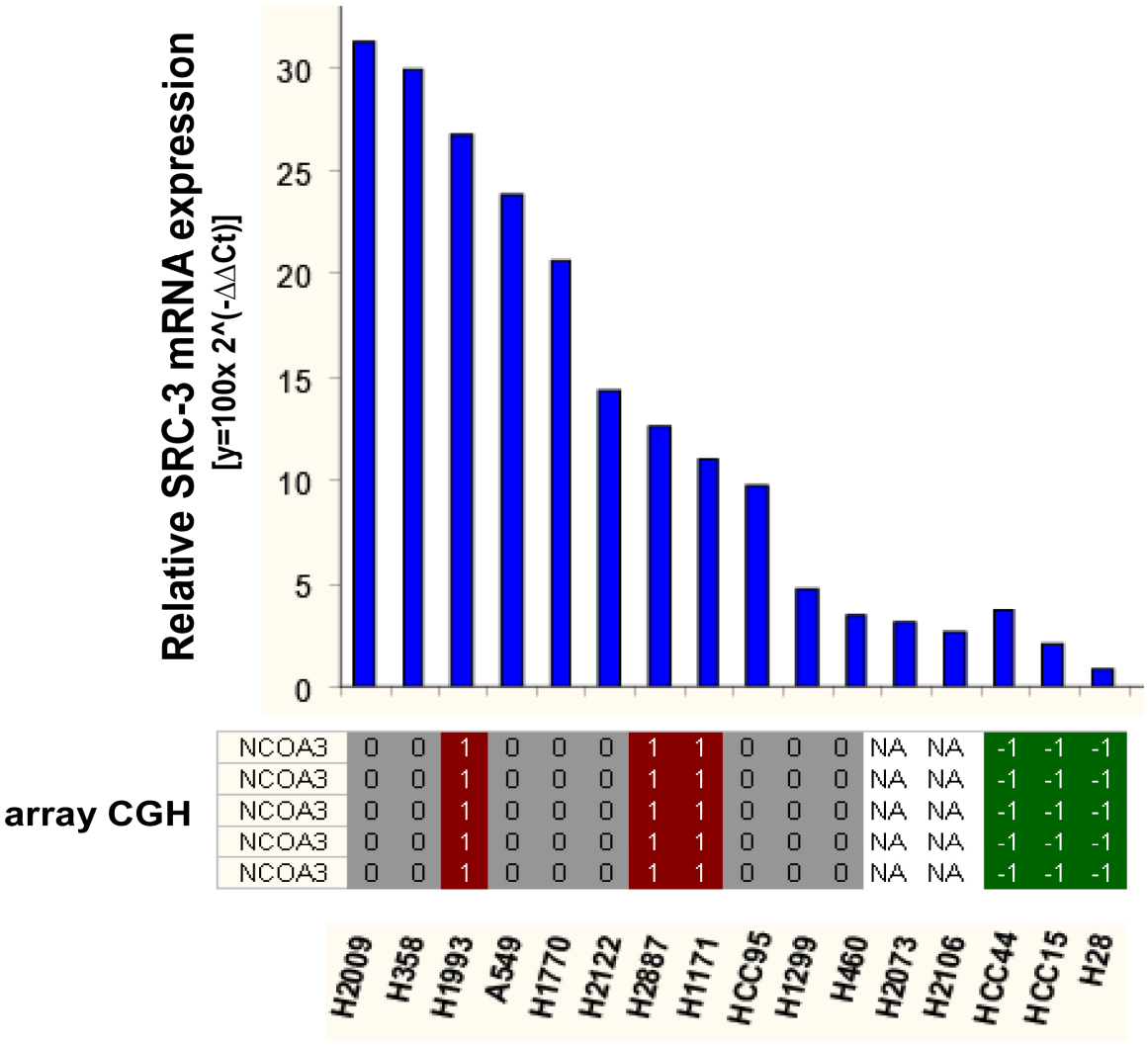


Figure 2-8 SRC-3 gene copy number change correlates with its mRNA level expression.



CHAPTER THREE

INTRODUCTION OF PHOSPHOSPECIFIC LIGAND

Signaling Pathways Regulated by Protein Phosphorylation

Definition of Protein Phosphorylation

Proteins phosphorylation, by definition, is the enzyme-catalyzed tethering transfer of a piece of the ATP molecule (a phosphate) to serine, threonine, or tyrosine on a protein. In prokaryotes, phosphorylation can also occur to histidine residue [65].

Importance of Protein Phosphorylation

Protein phosphorylation is one of the most crucial regulatory mechanisms in nature. It transfers information through cells and tissues and plays an important role in modulating cell decision processes and phenotypes [66]. Protein phosphorylation is involved in most cellular events including proliferation, differentiation, metabolism, transcriptional and translational regulation, protein degradation, cellular signaling and cell survival [67]. Aberrant phosphorylation often results in defective or altered signaling pathways that leads to various diseases, such as cancer, inflammation and metabolic disorders [68]. It has been estimated that more than 50% of all proteins are phosphorylated during their lifetime, and that more than 100,000 phosphorylation sites may exist in the human proteome. Phosphorylation occurs on serine, threonine and tyrosine residues at a frequency ratio of 1800:200:1 [65, 67]. However, protein phosphorylation is usually a very dynamic process, and phosphoproteins are often present at low levels.

Desire for Analytical Strategies for Phosphoproteins

Due to the essential role phosphorylation plays in regulating cell function, there is a need to develop analytical methods and strategies for the characterization of phophorylated proteins. In this chapter, an overview of different approaches for the preparation and enrichment of phosphoprotein samples will be given, with a summary of their advantages, drawbacks and applications.

Phosphoprotein Enrichment Technology

Many approaches have been developed to enrich biological samples for phosphorylated proteins/peptides for phosphoproteomic analysis. The low relative abundance, low phosphorylation stiochiometry and highly dynamic modulation of phosphoprotein demand high sensitivity and specificity of the analytical strategies. So far, many methods have been established, including affinity chromatography and immunoprecipitation.

Chromatofocusing (CF)

The principle of chromatofocusing is to separate the protein based on their isoelectric point (pI) on a liquid chromatography column. Phosphorylation at one site sometimes can cause no change or an acidic shift of pI, which is a decrease of 1-2pH units, depending on the primary and secondary structure of the protein [69]. CF has excellent resolution in separating complex protein mixtures. It creates a linear pH gradient, it is carried out under mild elution conditions at the pI of the proteins, maintaining them in their native conformations. Moreover, the experiment is performed in solution, which simplifies follow-up applications. However, it is very difficult to remove ampholytes from proteins, to scale up the experiment, and CF is sometimes limited by low yields [69]. The common application for CF is to separate a phosphorylated protein from its native unmodified counterpart or protein isoforms with different number of phosphate groups.

Ion Exchange (IEX)

The principle of ion exchange chromatography is to separate the protein based on the differences in the net surface charge. The net charge of a protein is highly pH-dependent because amino acids with different groups can contribute differently to the overall charge at various pHs. At a pH below its pI, a protein carries a net positive charge and will bind to a cation exchanger, whereas at a pH above its pI, a protein carries a net negative charge and will bind to an anion exchanger.

IEX can greatly improve phosphopeptide recovery when used as a prefractionation step prior to subsequent enrichment for large-scale phosphoproteomic studies [67]. However, a very flat and long gradient is required for the elution. Another drawback of this technique is that it is not selective for phosphoprotein from the complex protein mixtures [69]. IEX can be used to separate phosphoprotein from its unmodified form; it can also separate different phospho-isoforms of the same protein.

Immobilized Metal Ion Affinity Chromatography (IMAC)

Phosphoprotein exhibits a strong tendency to chelate metal ions. Metal ions (Fe^{3+}, Al^{3+}, Ga^{3+}, or Co^{2+}) are chelated to iminodiacetic acid (IDA) or nitrilotriacetic acid (NTA) coated beads, forming a stationary phase to which negatively charged phosphopeptides in a mobile phase can bind [67]. Agarose are used to bind phosphorylated protein at pH in the range of 2.0-3.5. While a newly developed resin alkoxide-bridged di-nuclear Zn (II) centre (termed phos-tag) can selectively bind phosphorylated proteins at physiological pH [70, 71]. The bound phosphoyrlated proteins could be eluted either by the addition of free phosphate or by a shift in pH.

A major drawback of this technique is the use of low pH to block binding of aspartate- and glutamate-containing peptides to achieve optimal binding, yet some proteins can not tolerate such condition and will not be recovered in its native form [71]. Another issue is the non-specific binding of acidic proteins and peptides to the IMAC matrix.

Recently, titanium dixodie (TiO_2) chromatography was introduced as an alternative to IMAC. Yet it also has the problem of nonspecifically binding to the acidic proteins. It is generally assumed that TiO2 chromatography has a preference for mono-phosphorylated peptides, while IMAC has a preference for multiply phosphorylated peptides [67]. IMAC and TiO_2 chromatography can be applied for phosphoprotein enrichment, and separation of different phosphoisoforms of the same protein.

Immunoaffinity Isolation

Phosphorylation state-specific antibodies (PSSAs) are useful for studying phosphoproteins involved in numerous biological processes. Polyclonal and monoclonal antibodies have been developed both specifically against primary structure elements of phosphoprotein isoforms and to p-Ser, p-Thr and p-Tyr as such. The phosphoamino acids or phosphopeptides have been coupled to KLH (p-Ser/p-Thr) or BSA (p-Tyr) for immunization. P-Tyr antibodies have a much higher efficiency than the p-Ser and p-Thr antibodies, which are generated mainly based on a consensus sequence flanking the p-Ser or p-Thr residues. P-Ser/p-Thr have much lower immunogenicity compared to p-Tyr [72, 73]. Anti-p-Ser/p-Thr antibodies have much lower specificity than the anti p-Tyr antibody. This fact also explains why phosphorylation in Tyr residues has been

investigated so intensively despite its low abundance compared to phosphorylation in Ser and Thr residues [74].

Major problems associated with PSSA immunoaffinity isolation include the difficulty in generating p-Ser/p-Thr specific antibody, low yield in antibody purification, high costs of immunoaffinity media, and loss of structure and activity. Most p-Ser/Thr-specific antibodies are not suitable for phosphoprotein immunoprecipitation for unknown reason, though they are very specific to certain consensus motif [74]. This may be close to the fact that the linear epitope that they recognize is not exposed completely in the native protein.

Conclusion remarks

Most of the techniques described above are not totally selective for phosphoproteins, and they are limited to some certain scales. They might provide some specific advantages, but their drawbacks are also quite obvious. One way to optimize the current approaches is to use a combination of different methods in the right order, for example, Sequential elution from IMAC (SIMAC) that combines the strengths of IMAC and those of TiO2 has provided a significant improvement of phosphopeptide enrichment [69]. Alternatively, novel research tools need to be developed to analyze phosphoproteins in a specific manner.

Synthesis and Screening of Peptoid Libraries for the Isolation of Protein Ligands

Background of Peptoid Chemistry

Peptoids, oligomers of N-substituted glycines, are developed as a motif for the generation of chemically diverse libraries. Peptoids are synthesized by "sub-monomer" method involving addition of an activated derivative of α-bromoacetate to the N-terminal nitrogen of the growing chain, followed by displacement of the bromide with a primary amine [75]. Unlike peptides, peptoids are not sensitive to peptidases or proteases, this metabolic stability makes possible the *in vivo* application of peptoids, which traditional peptides can hardly achieved [76, 77].

Synthesis of Peptoid Libraries

Large combinatorial libraries of peptoids can be synthesized easily and inexpensively. Split and pool synthesis is used to construct a one-bead one-compound peptoid library on hydrophilic TentaGel beads. Primary amines, hundreds of which are available commercially, serve as the building blocks of the library to represent the chemical diversity of the library.

Screening of Peptoid Libraries

For the screening experiment, the peptoid library is exposed to the protein of interested (POI), followed by primary antibody recognizing the POI and secondary antibody conjugated with quantum dots, which are nanoparticle semiconductors that exhibit a tremendous Strokes shift and emit intensely red fluorescence, overcoming the relatively high autofluorescent background emission of the beads when they are

irradiated with UV light [78, 79] . The hits from the peptoid library are identified by Edman sequencing or tandem MS-MS. Finally, biochemical assays are performed to validate the hits from the screening experiments.

Solid Support Selection for Peptoid Library

TentaGel resins were selected for both library synthesis and target protein screening for the below reasons: 1) They have good swelling properties in both organic and aqueous solutions which facilitates both for efficient synthesis using organic solvents and allows the library to be screened on the same support in aqueous buffer, 2) The poly(ethylene glycol) (PEG) chains that comprise the outer shell of the resins provide a "nonsticky" surface that is ideal for reducing nonspecific protein binding during screening experiments. 3) They are mechanically stable, and are not easily broken during the shaking and washing steps in the screen. 4)They have a sufficiently high loading capacity that the sequence of the hits from a screen can be determined sensitively by direct Edman sequencing or mass spectrometry [80].

Examples of Protein Ligands Identified from Peptoid Library

Many studies from our laboratory and others have shown that peptoid libraries are rich sources of protein-binding ligands [81-85]. The strategy for ligand identification is usually bead-based screens for binders of your protein of interest. The molecules selected from these simple binding assays generally affect the function of the target protein. For example, Xiao et. al have reported the discovery of a cell-permeable synthetic transcription factor mimic which is a chimera containing a DNA-binding

hairpin polyamide linked to a CBP-binding peptoid [83]. Lim et.al has identified a peptoid inhibitor of the proteasome 19S regulatory particle with Sug2/Rpt4 as its molecular target [84, 86]. Udugamasooriya et.al have identified a peptoid antagonist of VEGF receptor 2 from a two-color, cell based screen [82]. Very recently, Simpson et.al isolated a Cholera toxin (CT)-binding peptoid, when immobilized on TentaGel beads, can act as a high affinity toxin capture agent to sequester bacterial toxins [81].

Taken together, these data suggest that peptoid libraries are an inexpensive and convenient source of protein ligands. The on-bead screening strategy using quantum dots for visualization provide a high throughput and high efficiency method for the identification of protein binding molecules.

DOPA Cross-linking Chemistry

Periodate-triggered DOPA Cross-linking Reaction

Protein-protein interactions are of fundamental importance in almost all biological processes. Chemical cross-linking is a potentially powerful method to analyze protein-protein interaction and protein-small molecule interaction. Among all the chemical cross-linking strategies, oxidative coupling reactions stand out from the rest, because of its short reaction times and high yield. Previously, our lab explored 3,4-dihydroxylphenylalanine (DOPA) as a protein cross-linking reagent upon periodate oxidation [87]. Under these conditions, DOPA is converted to an ortho-quinone intermediate that can be attacked by nucleophiles in close proximity, resulting in a stable cross-link [88, 89]. Unlike most chemical cross-linking methods, DOPA cross-linking occurs in high yield with little or no nonspecific products observed even in complex protein mixtures [90].

Applications of DOPA Cross-linking Chemistry

DOPA cross-linking is an operationally simple, highly specific and high yield reaction. This technique is useful for the analysis of protein-protein and protein-ligand interactions in large complexes. For example, the methodology has been applied to identify the direct binding partners of the Gal4 transactivation domain in the proteasome, a >2M Da complex [91]. DOPA contained in RIP-1 (regulatory particle inhibitor peptoid-1) was used to identify Sug2/Rpt4 subunit in the 19S proteasome as the molecular target of RIP-1 [86]. Moreover, periodate-triggered cross-linking identified a residue-to-residue cross-link between Ste2p, a yeast GPCR, and its peptide ligand alpha-factor with a

DOPA moiety, defining a specific contact point between the bound ligand and its receptor [92]. Furthermore, a novel label transfer system using DOPA as cross-linker can be used to better characterize protein-protein interactions without covalent modification of the protein of interest and hopefully to identity unknown protein-binding partners in cell lysates [93].

Mechanism of DOPA Cross-linking Chemistry

A mechanistic study on the chemistry of periodate-mediated DOPA-protein cross-linking using a peptide nucleic acid templated system revealed that the terminal α-amino, ε-amino groups of Lys, the imidazole of His, and the thiol of Cys are capable of attacking the DOPA-derived ortho-quinone, respectively. The researchers also showed that other aliphatic 1,2-diols that are abundant in cells, such as lactose and ATPs, do not interfere with the periodate-mediated cross-linking reaction, indicating this methodology can be useful for the analysis of macromolecular interactions on cell surfaces or other environments that are rich in 1,2-diols [89].

Conclusion Remarks

Taken together, these and other results obtained using DOPA cross-linking reaction suggest that it is of potentially powerful utility in probing intereactions between biomolecules in complex environments. Therefore, I will use this approach in my phosphospecific peptoid ligand to cross-link the target protein, facilitating its use as a "western blotting" reagent surrogate.

Background of Brd4 Biology and the Discovery of the PDID Domain

Background of Brd4 and HPV E2

Brd4 is a double bromodomain-containing protein that binds preferentially to acetylated chromatin. Brd4 plays crucial role in cell growth, it has been implicated in cell cycle control, DNA replications and many other cellular processes [94]. Brd4 is used as a cellular adaptor by some animal and human papillomaviruses (HPV) for anchoring viral genomes to mimotic chromosomes. This tethering, mediated by Brd4 interaction with virus-encoded transcription factor E2 protein, facilitates viral genome segregation during mitosis [95, 96]. It has been observed that E2 interacts with the C-terminal motif (CTM) of Brd4 for its transactivation activity [97].

Discovery of phos-PDID as a Novel E2-intereaction Domain

Lee & Chiang have identified a novel E2-interacting region of Brd4, phosphorylation-dependent interaction domain (PDID). Interestingly, PDID domain interacts only with the cancer-inducing high risk E2, such as HPV-16E2 and HPV-18E2, in a phosphorylation-dependent manner, but not with the wart-causing low risk HPV-11E2. Moreover, a cell-based assay showed that the over-expression of PDID domain serves as a dominant negative mutation specific for high-risk E2, indicating the interaction between PDID and high-risk E2 is important for the E2-responsive transcriptional activation. Taken together, their study demonstrates that the phosphorylation of Brd4 provide selectivity for high-risk E2, thereby presenting an alternative regulatory mechanism for the protein encoded by cancer-inducing HPVs [98].

CHAPTER FOUR

DEVELOPMENT OF PHOSPHOSPECIFIC LIGANDS

Abstract

Most proteins can exist in a variety of post-translationally modified forms. Chemical methods that would allow one to specifically purify or pharmacologically target a particular form of the protein would be of great interest. Here, we report the first peptidomimetic compounds that bind specifically to a serine-phosphorylated PDID domain of Brd4 protein, identified by screening a library of 40,000 peptoids for PDID binders. The isolated hit peptoids are only specific to phos-PDID, but not the non-phosphorylated form of the protein, or other phosphoserine- or phosphothreonine-containing proteins. . Phos-PDID-binding peptoids can specifically capture a recombinant phos-PDID from a crude insect cell extract, without binding to the unmodified PDID in bacteria lysate. Moreover, the phosphospecific peptoid ligand engineered with a Biotin tag and a DOPA crosslinker can specifically detect phos-PDID from whole cell lysate, demonstrating its potential as a "Western blotting"-like reagent. Furthermore, GST pull-down assay and reporter gene assay reveal that the peptoid ligand can specifically disrupt the interaction between phos-PDID and high-risk HPV 18E2, and hence inhibits the Brd4-dependent transcription activation in human cervical cancer cells. Taken together, these data showed that our phosphospecific peptoid ligand is able to substitute phosphospecific antibodies for the detection and isolation of phosphoproteins; it can also

perhaps be developed as a drug-like compound targeting the active form of protein in cells.

Introduction

Protein phosphorylation is one of the most crucial regulatory mechanisms in nature. It plays an important role in modulating cell decision processes and phenotypes [66]. Aberrant phosphorylation often results in defective or altered signaling pathways that leads to various diseases [68]. So far, many analytical methods and strategies used for phosphoprotein enrichment and characterization have been established, including chromatofocusing (CF), ion exchange column and immobilized metal ion affinity chromatography (IMAC), which separates proteins based on the differences of their isoelectric points (pI), net surface charge, and tendency to chelate metal ions, respectively. However, the major limitation of all of these methods is the non-selectivity for phosphorylated protein [69]. Phosphoylation state-specific antibodies (PSSAs) can be developed with specificity against either the primary structure elements of phosphoprotein isoforms or to p-Ser, p-Thr and pTyr as such. However, due to the low immunogenicity of p-Ser and p-Thr residues, it is very difficult to generate highly specific anti-p-Ser/p-Thr antibodies [72, 73]. Other limitations of PSSA immunoaffinity isolation are the low yield and high cost during antibody generation process, and a lack of capability for phosphoprotein immunoprecipitation, probably due to the linear epitope that they recognize is not exposed completely in the native protein [74].

Because of the limitations of above-described methods, it is of great interest to develop novel strategies using inexpensive and easy-to-make synthetic compounds that display high phosphospecificity as PSSAs for the analysis of phosphorylated proteins. We and others have demonstrated that libraries of peptoids (oligo-N-substituted glycines) are rich sources of protein-binding ligands[81-85]. Unlike peptides, peptoids are

insensitive to peptidases or proteases and are very simple and economic to synthesize [76, 77]. Moreover, our lab has previously developed on-bead screening strategy using quantum dots for visualization, providing high throughput and high efficiency for the identification of protein binders from large peptoid libraries[78, 84]. In this work, we demonstrate the identification of a ligand specific to a phosphorylated protein from a peptoid library, and its application in "Western blot-like" analysis and "immuprecipitation" of the targeted phosphorylated proteins.

The phosphorylation-dependent interaction domain (PDID) of Brd4 protein was chosen as a target for investigating this strategy. Brd4 is a double bromodomain-containing protein that binds preferentially to acetylated chromatin. It is a member of the BET (bromodomains and extraterminal) family that includes mammalian Brd2, Brd3, yeast Bdf1, Bdf2 and corresponding homologues in other species[94]. Two tandem bromodomians (BDI and BDII) and a SEED (Ser/Glu/Asp-rich regin) motif are highly conserved among BET family members[99]. Mammalian Brd4 plays a crucial role in cell growth, it has been implicated in cell cycle control, DNA replication and many other cellular processes[94]. Brd4 is used as a cellular adaptor by some animal and human papillomaviruses (HPV) for anchoring viral genomes to mitotic chromosomes. This tethering is mediated by the interaction between the C-terminal motif (CTM) of Brd4 and virus-encoded transcription factor E2 protein [95-97].

Lee & Chiang have identified a novel E2-interacting region of Brd4, phosphorylation-dependent interaction domain (PDID). Interestingly, the PDID domain interacts only with the cancer-inducing high risk E2, such as HPV-16E2 and HPV-18E2, in a phosphorylation-dependent manner, but not with the wart-causing low risk HPV-

11E2. Moreover, a cell-based assay showed that the over-expression of the PDID domain serves as a dominant negative mutation specific for high-risk E2, indicating the interaction between PDID and high-risk E2 is important for the E2-responsive transcriptional activation. Taken together, their study demonstrates that the phosphorylation of Brd4 provides selectivity for high-risk E2, thereby presenting an alternative regulatory mechanism for the protein encoded by cancer-inducing HPVs[98].

Because of its biological importance and availability, phos-PDID protein was used as a target for our exploration of phosphospecific ligands. We report here the first peptidomimetic compounds that bind specifically to the phospho-PDID, but not the non-phosphorylated version of the protein, or other phosphoserine- or phosphothreonine-containing protein tested. Phos-PDID-binding peptoids can specifically capture phos-PDID from crude insect cell extract, but not the unmodified PDID in bacteria lysate. Moreover, the phosphospecific peptoid ligand engineered with Biotin tag and a DOPA crosslinker can specifically detect recombinant phos-PDID from a whole cell lysate, potentiating its use as a "Western bloting"-like reagent. Furthermore, GST pull-down assay and reporter gene assay reveal that the peptoid ligand can specifically disrupt the interaction between phos-PDID and high-risk HPV 18E2, and hence inhibit the Brd4-dependent transcription activation in human cervical cancer cells. Taken together, these data showed that our phosphospecific peptoid ligand is able to substitute for phosphospecific antibodies for the detection and isolation of phosphoproteins. This methodology may also provide a potentially general way to identify phosphorylated protein-binding ligands.

Brd4 Protein Contains Two CK2 Phosphorylation Regions: Brd4(287-530) and Brd4 (698-785)

One of the BET family member, yeast Bdf1 (yBdf1), has two separate CK2-mediated phosphorylation regions and the phosphorylation in these regions is necessary for *in vivo* yBdf1 function [100]. Since Brd4 belongs to the same protein family as Bdf1, it is possible that it is also phosphorylated in the corresponding regions and uses a similar mechanism for functional regulation. Thus, Lee and Chiang attempted to search for putative CK2 phosphorylation sites on Brd4. Two dominant phosphorylation regions were predicted on Brd4 and are similarly positioned to those in yBdf1: One is Brd4(287-530) domain which locates in the region approximately 20 amino acids away from the C-terminal BDII, and the other is Brd4 (598-785) which locates in the SEED domain. To determine if the proteins from insect cells are indeed phosphorylated, Brd4 (287-530) and Brd4 (598-785) purified from Sf9 cells were treated with alkaline phosphatase from calf intestine (CIP), and a clear mobility shift is observed in SDS-PAGE (**Figure 4-1 B & C**). In a reciprocal experiment, *in vitro* kinase assay was performed with recombinant human CK2 and bacterial Brd4 (287-530) as its substrate. Commassie gel staining showed that CK2-treated bacterial protein co-migrated with insect cell protein(**Figure 4-1 B**). These data indicate that both Brd4 (287-530) and Brd4 (698-785) are indeed phosphorylated in Brd4 when expression in Sf9 cells. Brd4 (287-530) is named as phosphorylation-dependent interaction domain (PDID), the biological function of PDID domain will be described elsewhere[98]. This paper will focus on using PDID domain as a target to identify phosphospecific ligand.

Peptoid Library Construction and Hit Screening

To identify phos-PDID-specific ligands, a "one bead one comound" peptoid library with a theoretical diversity of 14^4 (38, 416) compounds was synthesized on Tentagel beads using the "split and pool" method [76]. This was achieved by incorporating the fourteen amines shown in **Figure 4-9** via microwave-assisted submonomer peptoid synthesis[101]. These amines were chosen so that the library would include a variety of physiochemical properties. Two Nlys (diaminobutane) residues were conjugated at C-terminus as constant spacers to facilitate identification of hit compounds by Edman sequencing.

For screening, approximately 150, 000 library beads were blocked with 4% BSA and then exposed to Flag-tagged phos-PDID (500nM) in the presence of 4% BSA in physiological buffer (PB buffer, composition described in **Materials and Methods**). After incubation, the beads were washed thoroughly and probed with anti-Flag monoclonal antibody and a secondary anti-IgG antibody conjugated with red-emitting quantum dots. The beads were then visualized under a fluorescent microscope, and those that displayed a red halo (16 total, a representative picture shown in **Figure 4-2**) were isolated manually using a micropipette. The isolated beads were stripped of protein with 1% SDS and then reprobed with anti-Flag antibody and Qdot-conjugated secondary antibody to identify compounds that bind directly to primary or secondary antibodies rather than the phos-PDID. After this post-screening, only two beads (0.0013% of the library) did not bind primary or secondary antibodies directly and thus presumably displayed phos-PDID ligands. These beads were subjected to Edman degradation, and the

sequence of each peptoid was identified. We named one of the peptoids as DC-pPDID-1 and the other as DC-pPDID-2 (**Figure 4-2**).

Both DC-pPDID-1 and DC-pPDID-2 Specifically Bind to the Phos-PDID, but not the Non-phosphorylated Form of the Protein

The strategy we used for the screening was designed to isolate ligand for phos-PDID protein, yet we cannot rule out the possibility that the identified hit could bind both phos-PDID and non-modified PDID, since no counterscreen aginst the unmodified protein was done. To determine whether these peptoids can selectively bind to the phosphorylated form of PDID, both compounds were re-synthesized on TentaGel macrobeads for the on-bead binding assay. Phos-PDID protein purified from sf9 cells was dephosphorylated with calf intestinal alkaline phosphatase (CIP), and nonphos-PDID protein purified from bacterial cells was phosphorylated by casein kinase 2 (CK2) (**Figure 4-1**). PDIDs with or without enzymatic treatments were incubated with hit peptoids displayed on beads, followed by anti-FLAG monoclonal antibody incubation and Qdot-655-conjugated secondary antibody incubation using same conditions as described in the library screening section. The hit peptoids incubated with phos-PDID protein emitted red-fluorescence, confirming them as true phos-PDID binder, this signal was strongly diminished if the phos-PDID was dephosphorylated with CIP prior to incubation. In the reciprocal experiment, a signal was not observed when both peptoids were incubated with non-modified PDID, but the beads emitted red fluorescence when the non-modified protein was phosphorylated by CK2. These results demonstrate that the hit peptoids isolated from an unbiased library can bind specifically to the phosphorylated PDID, but not the non-phosphorylated form of the protein (**left and middle columns in Figure 4-3 A & B**).

To test whether the hit peptoids are general serine-, threonine-phosphorylated protein binder, another phosphorylation domain of Brd4 protein which covers amino acids (598-785) was also included in the on-bead binding assay. Brd4 (598-785) has comparable molecular weight and similar number of tentative CK2 phosphorylation sites as phos-PDID[100]. The same protocol was applied to phos-Brd4 (598-785) and Brd4 (598-785) with both hit peptoids. DC-pPDID-1 shows no detectable signal to either protein, indicating this peptoid does not bind the phosphorylated or unphosphorylated Brd4 (598-785). A weak signal was observed when DC-pPDID-2 was incubated with phos-Brd4 (598-785), suggesting it might bind to phos-Brd4 (598-785) with an affinity much lower than with phos-PDID (**right column in Figure 4-3A and 4-3B**). This result suggests that these peptoids are specific to phos-PDID, rather than general binders to any phos-Serine/Threonine –containing proteins.

DC-pPDID-1 can Specifically Capture Phos-PDID Protein in the Presence of Whole Cell Lysate, without Interacting with Nonphos-PDID Protein

Since DC-pPDID-1 seems to have higher phosphospecificity than DC-pPDID-2, subsequent characterization experiments were mainly focused on DC-pPDID-1 peptoid. To evaluate the capacity of this compound to bind and isolate phos-PDID protein in the presence of large amounts of competitor proteins, DC-pPDID-1 was resynthesized on Rink Amide beads with a cysteine reside at the C-terminus of the peptoid to facilitate compound quantification and immobilization on agarose gels. The HPLC purified Cys-DC-pPDID-1 was quantified using Elman's reagent and immobilized to SulfoLink Coupling Gel (Pierce) following the manufacturer's protocols. L-cysteine was also conjugated to the Sulfolink agarose beads to serve as negative control. An equal amount of DC-pPDID-1 or L-Cysteine-coated resins (~6ul) were incubated with both sf9 cell lysate that expresses phos-PDID, and bacterial cell lysate that expresses nonphos-PDID individually (**Figure 4-4A**). After removing the supernatant (unbound proteins) and washing, bound proteins were stripped from the beads with 1% SDS in a volume equal in each experiment to permit comparison between the output (elute) from different combinations of bead incubation by Western blotting. As expected, control beads that are not conjugated with a peptoid ligand do not bind the phos-PDID. Cys-DC-pPDID-1-conjugated agarose beads can specifically capture phos-PDID protein from the whole insect cell lysate with a 20% recovery rate, but showed no binding affinity with nonphos-PDID protein in the presence of bacterial cell lysate (**Figure 4-4B**). We conclude from these data that DC-pPDID-1 peptoid shows high specificity towards phos-PDID among

the existence of whole cell lysate, indicating this compound can be developed into affinity-purification tools to enrich the targeted phosphorylated protein.

Chemically Engineered DC-pPDID-1 Peptoid Serves as Analytical Tool for Phos-PDID Protein Detection

With a phosphospecific peptoid in hand, we then tried to determine if this peptoid can be developed into a "Western blotting"-like reagent to detect phos-PDID proteins from complex protein mixtures. To achieve this goal, a highly efficient chemical cross-linker should be anchored to "stabilize" the peptoid when it specifically recognizes and binds to the target protein; and a Biotin tag should be incorporated to facilitate the followup probing on the PVDF membrane using neutraavidin-HRP (NA-HRP) conjugate (**Figure 4-5A**). Previously, our lab have demonstrated that 3, 4-dihydroxylphenylalanine (DOPA) can cross-link efficiently to proteins when triggered by peroidate. DOPA is converted to an ortho-quinone intermediate that can react with nucleophiles, such as lysine, histidine, and cysteine residues on the proteins in close proximity [87-89]. Unlike most chemical cross-linking methods, DOPA cross-linking occurs in high yield with little or no nonspecific products observed even in complex protein mixtures[90]. Therefore, a DOPA-conjugated DC-pPDID-1 peptoid with a Biotin tag on its C-terminal is synthesized and HPLC purified (**Figure 4-5B**).

0.1uM and 0.5uM of Biotin-DOPA-DC-pPDID-1 (BD-DC-pPDID-1) were incubated with 0.5uM immunopurified Flag-tagged phos-PDID. After addition of sodium periodate ($NaIO_4$) and a brief incubation, the reaction was quenched and cross-linked products were analyzed by SDS-PAGE and transferred to PVDF membrane. When probed with NA-HRP for the detection of biotin-conjugated peptoid-protein complex, a band (~40kD) corresponding to the size of phos-PDID was observed in the presence of 0.5uM peptoid and 2mM $NaIO_4$ (**Figure 4-5 C-(2), lane 2**). No signal is detected when

peptoid concentration is reduced to 0.1uM or when $NaIO_4$ is absent, indicating BD-DC-pPDID-1 can only cross-link to phos-PDID protein at a concentration ≈ [its target protein] in the presence of $NaIO_4$. (**Figure 4-5 C-(2), lane 1&3**).

Next, we tried to determine whether BD-DC-pPDID-1 can specifically cross-link to phos-PDID in the context of whole cell lysate. As shown in the colloidal blue staining gel in **Figure 4-5 C-(1)**, sf9 cell lysates containing either 10% or 20% of phos-PDID protein were used for the cross-linking experiment with the same conditions for the purified protein alone. When blotted with NA-HRP, a major band corresponding to the size of phos-PDID was detected, together with three other bands of unknown identities (**Figure 4-5 C-(3)**). Same experiment was also performed in sf9 lysate containing 5% of phos-PDID, yet only a smear of bands were seen on the NA-HRP blots (data not shown). These results demonstrate that the sensitivity of BD-DC-pPDID-1 reagent in the current cross-linking protocol is limited to detect its target protein at ~10% abundance in the lysate.

The Hit Peptoid Specifically Inhibits the Interaction between Phos-PDID and High-risk HPV-18E2 in a Dose-dependent Manner, without Affecting the Binding between the CTM of Brd4 and 18E2

Lee & Chiang have discovered that the phosphorylated PDID domain of Brd4 can selectively bind to the high-risk HPV E2 protein, such as HPV-16E2 and HPV-18E2, while the C-terminal domain (CTM) of Brd4 has no selectivity in binding with high-risk or low-risk E2[98]. We were curious whether the phos-PDID-specific peptoids are able to inhibit the interaction between phos-PDID and HPV E2 proteins *in vitro*. To examine this protein-protein interaction disruption effect of the Hit peptoids, we employed a GST-pull down assay that involves exposing the mixture of Flag-tagged phos-PDID protein and GST-tagged 18E2 on glutathione beads, with gradient concentration of peptoids in solution. A reaction with bacterially purified CTM also was performed as a control experiment. As shown in **Figure 4-6 B**, DC-pPDID-1 disrupts the interaction between phos-PDID and 18E2 protein in a dose-dependent fashion, with an IC_{50} of ≈100-200 uM. However, the same concentration of DC-pPDID-1 peptoid had no effect on the binding between C-terminal motif (CTM) of Brd4 and 18E2 (**Figure 4-6 C**). Together these data substantiate the high specificity of hit peptoids towards the phos-PDID protein during protein-protein interaction event.

Phosphorylated-PDID-specific Hit Peptoids can Inhibit the Brd4-dependent Transactivation in Mammalian Cells

Brd4 is known be very important for the E2-responsive transcription activation through the interaction between CTM of Brd4 and E2 in cultured cervical cancer cell C33A that expresses endogenous level of Brd4 [95-97]. Recently, Lee & Chiang reported that the over-expression of PDID domain serves as a dominant negative mutation specific for high-risk E2 in a cell-based assay, indicating the interaction between PDID and high-risk E2 is important for the E2-responsive transcriptional activation(**Figure 4-7 A**) [98].

To determine whether phos-PDID-specific peptoid can inhibit the Brd4-dependent transcription activation in mammalian cell culture, 100uM of DC-pPDID-1 & DC-pPDID-2 were transfected into C33A cells using Chariot delivery reagent, followed by the cotransfection of reporter gene and HPV-18E2 expression plasmids. As shown in **Figure 4-7 B**, both peptoids can inhibit 18E2-responsive luciferase activity by two to three fold with 10ng E2 transcription factor or 50ng, respectively. This result mimics the effect of over-expressing phos-PDID domains as described above. To test whether this result can be achieved without artificially facilitating entry of the compound into the cells by transfection, 100uM of DC-pPDID-1 & DC-pPDID-2 were directly added to the medium, before the reporter gene and 18E2 plasmids were co-transfected. Unfortunately, no inhibitory effect is observed without the assistance of Chariot (data not shown), indicating both the peptoids have insufficient cell permeability to achieve effective concentration inside the cells. In summary, it can be concluded that peptoids Hit G and Hit I which specifically bind to phosphorylated PDID are able to disrupt the Brd4-dependent transcription activation in cell culture when assisted by transfection reagent.

Identification of the Pharmacophore of Phos-PDID Specific Peptoids.

To identify residues in DC-pPDID-1 and DC-pPDID-2 critical for phos-PDID binding, a series of cysteine-conjugated derivatives was synthesized in which each side chain was, in turn, replaced by a methyl group, hence mutating each residue of the hit peptoids into sarcosine(**Figure 4-8 A**). Similar to the alanine scanning technology used in biochemistry to identify important amino acids for protein-protein interaction, this "sarcosine scanning" approach generates "point mutants" synthetically rather than through encoding at the DNA level[102]. All the compounds were purified by HPLC, the molar concentration was quantified by Elman's Reagent. Equal amount of the Sarcosine-scanned derivatives were immobilized to the Sulfolink agarose beads, and tested for phos-PDID pull down abilities from whole cell lysate using the same assay as described in **Figure 4-4**. Equal amount of the SDS stripped sample from the Sulfolink beads were loaded onto SDS-PAGE gel. As shown in **Figure 4-8 B**, substitution of G2 (N_{Lys}), G4(N_{Lys}), G6(N_{Tyr}), I2(N_{Lys}) and I4 (N_{Lys}) into sarcosine essentially abolished the binding ability to phos-PDID protein, about 5-fold less potent than the wildtype compounds. Mutations at G3 (N_{Mea}-Sar), G6 (N_{Leu}-Sar), I3 (N_{Ffa}-Sar), I5 (N_{Leu}-Sar), I6(N_{App}-Sar) residues show little or no effect on the protein capture activity, indicating these residues are dispensable for phos-PDID binding. In summary, we have identified the pharmacophore in the phos-PDID-specific peptoid as N_{Lys}-X-N_{Lys}-N_{Tyr} where X can be any residue.

Discussion

Herein, we report the first example of peptidomimetic compounds that bind specifically to a serine-phosphorylated protein. We stress the phosphospecificity of this molecule, without cross-binding to the non-phosphorylated form of same protein, or other serine-phosphorylated proteins, a feature that is similar to what a phospho-state specific antibody (PSSA) possesses. The power of this molecule is that it not only specifically recognizes its target protein, but is also able to capture phos-PDID from crude cell extract, without interaction with the unmodified PDID in bacteria lysate, enabling peptoids to specifically pull down your phosphoprotein of interest (pPOI) in an "immunoprecipitation"-like manner. Moreover, DC-pPDID-1 engineered with a Biotin tag and a DOPA crosslinker can detect phos-PDID specifically in a whole cell lysate on PVDF membrane using neutraavdin-HRP conjugate. However, the sensitivity of the DOPA-Biotin- DC-pPDID-1 reagent in the cross-linking reaction is relatively low, which is most likely due to the weak affinity between the peptoid and its target protein.

In addition to distinguishing between the phosphorylated and non-phosphorylated form of the same protein, our peptoid DC-pPDID-1 is also able to specifically disrupt a protein-protein interaction. As shown in the GST pull-down assay (**Figure 4-6**), DC-pPDID-1 can specifically inhibit the interaction between phos-PDID and high-risk HPV 18E2, without interfering with the binding between CTM of Brd4 and E2 protein, again suggesting our peptoid is a highly specific ligand. It is possible that the peptoids and high risk 18E2 bind to the same interaction surface on phosphorylated PDID so that they compete with each other for interaction with this protein. Since the interaction between phos-PDID and HPV-E2 is important for the E2-responsive gene transcription, our

peptoid, which specifically interrupts this interaction *in vitro,* may also have an effect on the transactivation activity *in vivo*. Indeed, the luciferase assay in cervical cancer cells that express endogenous level of Brd4 confirmed our hypothesis (**Figure 4-7**), suggesting this peptoid might be useful to study the biology of phosphoproteins in the cell-based assays. However, our phosphospecific peptoid has relatively low cell permeability, as it can only be delivered into the cells via the assistance of transfection reagents such as Chariot. The low permeability probably results from the three positively charged Nlys residues, as the charge-carrying monomers impair the ability of peptoids to penetrate into cells[103]. This drawback restricts its use in *in vivo* animal experiment or other assays that require high cellular permeability.

One way to transform this high specificity, low affinity compound into a high specificity, high affinity one is through medicinal chemistry modification. This strategy would involve the identification of residues required for specific binding to phos-PDID, the creation of large numbers of derivatives of the minimal pharmacophore, followed by ligand screening of this new library using a much more stringent condition. To identify residues that are essential for phosphospecific peptoid DC-pPDID-1, sarcosine scanning was performed which revealed a pharmacophore of N_{Lys}-X-N_{Lys}-N_{Tyr} (X can be any residue) as its minimum required motif for interaction with phos-PDID protein. We are speculating that the positively charged N_{Lys} residues may provide some binding affinity for peptoid interaction with the negatively charged phos-PDID proteins and the N_{tyr} provide some kind of specificity towards PDID protein. Some of our unpublished data also indicate that the sequence containing the motif Nlys-X-Nlys is important for the binding with PDID, as motif Nlys-X-X-Nlys and Nlys-X-X-X-Nlys show decreasing

binding affinity with phos-PDID proteins. Investigation of the structure-activity relationship (SAR) of DC-pPDID-1 peptoid would give us some insight into how the phosphospecific peptoids bind the phosphorylated proteins. The next step in this work will be to synthesize new library containing the Nlys-X-Nlys-Ntyr pharmacophore and assay it for compounds with significantly increased potency.

It is well known that protein phosphorylation is one of the most important regulatory mechanisms in nature. A lot of proteins are not active to exert their physiological or pathological roles unless being phosphorylated. This work is a proof of principle study demonstrating the ability of peptoids to distinguish between an activated (phosphorylated) protein and its inactive form. Our ultimate goal is to determine whether a certain phosphorylation-dependent signaling cascade is active or not in a selected cell environment, which can be achieved by isolation an array of peptoids recognizing different phosphorylated proteins in the pathway of interest.

In conclusion, we have demonstrated that a highly phosphospecific ligand can be identified from a modest-sized peptoid library. The peptoid can recognize and precipitate its target phosphoprotein from a complex protein mixture in a specific manner. These findings strongly support further development of DC-pPDID-1 for application as phosphospecific antibody surrogate. Our data also set the stage for technology platform to the isolation of phosphospecific peptoids to a broad range of phosphorylated proteins. Finally, this work provides the possibility of identifying peptoid ligands specific for other forms of post-translational modifications, such as methylation, acetylation and ubiquitination.

Materials and Methods

General Remarks: All chemicals and solvents were purchased from commercial suppliers and used without further purification. TentaGel macrobeads (140-170 um diameter, 0.48 mmol/g capacity, Rapp Polymere) were used to construct the library. Synthesis of the compounds was performed on Rink Amide MBHA resin (Nova Biochem). Preparative HPLC was performed on a Waters Breeze HPLC system with a Vydac C18 preparative column. Flow rate was 10ml/min. HPLC runs used linear gradients of 0.1% TFA and 90% acetonitrile plus 0.1% TFA. Mass spectrometry was performed on all synthesized compounds with the MALDI-Voyager DE Pro instrument.

Peptoid Library Construction: The tetramer peptoid library with two constant C-terminal Nlys (diaminobutane) residues was synthesized by split-and-pool method on TentaGel macrobeads (3.0g, 1.44mmol) using sub-monomer approach by a microwave-assisted protocol[80, 101]. Fourteen monomer amines with diversified functional groups were used as building blocks. The synthesized library has a theoretical diversity of 14^4=38, 416 compounds (**Figure 4-9**). The library beads were then washed thoroughly with DMF and dichloromethane, a solution (5ml) of 95% trifluoroacetic acid, 2.5% triisopropylsilane and 2.5% water was added and stirred at 220rpm at room temperature for 2hr. The solution was drained and library beads were neutralized with 20% diisopropylethylamine (DIPEA) in DMF, then washed with DMF 6 times, and finally rinsed with dichloromethane.

Screening of the peptoid library. The peptoid library displayed on TentaGel macrobeads (~150,000) were equilibrated with physiological buffer (PB buffer, 10 mM Tris-acetate (pH 7.4), 10 mM magnesium acetate, 1 mM dithiothreitol, 0.5 mM EDTA,

90 mM potassium acetate, 10mM b-glycerophosphate and complete mini protease inhibitor cocktail (Sigma),).for 1hr. After decanting PB buffer, the resulting beads were incubated with PB buffer containing 5% BSA at 4°C overnight. The beads were subsequently incubated with 500nM purified phos-PDID protein in 5% BSA in PB buffer at 4°C for 4hr, the buffer was removed and beads were washed with 1xTBST 3 times. Then the beads were treated with anti-Flag M2 monoclonal antibody (Sigma, 1:500 dilution) in 4% BSA-containing PB buffer at 4°C overnight. After washing with 1x TBST 3 times, the beads were treated with Qdot 655-coated secondary antibody (Quantum Dot Corp, 1:300 dilution) in 4% BSA-containing PB buffer for 2h at 4°C after washing. The buffer was removed and the beads were washed with TBST 3 times. The peptoids that bind to the phos-PDID protein, and/or anti-FLAG M2 monoclonal antibody and/or Qdot 655-coated secondary antibody were visualized under a fluorescence microscope by irradiation of the beads through a DAPI filter (390-410 nm band pass filter), and the red-emitting beads were picked up using a micropipette. All the red beads were heated at 95°C in 1% SDS for 30 min. The beads were then washed successively with 1x TBST 3 times. Again all the beads were incubated with anti-FLAG primary antibody and Qdot 655-coated secondary antibody with exactly the same conditions as mentioned above, and subjected to washing and fluorescence microscope visualization. This time, the red-emitting beads, which are binders to the primary or secondary antibodies, are discarded.The remaining two beads, named as “Hit G” and “Hit I”, are heated in 1% SDS and washed as described above. The washed hit beads were then subjected to automated on-bead Edman sequencing analysis.

Synthesis of DC-pPDID-1 & -2 and the sarcosine-scanned derivatives. After removing the Fmoc protecting group on Rink Amide MBHA resin, Fmoc-Cys (tert)-OH was conjugated to the beads, then the peptoids residues involved in Hit G, Hit I and all the sarsosine-scanned derivatives were conjugated using a microwave-assisted protocol in a scale of 0.35 mmol, respectively[80, 101], after The resulting compounds were cleaved from the resin by 95% TFA, 2.5% water and 2.5% triisopropylsilane at room temperature for 2hr. TFA was removed and the crude products were purified by C18 RP-HPLC. The identities of the hit peptoids and all the sarcosine-scanned derivatives (Sar G-1~Sar G-6, Sar I-1~Sar I-6) were confirmed by MALDI-TOF analysis: DC-pPDID-1: observed 910.8, expected 910.5 for $C_{42}H_{75}N_{11}O_9S$ + H. DC-pPDID-2: observed 937.4, expected for 937.6 for C43H76N12O9S + H.

Plasmid construction

Bacterial expression plasmid pF:hBrd4(287-530)-11d was constructed to prepare Brd4 deletion mutant proteins from BL21(DE3)pLysS. The corresponding coding sequence was generated by PCR amplification with an NdeI site-containing sense primer and BamHI site-containing antisense primer using pcDNA3-F:hBrd4(FL) as a template [97]. After purification and restriction enzyme treatment of the PCR product, it was cloned between NdeI and BamHI sites by swapping the insert with pF:TBP-11d. For the baculovirus expression plasmids, the intermediate plasmid pF:hBrd4(287-530)-7 was cloned as described previously for the generation of pF:hBrd4(1-722)-7 [97]. Then, the XbaI-BamHI fragment containing Brd4-coding sequence was transferred between XbaI and BamHI sites of pVL1392 (Invitrogen) to generate pVL-F:hBrd4(287-530).

Affinity purification of PDID proteins from insect and bacterial cells. FLAG-tagged PDID protein was purified from insect Sf9 cells infected with baculoviruses harboring pVL-F:hBrd4(287-530) following the published protocol [104]. Bacterially expressed FLAG-tagged PDID proteins were purified following the published protocol [105].

***In vitro* dephosphorylation assay.** 200 ng of purified protein from insect cells was incubated with or without 1 µl of alkaline phosphatase (1 U/µl, Roche) at 30°C for 1 hr in a 20 µl reaction. To analyze the mobility shift by dephosphorylation reaction, the sample was separated by SDS-PAGE and visualized through Coomassie staining.

Kinase assay. 200 ng of bacterially purified Brd4 protein was incubated with 200µM of ATP-containing buffer in the absence or presence of 50 U of CK2 (New England Biolabs) at 30°C for 30 min. in a 20 µl reaction. The product was analyzed through Coomassie staining. The reaction was processed using a Typhoon 9200 PhosphorImager.

GST pull-down assay. To analyze protein-protein interactions, GST pull-down was performed as described previously[106]. Briefly, GST alone or GST-tagged 18E2 were expressed in BL21(DE3)pLysS and the bacterial lysates were prepared. For the assay, 400 ng of GST proteins were immobilized onto 10 µl of glutathione-SepharoseTM 4B beads (Amersham Pharmacia Biotech) at 4°C overnight. After washing, the beads were incubated with 100 ng of purified PDID or CTM of Brd4 at 4°C for 1 hr and washed again. 50 µl of 2X protein sample buffer was added to the protein-bound beads and the sample was analyzed using Western blotting.

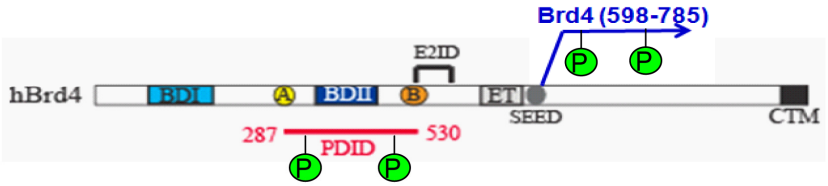
HPV-18E2-responsive luciferase assay. Human cervical cancer C33A cells were placed in 24-well plate and grown in Dulbecco's modified Eagle medium (Invitrogen) supplemented with 10% fetal calf serum (Invitrogen) for 24h. 100uM of peptoids were

transfected into the cells using Chariot reagent following manufacturer's instruction. After 24h., 10ng or 50ng of HPV-18E2 expression plasmid was co-transfected into the cells with 100ng reporter gene. Luciferase activity was measured 36h. after the transfection. The luminescence from the blank that contains no 18E2 or reporter gene expression was subtracted from each measured firefly luciferase luminescence. The result for the control cells to which no 18E2 was transfected was treated as 1, the baseline for the induction. Other results were normalized based on this baseline and finally presented as fold of inductions.

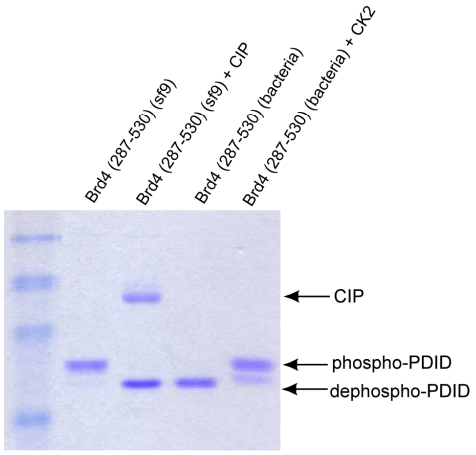
Chemical Cross-linking. Cross-linking reactions were done in PB buffer (see components above). Concentrations of purified phos-PDID protein or sf9 lysate were determined by BSA assay. Different concentrations of phos-PDID protein in the presence or absence of sf9 lysate were incubated with 0.5uM of biotin-DOPA-DC-pPDID-1 in a 10ul reaction. The mixture was left for 10min at room temperature until $NaIO_4$ was added at a final concentration of 2mM. After 1min incubation, cross-linking reaction was quenched with 6x loading buffer containing 100mM DTT. The samples were separated using standard SDS-PAGE.

Figure 4-1 Background of Bromodomain4 (Brd4) and PDID domain.

(A) Both Brd4 (287-530), named as PDID (phosphorylation-dependent intereaction domain) and Brd4 (598-785) are highly phosphorylated by CK2.



(B) *In vitro* kinase assay and dephosphorylation assay in PDID domain.



(C) *In vitro* dephosphorylation assay in Brd4 (598-785).

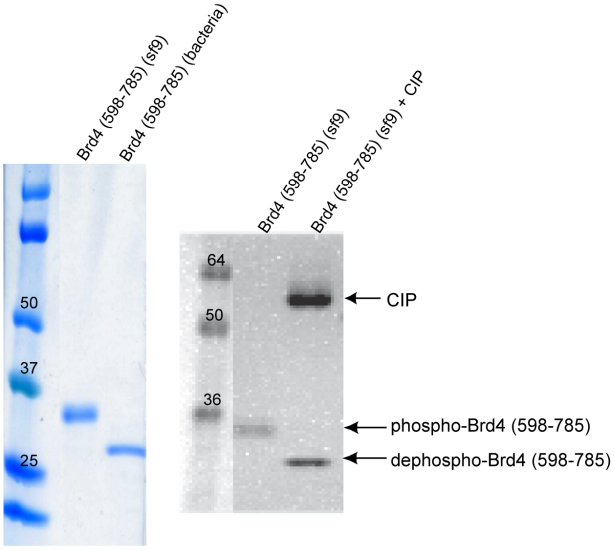


Figure 4-2 Isolation and characterization of phospho-PDID binding peptoid.

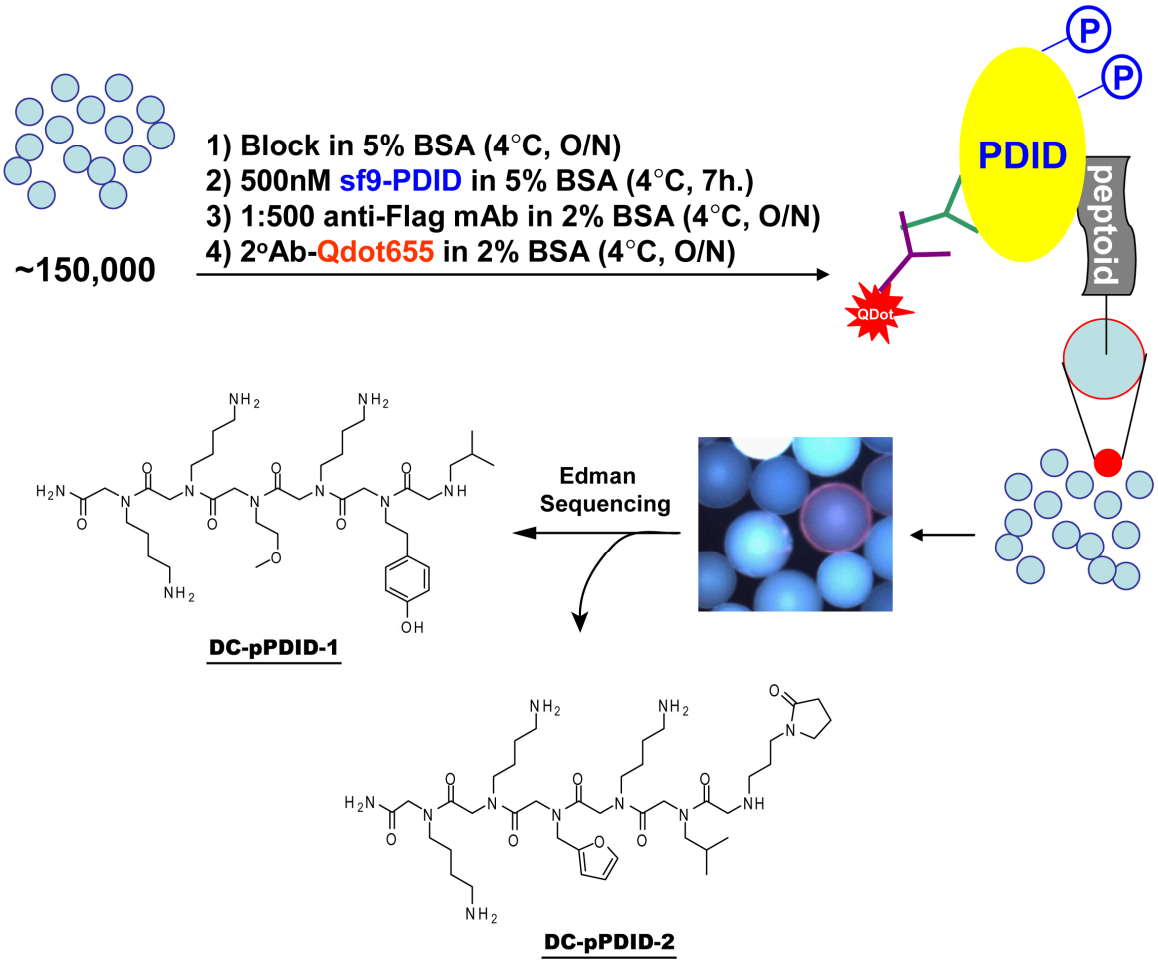
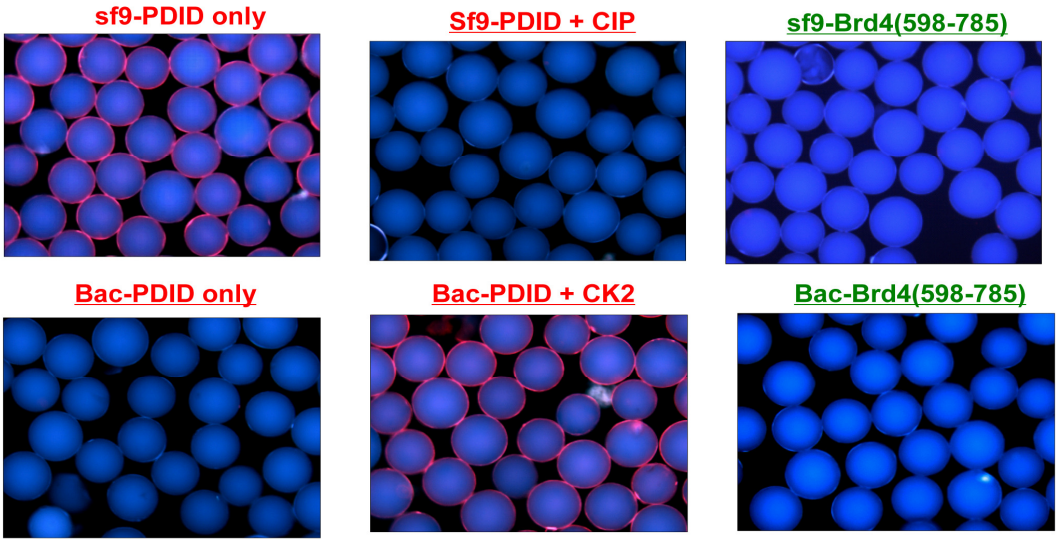


Figure 4-3 DC-pPDID-1 and DC-pPDID-2 specifically bind to the phospho-PDID, but not the non-phosphorylated form of the protein, or another highly phosphorylated domain from the same Brd4 protein.

(A) On bead binding assay between immobilized-DC-pPDID-1 peptoid and i) phos-PDID; ii) phos-PDID +CIP; iii) phos-Brd4 (598-785); iv) nonphos-PDID; v) nonphos-PDID +CK2 and vi) nonphos-Brd4 (598-785) proteins in solution.



(B) On bead binding assay between immobilized-DC-pPDID-2 peptoid and protein samples as as described as in (A).

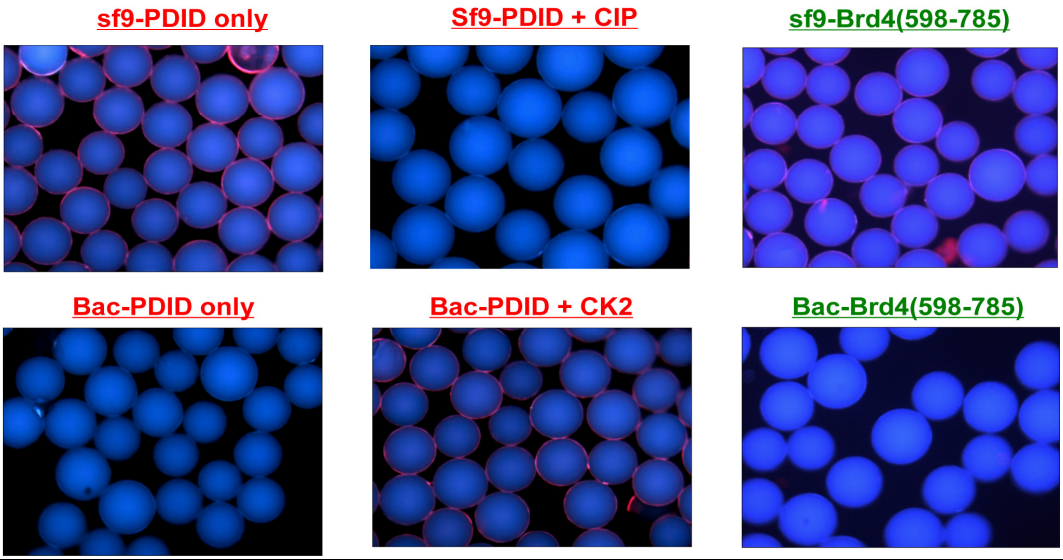
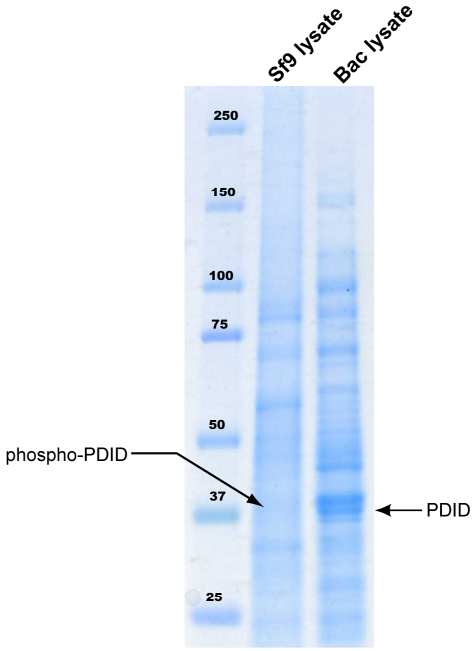


Figure 4-4 Peptoid DC-pPDID-1 can capture phosphorylated-PDID protein from sf9 cell lysate, but not the non-phosphorylated form from bacterial cell extract.

(A) Colloidal blue staining of input lysates of sf9 cells that express phos-PDID, and bacterial cells that express unphospho-PDID.



(B) DC-pPDID-1 peptoid immobilized on agrose gel can "immuno-precipitate" phosphorylated PDID protein from the insect cell lysate with a ~20% recovery, but not the non-phosphorylated PDID from bacterial cell extract. Lane 1: 20% of the total input sf9 lysate. Lane 2: 50% of the protein retained by negative control cysteine. Lane 3: 50% of the protein retained by DC-pPDID-1. Lane 4: 20% of the total input bacterial lysate. Lane 2: 50% of the protein retained by negative control cysteine. Lane 3: 50% of the protein retained by DC-pPDID-1.

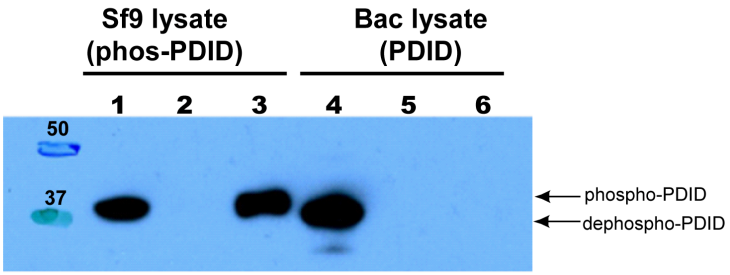
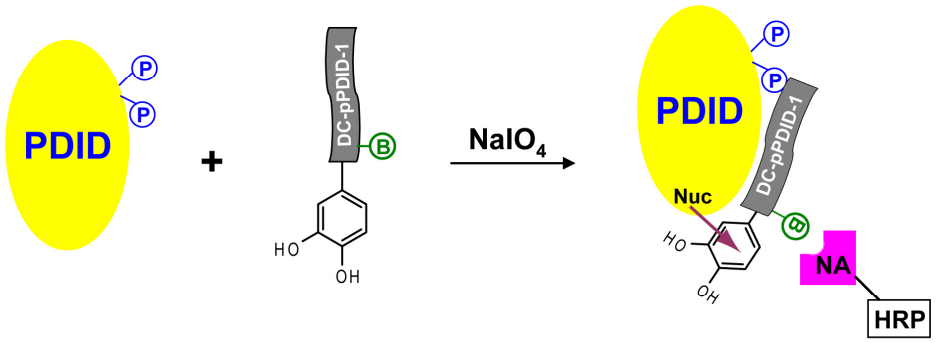
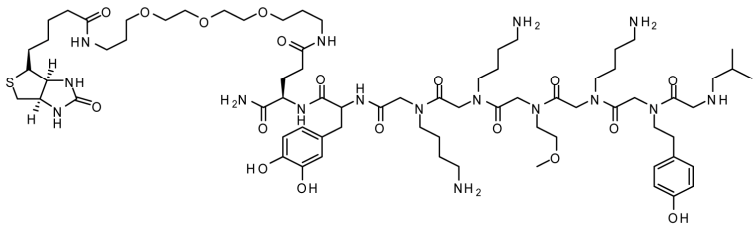


Figure 4-5 Biotin- and DOPA-conjugated DC-pPDID-1 peptoid serves as analytical tool for phos-PDID protein detection.

(A) Cartoon illustratiing mechanism of action of Biotin-DOPA conjugated DC-pPDID-1 peptoid. B=Biotin, NA-HRP = Neutra-avidin horse radish peroxidase.



(B) Structure of biotinylated DC-pPDID-1 conjugated with DOPA crosslinker (BD-DC-pPDID-1).



Biotin-DOPA-DC-pPDID-1 (BD-DC-pPDID-1)

(C) The cross-linking product between Biotin-DOPA-DC-pPDID-1 and phos-PDID can be detected by Neutra-Avidin-HRP, even in the presence of crude cell extract.

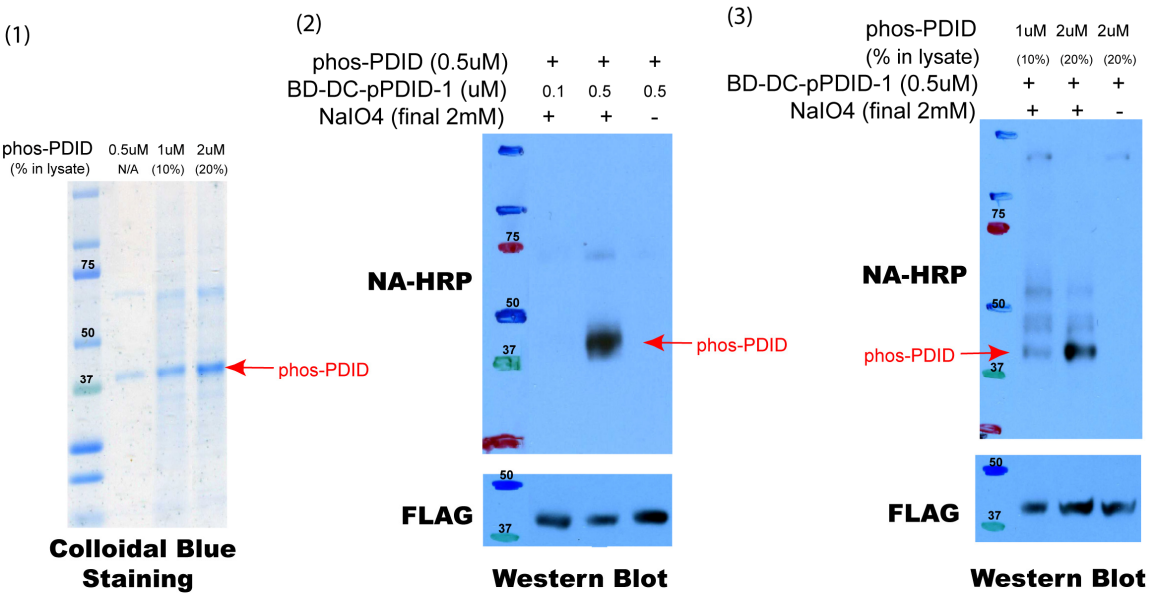
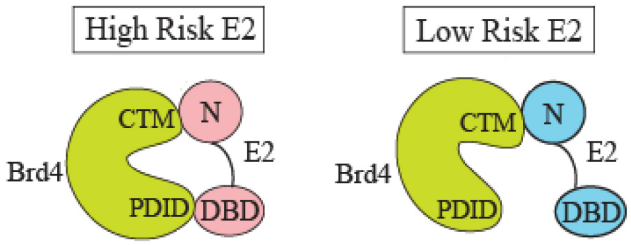
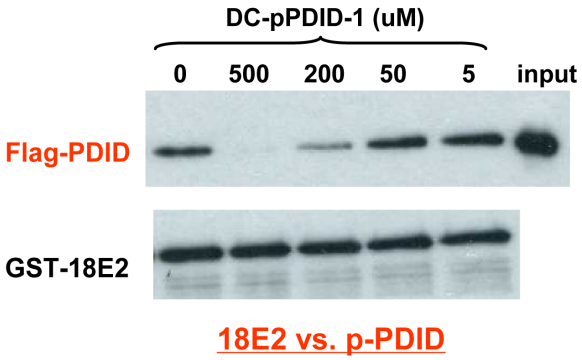


Figure 4-6 DC-pPDID-1 specifically inhibits the intereaction between phos-PDID and high risk HPV-18E2 in a dose-dependent manner, without affecting the binding between CTM of Brd4 and 18E2.

(A) Proposed model of intereaction between CTM and PDID domains of human Brd4 protein and high risk or low risk HPV E2.



(B) GST pull down experiment shown that Peptoid DC-pPDID-1 specifically inhibits the intereaction between phos-PDID and high risk HPV-18E2 in a dose-dependent manner, without affecting the binding CTM of Brd4 and 18E2.



(C) GST pull down experiment shown that Peptoid DC-pPDID-1 does not affect the binding between C-terminal motif (CTM) of Brd4 protein and HPV-18E2.

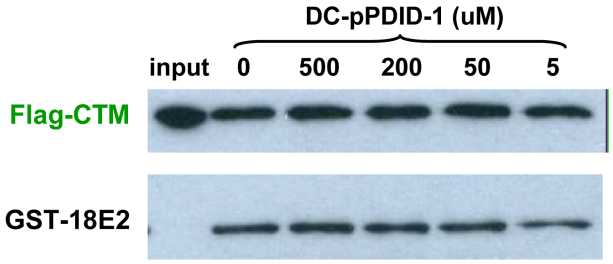


Figure 4-7 Phosphorylated-PDID-specifc peptoids can inhibit the Brd4-dependent transcription activation in mammalian cell culture.

Both DC-pPDID1 and DC-pPDID-2 can inhibit the Brd4-dependent transcription activation in C33A cells. 100uM of both peptoids were transfected into cells using Chariot reagent, followed by co-transfection of reporter gene and 10ng or 50ng 18E2. Firefly luciferase activity was measured 36 hours after transfection.

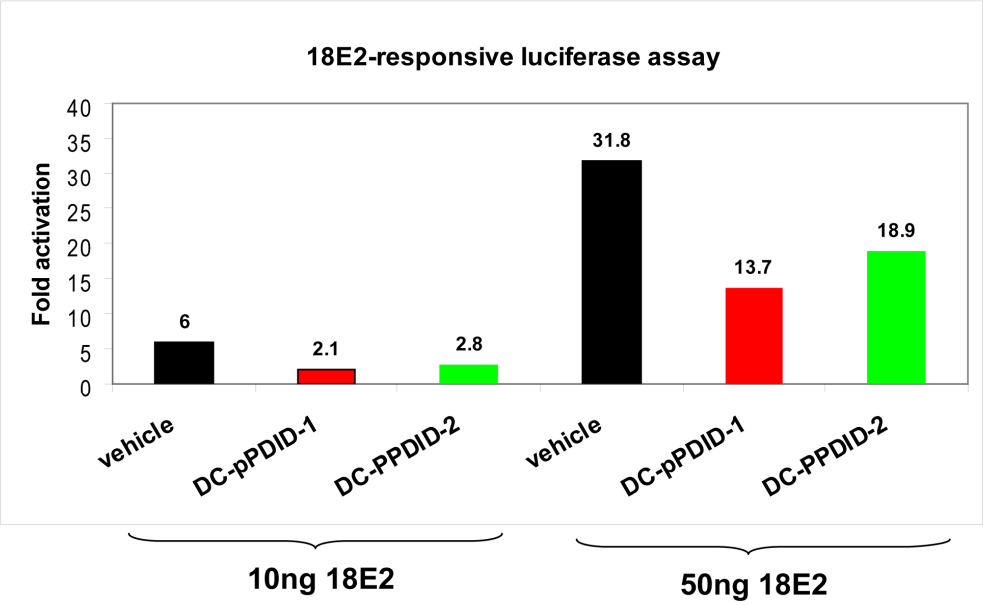
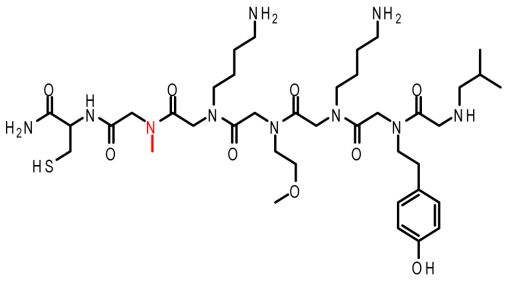
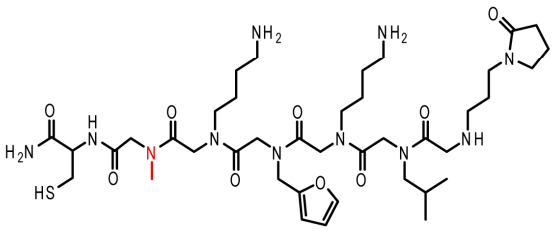


Figure 4-8 Identification of the pharmacophore of phos-PDID-specific peptoids.

(A) Mutation strategy to identify the minimum pharmacophore of phos-PDID-specific peptoids.



Sar G-1:	Cys-Sar-Nlys-Nmea-Nlys-Ntyr-Nleu
Sar G-2:	Cys-Nlys-Sar-Nmea-Nlys-Ntyr-Nleu
Sar G-3:	Cys-Nlys-Nlys-Sar-Nlys-Ntyr-Nleu
Sar G-4:	Cys-Nlys-Nlys-Nmea-Sar-Ntyr-Nleu
Sar G-5:	Cys-Nlys-Nlys-Nmea-Nlys-Sar-Nleu
Sar G-6:	Cys-Nlys-Nlys-Nmea-Nlys-Ntyr-Sar



Sar I-1:	Cys-Sar-Nlys-Nffa-Nlys-Nleu-Napp
Sar I-2:	Cys-Nlys-Sar-Nffa-Nlys-Nleu-Napp
Sar I-3:	Cys-Nlys-Nlys-Sar-Nlys-Nleu-Napp
Sar I-4:	Cys-Nlys-Nlys-Nffa-Sar-Nleu-Napp
Sar I-5:	Cys-Nlys-Nlys-Nffa-Nlys-Sar-Napp
Sar I-6:	Cys-Nlys-Nlys-Nffa-Nlys-Nleu-Sar

(B) Pull-down ability of various sarcosine-scanned derivatives of Hit G and Hit I peptoids in the on-bead binding assay with sf9 lyaste that overexpresses phos-PDID.

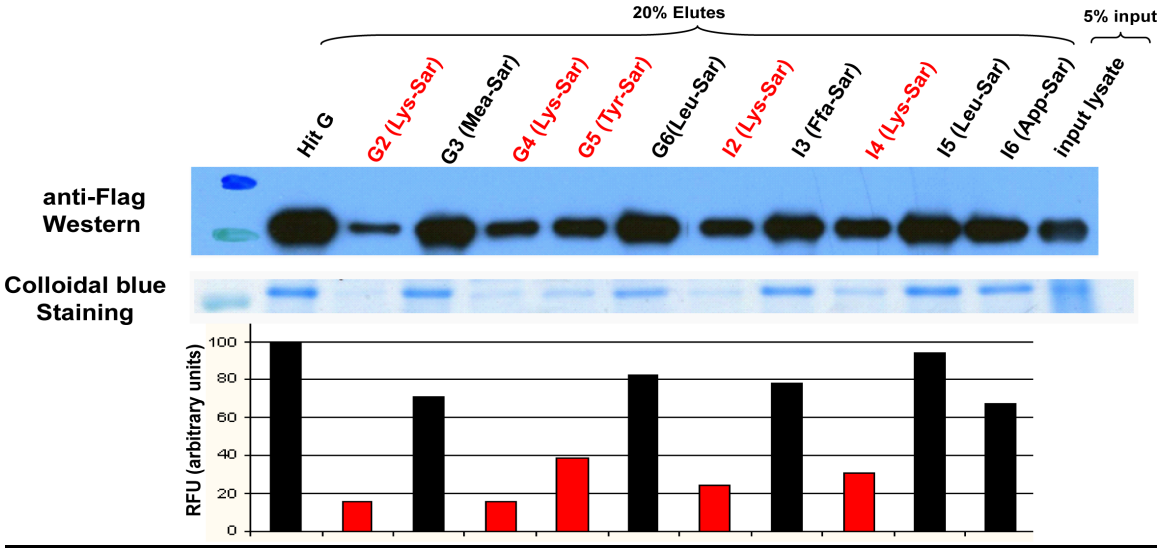
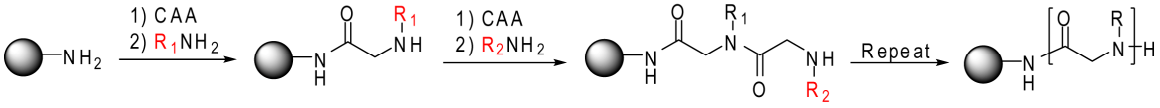
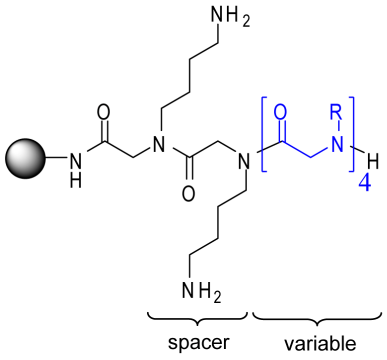


Figure 4-9 Synthesis scheme for peptoid library and chemical structure of the amines used.

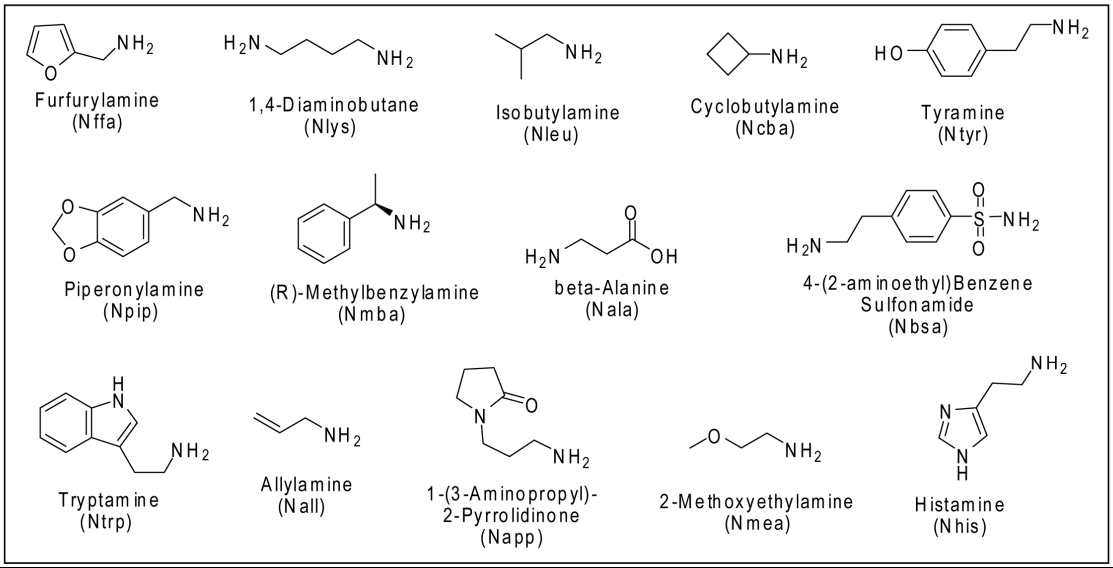
(A) Library synthesis scheme.



(B) Chemical structure of peptoids in the synthesized library.



(C) Amine monomers used for the library construction.



References

1. Lonard, D.M. and B.W. O'Malley, *The expanding cosmos of nuclear receptor coactivators.* Cell, 2006. **125**(3): p. 411-4.
2. Lonard, D.M. and W. O'Malley B, *Nuclear receptor coregulators: judges, juries, and executioners of cellular regulation.* Mol Cell, 2007. **27**(5): p. 691-700.
3. Picard, F., et al., *SRC-1 and TIF2 control energy balance between white and brown adipose tissues.* Cell, 2002. **111**(7): p. 931-41.
4. Lonard, D.M., R.B. Lanz, and B.W. O'Malley, *Nuclear receptor coregulators and human disease.* Endocr Rev, 2007. **28**(5): p. 575-87.
5. Bertolotti, M., et al., *Decreased hepatic expression of PPAR-gamma coactivator-1 in cholesterol cholelithiasis.* Eur J Clin Invest, 2006. **36**(3): p. 170-5.
6. Vianna, C.R., et al., *Hypomorphic mutation of PGC-1beta causes mitochondrial dysfunction and liver insulin resistance.* Cell Metab, 2006. **4**(6): p. 453-64.
7. Nawaz, Z., et al., *The Angelman syndrome-associated protein, E6-AP, is a coactivator for the nuclear hormone receptor superfamily.* Mol Cell Biol, 1999. **19**(2): p. 1182-9.
8. Roelfsema, J.H., et al., *Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease.* Am J Hum Genet, 2005. **76**(4): p. 572-80.
9. Anzick, S.L., et al., *AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer.* Science, 1997. **277**(5328): p. 965-8.
10. Chen, H., et al., *Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300.* Cell, 1997. **90**(3): p. 569-80.
11. Li, H., P.J. Gomes, and J.D. Chen, *RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC-1 and TIF2.* Proc Natl Acad Sci U S A, 1997. **94**(16): p. 8479-84.
12. Torchia, J., et al., *The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function.* Nature, 1997. **387**(6634): p. 677-84.
13. Huang, Z.J., I. Edery, and M. Rosbash, *PAS is a dimerization domain common to Drosophila period and several transcription factors.* Nature, 1993. **364**(6434): p. 259-62.
14. Spencer, T.E., et al., *Steroid receptor coactivator-1 is a histone acetyltransferase.* Nature, 1997. **389**(6647): p. 194-8.
15. Yan, J., S.Y. Tsai, and M.J. Tsai, *SRC-3/AIB1: transcriptional coactivator in oncogenesis.* Acta Pharmacol Sin, 2006. **27**(4): p. 387-94.
16. Han, S.J., D.M. Lonard, and B.W. O'Malley, *Multi-modulation of nuclear receptor coactivators through posttranslational modifications.* Trends Endocrinol Metab, 2009. **20**(1): p. 8-15.
17. Wu, R.C., et al., *Selective phosphorylations of the SRC-3/AIB1 coactivator integrate genomic reponses to multiple cellular signaling pathways.* Mol Cell, 2004. **15**(6): p. 937-49.

18. Osborne, C.K., et al., *Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer.* J Natl Cancer Inst, 2003. **95**(5): p. 353-61.
19. Gnanapragasam, V.J., et al., *Expression of RAC 3, a steroid hormone receptor co-activator in prostate cancer.* Br J Cancer, 2001. **85**(12): p. 1928-36.
20. Zhou, H.J., et al., *SRC-3 is required for prostate cancer cell proliferation and survival.* Cancer Res, 2005. **65**(17): p. 7976-83.
21. Bouras, T., M.C. Southey, and D.J. Venter, *Overexpression of the steroid receptor coactivator AIB1 in breast cancer correlates with the absence of estrogen and progesterone receptors and positivity for p53 and HER2/neu.* Cancer Res, 2001. **61**(3): p. 903-7.
22. Sakakura, C., et al., *Amplification and over-expression of the AIB1 nuclear receptor co-activator gene in primary gastric cancers.* Int J Cancer, 2000. **89**(3): p. 217-23.
23. Wang, Y., et al., *Prognostic significance of c-myc and AIB1 amplification in hepatocellular carcinoma. A broad survey using high-throughput tissue microarray.* Cancer, 2002. **95**(11): p. 2346-52.
24. Ghadimi, B.M., et al., *Specific chromosomal aberrations and amplification of the AIB1 nuclear receptor coactivator gene in pancreatic carcinomas.* Am J Pathol, 1999. **154**(2): p. 525-36.
25. Xu, F.P., et al., *SRC-3/AIB1 protein and gene amplification levels in human esophageal squamous cell carcinomas.* Cancer Lett, 2007. **245**(1-2): p. 69-74.
26. Torres-Arzayus, M.I., et al., *High tumor incidence and activation of the PI3K/AKT pathway in transgenic mice define AIB1 as an oncogene.* Cancer Cell, 2004. **6**(3): p. 263-74.
27. Azorsa, D.O., H.E. Cunliffe, and P.S. Meltzer, *Association of steroid receptor coactivator AIB1 with estrogen receptor-alpha in breast cancer cells.* Breast Cancer Res Treat, 2001. **70**(2): p. 89-101.
28. Planas-Silva, M.D., et al., *AIB1 enhances estrogen-dependent induction of cyclin D1 expression.* Cancer Res, 2001. **61**(10): p. 3858-62.
29. Shao, W., et al., *Coactivator AIB1 links estrogen receptor transcriptional activity and stability.* Proc Natl Acad Sci U S A, 2004. **101**(32): p. 11599-604.
30. Xu, J., et al., *The steroid receptor coactivator SRC-3 (p/CIP/RAC3/AIB1/ACTR/TRAM-1) is required for normal growth, puberty, female reproductive function, and mammary gland development.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6379-84.
31. Oh, A., et al., *The nuclear receptor coactivator AIB1 mediates insulin-like growth factor I-induced phenotypic changes in human breast cancer cells.* Cancer Res, 2004. **64**(22): p. 8299-308.
32. Lin, A. and M. Karin, *NF-kappaB in cancer: a marked target.* Semin Cancer Biol, 2003. **13**(2): p. 107-14.
33. Pages, F., et al., *Control of tumor development by intratumoral cytokines.* Immunol Lett, 1999. **68**(1): p. 135-9.

34. Fereshteh, M.P., et al., *The nuclear receptor coactivator amplified in breast cancer-1 is required for Neu (ErbB2/HER2) activation, signaling, and mammary tumorigenesis in mice.* Cancer Res, 2008. **68**(10): p. 3697-706.
35. Herbst, R.S., J.V. Heymach, and S.M. Lippman, *Lung cancer.* N Engl J Med, 2008. **359**(13): p. 1367-80.
36. Dihge, L., et al., *Epidermal growth factor receptor (EGFR) and the estrogen receptor modulator amplified in breast cancer (AIB1) for predicting clinical outcome after adjuvant tamoxifen in breast cancer.* Breast Cancer Res Treat, 2008. **109**(2): p. 255-62.
37. Lahusen, T., et al., *Epidermal growth factor receptor tyrosine phosphorylation and signaling controlled by a nuclear receptor coactivator, amplified in breast cancer 1.* Cancer Res, 2007. **67**(15): p. 7256-65.
38. Jemal, A., et al., *Cancer statistics, 2005.* CA Cancer J Clin, 2005. **55**(1): p. 10-30.
39. Gazdar, A.F. and J.D. Minna, *Deregulated EGFR signaling during lung cancer progression: mutations, amplicons, and autocrine loops.* Cancer Prev Res (Phila Pa), 2008. **1**(3): p. 156-60.
40. Zhou, B.B., et al., *Targeting ADAM-mediated ligand cleavage to inhibit HER3 and EGFR pathways in non-small cell lung cancer.* Cancer Cell, 2006. **10**(1): p. 39-50.
41. Morgillo, F., et al., *Implication of the insulin-like growth factor-IR pathway in the resistance of non-small cell lung cancer cells to treatment with gefitinib.* Clin Cancer Res, 2007. **13**(9): p. 2795-803.
42. Stabile, L.P., et al., *Combined targeting of the estrogen receptor and the epidermal growth factor receptor in non-small cell lung cancer shows enhanced antiproliferative effects.* Cancer Res, 2005. **65**(4): p. 1459-70.
43. Stabile, L.P., et al., *Human non-small cell lung tumors and cells derived from normal lung express both estrogen receptor alpha and beta and show biological responses to estrogen.* Cancer Res, 2002. **62**(7): p. 2141-50.
44. Marquez-Garban, D.C., et al., *Estrogen receptor signaling pathways in human non-small cell lung cancer.* Steroids, 2007. **72**(2): p. 135-43.
45. Mourtada-Maarabouni, M., et al., *Candidate tumor suppressor LUCA-15/RBM5/H37 modulates expression of apoptosis and cell cycle genes.* Exp Cell Res, 2006. **312**(10): p. 1745-52.
46. Weinstein, I.B. and A. Joe, *Oncogene addiction.* Cancer Res, 2008. **68**(9): p. 3077-80; discussion 3080.
47. Weinstein, I.B., *Cancer. Addiction to oncogenes--the Achilles heal of cancer.* Science, 2002. **297**(5578): p. 63-4.
48. Weinstein, I.B. and A.K. Joe, *Mechanisms of disease: Oncogene addiction--a rationale for molecular targeting in cancer therapy.* Nat Clin Pract Oncol, 2006. **3**(8): p. 448-57.
49. Felsher, D.W., *Tumor dormancy and oncogene addiction.* Apmis, 2008. **116**(7-8): p. 629-37.
50. Luo, J., N.L. Solimini, and S.J. Elledge, *Principles of cancer therapy: oncogene and non-oncogene addiction.* Cell, 2009. **136**(5): p. 823-37.

51. Kuang, S.Q., et al., *AIB1/SRC-3 deficiency affects insulin-like growth factor I signaling pathway and suppresses v-Ha-ras-induced breast cancer initiation and progression in mice.* Cancer Res, 2004. **64**(5): p. 1875-85.
52. Li, C., et al., *Essential phosphatases and a phospho-degron are critical for regulation of SRC-3/AIB1 coactivator function and turnover.* Mol Cell, 2008. **31**(6): p. 835-49.
53. Gandhi, J., et al., *Alterations in genes of the EGFR signaling pathway and their relationship to EGFR tyrosine kinase inhibitor sensitivity in lung cancer cell lines.* PLoS ONE, 2009. **4**(2): p. e4576.
54. Kajiro, M., et al., *The ubiquitin ligase CHIP acts as an upstream regulator of oncogenic pathways.* Nat Cell Biol, 2009. **11**(3): p. 312-9.
55. Shepherd, F.A., et al., *Erlotinib in previously treated non-small-cell lung cancer.* N Engl J Med, 2005. **353**(2): p. 123-32.
56. Moroni, M., et al., *Gene copy number for epidermal growth factor receptor (EGFR) and clinical response to antiEGFR treatment in colorectal cancer: a cohort study.* Lancet Oncol, 2005. **6**(5): p. 279-86.
57. Travis, W.D., *Pathology of pulmonary vasculitis.* Semin Respir Crit Care Med, 2004. **25**(5): p. 475-82.
58. Travis, W.D., et al., *Tumours of the lung*, in *Pathology and Genetics: Tumours of the Lung, Pleura, Thymus and Heart*, W.D. Travis, et al., Editors. 2004, International Agency for Research on Cancer (IARC): Lyon. p. 9-124.
59. Mountain, C.F., *Revisions in the International System for Staging Lung Cancer.* Chest, 1997. **111**(6): p. 1710-7.
60. Shigematsu, H., et al., *Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers.* J Natl Cancer Inst, 2005. **97**(5): p. 339-46.
61. Tang, X., et al., *EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients.* Cancer Res, 2005. **65**(17): p. 7568-72.
62. Pollack, J.R., et al., *Genome-wide analysis of DNA copy-number changes using cDNA microarrays.* Nat Genet, 1999. **23**(1): p. 41-6.
63. Kwei, K.A., et al., *Genomic profiling identifies TITF1 as a lineage-specific oncogene amplified in lung cancer.* Oncogene, 2008. **27**(25): p. 3635-40.
64. Tibshirani, R. and P. Wang, *Spatial smoothing and hot spot detection for CGH data using the fused lasso.* Biostatistics, 2008. **9**(1): p. 18-29.
65. Kersten, B., et al., *Plant phosphoproteomics: a long road ahead.* Proteomics, 2006. **6**(20): p. 5517-28.
66. Hunter, T., *Signaling--2000 and beyond.* Cell, 2000. **100**(1): p. 113-27.
67. Thingholm, T.E., O.N. Jensen, and M.R. Larsen, *Analytical strategies for phosphoproteomics.* Proteomics, 2009. **9**(6): p. 1451-68.
68. Tedford, N.C., et al., *Quantitative analysis of cell signaling and drug action via mass spectrometry-based systems level phosphoproteomics.* Proteomics, 2009. **9**(6): p. 1469-87.

69. Schmidt, S.R., F. Schweikart, and M.E. Andersson, *Current methods for phosphoprotein isolation and enrichment.* J Chromatogr B Analyt Technol Biomed Life Sci, 2007. **849**(1-2): p. 154-62.
70. Kinoshita, E., et al., *Novel immobilized zinc(II) affinity chromatography for phosphopeptides and phosphorylated proteins.* J Sep Sci, 2005. **28**(2): p. 155-62.
71. Andersson, L. and J. Porath, *Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography.* Anal Biochem, 1986. **154**(1): p. 250-4.
72. Czernik, A.J., et al., *Production of phosphorylation state-specific antibodies.* Methods Enzymol, 1991. **201**: p. 264-83.
73. Heffetz, D., M. Fridkin, and Y. Zick, *Antibodies directed against phosphothreonine residues as potent tools for studying protein phosphorylation.* Eur J Biochem, 1989. **182**(2): p. 343-8.
74. Paradela, A. and J.P. Albar, *Advances in the analysis of protein phosphorylation.* J Proteome Res, 2008. **7**(5): p. 1809-18.
75. Kodadek, T. and K. Bachhawat-Sikder, *Optimized protocols for the isolation of specific protein-binding peptides or peptoids from combinatorial libraries displayed on beads.* Mol Biosyst, 2006. **2**(1): p. 25-35.
76. Simon, R.J., et al., *Peptoids: a modular approach to drug discovery.* Proc Natl Acad Sci U S A, 1992. **89**(20): p. 9367-71.
77. Zuckermann, R.N., et al., *Efficient method for the preparation of peptoids [oligo(N-substituted glycines)] by submonomer solid-phase synthesis.* Journal of the American Chemical Society, 1992. **114**: p. 10646-7.
78. Olivos, H.J., K. Bachhawat-Sikder, and T. Kodadek, *Quantum dots as a visual aid for screening bead-bound combinatorial libraries.* Chembiochem, 2003. **4**(11): p. 1242-5.
79. Kodadek, T., et al., *Synthetic molecules as antibody replacements.* Acc Chem Res, 2004. **37**(9): p. 711-8.
80. Alluri, P.G., et al., *Isolation of protein ligands from large peptoid libraries.* J Am Chem Soc, 2003. **125**(46): p. 13995-4004.
81. Simpson, L.S., et al., *Selective Toxin Sequestrants for the Treatment of Bacterial Infections.* J Am Chem Soc, 2009.
82. Udugamasooriya, D.G., et al., *A peptoid "antibody surrogate" that antagonizes VEGF receptor 2 activity.* J Am Chem Soc, 2008. **130**(17): p. 5744-52.
83. Xiao, X., et al., *A cell-permeable synthetic transcription factor mimic.* Angew Chem Int Ed Engl, 2007. **46**(16): p. 2865-8.
84. Lim, H.S., C.T. Archer, and T. Kodadek, *Identification of a peptoid inhibitor of the proteasome 19S regulatory particle.* J Am Chem Soc, 2007. **129**(25): p. 7750-1.
85. Alluri, P., et al., *Isolation and characterization of coactivator-binding peptoids from a combinatorial library.* Mol Biosyst, 2006. **2**(11): p. 568-79.
86. Lim, H.S., et al., *Periodate-triggered cross-linking reveals Sug2/Rpt4 as the molecular target of a peptoid inhibitor of the 19S proteasome regulatory particle.* J Am Chem Soc, 2007. **129**(43): p. 12936-7.
87. Burdine, L., et al., *Periodate-triggered cross-linking of DOPA-containing peptide-protein complexes.* J Am Chem Soc, 2004. **126**(37): p. 11442-3.

88. Burzio, L.A. and J.H. Waite, *Cross-linking in adhesive quinoproteins: studies with model decapeptides.* Biochemistry, 2000. **39**(36): p. 11147-53.
89. Liu, B., L. Burdine, and T. Kodadek, *Chemistry of periodate-mediated cross-linking of 3,4-dihydroxylphenylalanine-containing molecules to proteins.* J Am Chem Soc, 2006. **128**(47): p. 15228-35.
90. Li, W.W., J. Heinze, and W. Haehnel, *Site-specific binding of quinones to proteins through thiol addition and addition-elimination reactions.* J Am Chem Soc, 2005. **127**(17): p. 6140-1.
91. Archer, C.T., L. Burdine, and T. Kodadek, *Identification of Gal4 activation domain-binding proteins in the 26S proteasome by periodate-triggered cross-linking.* Mol Biosyst, 2005. **1**(5-6): p. 366-72.
92. Umanah, G.K.E., et al., *Cross-Linking of a DOPA-Containing Peptide Ligand into Its G Protein-Coupled Receptor.* Biochemistry, 2009. **48**(9): p. 2033-2044.
93. Liu, B., et al., *Label transfer chemistry for the characterization of protein-protein interactions.* J Am Chem Soc, 2007. **129**(41): p. 12348-9.
94. Wu, S.Y. and C.M. Chiang, *The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation.* J Biol Chem, 2007. **282**(18): p. 13141-5.
95. Muchardt, C., et al., *The hbrm and BRG-1 proteins, components of the human SNF/SWI complex, are phosphorylated and excluded from the condensed chromosomes during mitosis.* Embo J, 1996. **15**(13): p. 3394-402.
96. Dey, A., et al., *The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 8758-63.
97. Lee, A.Y. and C.M. Chiang, *Chromatin adaptor Brd4 modulates E2 transcription activity and protein stability.* J Biol Chem, 2009. **284**(5): p. 2778-86.
98. Lee, A.Y. and C.M. Chiang, *Functional regulation of chromatin adaptor Brd4 by phosphorylation.* manuscript in preparation, 2009.
99. Paillisson, A., et al., *Bromodomain testis-specific protein is expressed in mouse oocyte and evolves faster than its ubiquitously expressed paralogs BRD2, -3, and -4.* Genomics, 2007. **89**(2): p. 215-23.
100. Sawa, C., et al., *Bromodomain factor 1 (Bdf1) is phosphorylated by protein kinase CK2.* Mol Cell Biol, 2004. **24**(11): p. 4734-42.
101. Olivos, H.J., et al., *Microwave-assisted solid-phase synthesis of peptoids.* Org Lett, 2002. **4**(23): p. 4057-9.
102. Lim, H.S., et al., *Rapid identification of the pharmacophore in a peptoid inhibitor of the proteasome regulatory particle.* Chem Commun (Camb), 2008(9): p. 1064-6.
103. Yu, P., B. Liu, and T. Kodadek, *A high-throughput assay for assessing the cell permeability of combinatorial libraries.* Nat Biotechnol, 2005. **23**(6): p. 746-51.
104. Wu, S.Y., et al., *Isolation of mouse TFIID and functional characterization of TBP and TFIID in mediating estrogen receptor and chromatin transcription.* J Biol Chem, 1999. **274**(33): p. 23480-90.

105. Chiang, C.M. and R.G. Roeder, *Expression and purification of general transcription factors by FLAG epitope-tagging and peptide elution.* Pept Res, 1993. **6**(2): p. 62-4.
106. Wu, S.Y. and C.M. Chiang, *TATA-binding protein-associated factors enhance the recruitment of RNA polymerase II by transcriptional activators.* J Biol Chem, 2001. **276**(36): p. 34235-43.