# Therapeutic targeting of ETS rearranged cancers

**by**

**John Chad Brenner**

**A dissertation submitted in partial fulfillment**
**of the requirements for the degree of**
**Doctor of Philosophy**
**(Cellular and Molecular Biology)**
**in The University of Michigan**
**2012**

**Doctoral Committee:**

**Professor Arul M. Chinnaiyan, Chair**
**Professor Thomas E. Carey**
**Professor Sofia D. Merajver**
**Associate Professor David O. Ferguson**
**Associate Professor Thomas E. Wilson**

**© John Chad Brenner**
**2012**

## Acknowledgements

This thesis has come together as a product of many people working together. In this brief section, I have the opportunity to list some of the contributions made by my friends and co-workers to develop these projects, but there is no way in which I can fully thank everyone who helped with these projects. I would like to thank the many members of the Chinnaiyan lab, especially George X. Wang, Bushra Ateeq and Nallasivam Palanisamy.

Several students have helped get much of this work done including Sonam Patel, Hallie Liang, Azalea Ayuningtyas, Shanker Menon, Laura Bou-Maroun, Yang Sheng and Sid Goyal. I hope that they all learned from me as much as I learned from each of them. Additionally, Jill Granger, Yong Li and Guoqing Hou all worked extremely hard as technicians to help with many of the experiments on a daily basis. I am extremely grateful for their patience and dedication. Xuhong Cao helped me tremendously in getting experiments done quickly. Without her, I am certain that this thesis would not have been completed. And Karen Giles who helped immensely with everything from understanding the boss to writing, submitting and reviewing papers. I thank Margie R. Conlon for her support throughout many years. I also need to especially thank all of the graduate students in Arul's lab that made coming in on the weekends so much fun including Matthew K. Iyer, John R. Prensner and O. Alejandro Balbin.

Felix Y. Feng has generously helped with many different aspects of this thesis, grant writing and management skills. I am very grateful for his support. The members of his lab have also helped finish several difficult experiments in a very rapid manner including Sumin Han, Meilan Liu and Kari Wilder-Romans. They have all truly made it possible to do way more than I ever thought could be accomplished. I thank Scott A. Tomlins and Daniel R. Rhodes for their support and willingness to discuss anything. Scott especially helped me make through the 2.5 years during which the ERG-PARP manuscript was in review.

Arul M. Chinnaiyan has made this entire experience possible, he fostered my development and put me in a position to achieve success. For this, I cannot thank him enough. Likewise, my entire committee including Sofia D. Merajver, Thomas E. Carey, Thomas E. Wilson, David O. Ferguson were all willing to meet with me to discuss everything from science to career advice. These discussions were all extremely helpful for me. Tom Carey has been a mentor and friend since I was in high school – I owe a large part of my passion for science to him. I am also certain that I would not have been awarded the Department of Defense pre-doctoral award without the strong support of Sofia Merajver. I thank my preliminary exam committee for their advice early in my graduate school career including Thomas E. Carey,

Kenneth J. Pienta, John K. Kim and Diane M. Robbins. I also thank Thomas A. Look, Fred W. Alt, Adolfo A. Ferrando, Christopher A. Maher and Hamilton O. Smith for their advice during my graduate career. I thank the CMB program and administration for their help throughout graduate school.

I want to thank many of the faculty members at UM who supported me throughout this thesis despite not being on my committee including Carol R. Bradford, Mark E. Prince, Max S. Wicha, Kenneth J. Pienta, Maha H. Hussain, Shaomeng Wang, Andrzej A. Dlugosz and formerly, Marisol Soengas. Collectively they have made doing research at UM an exciting experience and have helped translate the findings of this thesis into the clinic. I thank Nora M. Navone for teaching us how to study primary prostate cancer serial xenografts in the lab; these have been a tremendous addition to our arsenal of prostate cancer xenograft models. I want to thank Lori J. Pierce for her support of a project to study PARP inhibitors with radiation in breast cancer. I also want to thank the support of Howard R. Soule and Jonathan W. Simons at the Prostate Cancer Foundation (PCF). They have greatly expanded my network of collaborators and also generously allowed me to become the first ever graduate student to win a young investigator award from the PCF – I am very grateful for this honor.

Finally, I want to thank my family and friends for their support throughout the entire process. I especially thank my girlfriend Angela Liang who puts up with my high affinity for lab and work, even at night and on the weekends. I have missed or been late to many family functions in order to try to get experiments done in the lab. My family and friends have always been understanding and let me get by without too much trouble. I feel like I have the best job in the world and I am very grateful that they support me in my dedication to science.

# Table of Contents

Acknowledgements……ii

List of Figures…….……..vii

List of Tables……...x

List of Appendices…….…xi

Abstract…….xii

Chapter

I. Introduction: Rationale for Therapeutic Targeting of ETS Transcription Factors in Cancer……..1

INTRODUCTION……..1
THE ETS FAMILY OF TRANSCRIPTION FACTORS……..4
WHEN TO TARGET ETS TRANSCRIPTION FACTORS ……...20
CLINICAL CONSIDERATIONS……26
APPROACHES TO INHIBIT TRANSCRIPTION FACTORS……...33
CONCLUSION OF INTRODUCTION……...41
FIGURES……..43
REFERENCES…….45

II. PARP1 and DNA-PKcs are required for ETS-mediated invasion and metastasis in Prostate Cancer……..50

SUMMARY……..50
INTRODUCTION……51
MATERIALS AND METHODS……..51
RESULTS……..63
DISCUSSION……...74
CONCLUSION…….81
FIGURES……..82
TABLE……..99
REFERENCES……100

III. Mechanistic Rationale for Poly(ADP)-Ribose Polymerase inhibition in ETS Gene Fusion Positive Prostate Cancer……103

SUMMARY……103
INTRODUCTION……...104
MATERIALS AND METHODS……105
RESULTS……116
DISCUSSION……..124
FIGURES……128
REFERENCES……138

IV. PARP-1 inhibition as a targeted strategy to treat Ewing's sarcoma……139

SUMMARY……139
INTRODUCTION……...140
MATERIALS AND METHODS……142
RESULTS……151
DISCUSSION……156
FIGURES……157
REFERENCES……167

V. Development of peptidomimetic inhibitors of the ERG transcription factor in prostate cancer……168

SUMMARY……168
INTRODUCTION……...169
MATERIALS AND METHODS……170
RESULTS……188
DISCUSSION……200
FIGURES……202
TABLE……218
REFERENCES……219

VI. Discussion: Mechanistically-Informed Personalized Therapy for Transcription Factor Driven Disease……220

SUMMARY……220
IMPLICATIONS OF EMERGING STRUCTURAL CONCEPTS……..221
IMPORTANCE OF ETS INDUCED DNA DAMAGE……..……228
DEVELOPING EIPs AS CLINICAL THERAPEUTICS……...234
CLINICAL IMPLICATIONS OF PARP AND DNA-PKcs INHIBITORS……..……237

A MOVE TO PERSONALIZED MEDICINE..……..248
CONCLUSION……251
REFERENCES……254

Appendices……...……257

## List of Figures

1.1: Schematic representation of conserved domains in ETS proteins and ETS gene fusion protein products……..43

1.2: Schematic representation of strategies to inhibit the *TMPRSS2-ERG* transcription factor signaling axis……...44

2.1: Schematic representation of mass spectrometry workflow used for identification of ERG interacting proteins……82

2.2: The *TMPRSS-ERG* gene fusion product interacts with PARP1 and the DNA-PK complex……..83

2.3: ERG interacts with DNA-PKcs, PARP1, Ku70 and Ku80…….85

2.4: Several ETS proteins interact with DNA-PKcs directly, ERG requires Y373……...87

2.5: DNA-PKcs and PARP1 associate with ERG regulated genomic loci in VCaP cells and are required for ERG-mediated transcription……..89

2.6: PARP1 and DNA-PKcs are required for ERG-regulated transcription……..90

2.7: Quantification of knockdown in RWPE and VCaP models……...92

2.8: Quantification of knockdown in RWPE-ETV1, PC3 and RWPE-SLC45A3-BRAF models……....93

2.9: ERG-mediated invasion requires engagement of PARP1 and DNA-PKcs……..94

2.10: WST chemosensitivity curves for the prostate cancer cell lines used for both *in vitro* and *in vivo* experiments. ……...95

2.11: ERG-mediated invasion, intravasation and metastasis requires PARP1 activity *in vivo*……96

2.12: PARP1 and DNA-PKcs expression are required for ERG-mediated transcription of *EZH2 in vivo*……..98

3.1: Establishment of PC3-ERG model cell line and *in vivo* effects of Olaparib and characterization of primary prostate cancer serial xenograft model……128

3.2: Inhibition of PARP1 alters ETS positive, but not ETS negative cell line xenograft growth……..130

3.3: Response of primary prostate cancer serial xenograft model to Olaparib and effects of Olaparib on mouse body weight in all xenograft models……131

3.4: ETS transcription factors induce DNA damage which is potentiated by PARP inhibition……..133

3.5: ETS gene fusion proteins induce γH2A.X and 53BP1 foci formation…….135

3.6: ETS proteins for not alter homologous recombination efficiency in prostate cells..137

4.1: Ewing's sarcoma gene fusion products (EWS-FLI1 and EWS-ERG) interact with PARP1 and DNA-PKcs, and confer sensitivity to PARP1 inhibition…….157

4.2: Ewing sarcoma gene fusion products confer increased DNA damage and cell invasion, which can be modulated by genetic or pharmacologic approaches…….158

4.3: DNA breaks are potentiated by Olaparib in Ewing's sarcoma cell lines and by overexpression of EWS-FLI1 in PC3 cells……..159

4.4: Confirmation of knockdown…….160

4.5: FLI1 siRNA replicates EWS siRNA phenotype in EWS-FLI1 rearranged cells….161

4.6: Depletion of PARP1 potentiates the relative number of DNA double strand breaks in Ewing's Sarcoma…….162

4.7: Olaparib chemosensitivity of Ewing's cell lines……..163

4.8: PARP inhibition prevents EWS-FLI1 positive xenograft growth and metastasis…164

4.9: EWS-FLI1 maintains a feed forward loop that drives PARP1 expression……...165

4.10: EWS-FLI1 knockdown causes a decrease in PARP1 mRNA expression……..166

5.1: Identification and characterization of ERG inhibitory peptides……...202

5.2: Expression of recombinant proteins…….204

5.3: Point mutations in ETS domain abolished *TMPRSS2:ERG*-induced invasion and DNA damage……...205

5.4: Analysis of ERG ETS DNA binding domain structure……206

5.5: ERG binding peptides specifically inhibit ERG activities……...207

5.6: ELISA-based transcription factor binding assay……..209

5.7: EIP blocks ERG-mediated DNA damage and cell invasion…….210

5.8: EIPs block ERG-driven DNA damage…….211

5.9: Retroinverso EIPs block ERG-mediated tumor invasion and metastasis *in vivo*….212

5.10: RI-peptides retain the specific anti-ERG properties……...213

5.11: Retroinverso EIPs suppress tumor growth *in vivo*……..215

5.12: VCaP xenografts show sustained therapeutic response following RI-EIP…….217

## List of Tables

2.1: Summarized immunoprecipitation mass spectrometry data……...99

5.1: Sequences of 59 phage clones sequenced following final bio-pan……...218

## List of Appendices

I: Oligonucleotide primers…….257

II: Genes Differentially Regulated by PARP1 or DNA-PKcs siRNA……..……259

# ABSTRACT

## Therapeutic targeting of ETS rearranged cancers

A rate limiting step in the realization of personalized medicine is the ability to create viable therapeutic strategies for specific molecular lesions. Among the unique types of molecular lesions that are currently considered "undruggable" is a class of proteins called transcription factors, which broadly serve as regulators of gene expression. Recurrent fusions of ETS transcription factors are driving events in a diverse array of cancers including prostate cancer (PCa) and Ewing's sarcoma family of tumors (ESFTs). Here, I show that the protein products of ETS gene fusions from both PCa (*TMPRSS2-ERG*) and ESFT (*EWS-ERG* and *EWS-FLI1*) interact with the enzymes poly(ADP-ribose) polymerase1 (PARP1) and DNA protein kinase catalytic subunit (DNA-PKcs). ETS gene-mediated transcription, invasion and metastasis require PARP1 and DNA-PKcs expression and activity. Importantly, pharmacological inhibition of PARP1 inhibits ETS-positive, but not ETS-negative, PCa and ESFT xenograft growth. Addition of a PARP inhibitor to the second-line ESFT chemotherapeutic agent temozolomide resulted in complete and sustained responses of all treated tumors in one ETS positive xenograft model. Mechanistically, ETS fusion overexpression induces DNA damage, which is potentiated by PARP1 inhibition in a manner similar to that of *BRCA1/2* deficiency. Notably, EWS-FLI1 fusion proteins acted in a positive feedback

loop to maintain the expression of PARP1, which was required for EWS-FLI-mediated transcription, thereby enforcing oncogene-dependent sensitivity to PARP-1 inhibition. Because PARP inhibition inhibits ERG indirectly, we further sought to inhibit ERG directly. We identified a consensus ERG-binding peptide that disrupts the ERG:DNA-PKcs interaction from a phage display library. X-ray crystallography confirmed the interaction residues are accessible and led to the first structure of an ETS protein without DNA. ERG inhibitory peptides were found to efficiently disrupt ERG-mediated transcription, DNA damage and cell invasion as well as suppress tumor growth, intravasation and metastasis. Taken together, this work creates a viable therapeutic strategy for targeting transcription factors – historically considered "undruggable" – based on mechanistic rationale and a strategy that can be used to build therapeutics against other transcription factors. As such, this thesis continues to build on the work of many generations of researchers towards the ultimate goal of personalized medicine.

# Chapter I

# Introduction: Rationale for Therapeutic Targeting of ETS Transcription Factors in Cancer

## INTRODUCTION

Throughout history, technological advances are often followed by discoveries that dramatically alter our perceptions of both disease etiology and therapeutic course. In the mid-1840's, German pathologists began using techniques to compare gross mitotic changes in tissue sections of various human malignancies and, soon after, chromosomes were discovered (*1*). Almost half of a century later, based on the observation that sea urchin embryos would frequently engage in uncommon development following mitotic abnormality, Theodore Boveri published a critical hypothesis that, "mammalian tumors might be initiated by mitotic abnormalities that resulted in a change in the number of chromosomes in the cell (aneuploidy)" *translated* (*2*). As time passed, breakthroughs continued to arise that increased the quality and reproducibility of cytogenetic techniques and, as a result, the understanding of disease pathogenesis became increasingly clear. For example, new cytogenetic technologies allowed Lejeune to associate Down syndrome with an extra copy of chromosome 21 (*3, 4*).

Advances in technology once again spurred discovery in 1958 when Rothfels and Siminovitch published a new cytogenetic, air-drying technique for flattening chromosomes (*5*). This technology allowed Hungerford and Nowell to observe that two patients with chronic myelogenous leukemia (CML) had a characteristic small chromosome (*6*) and subsequently report that a series of seven CML patients all harbored the same abnormality (*7*), coined the "Philadelphia chromosome" after the city in which the abnormal chromosome was discovered (*8*). This "disruptive event" – genomic event with pro-malignant consequences (*9*)– is a translocation between chromosomes 9 and 22 (*10*) that fuses the breakpoint cluster region (*BCR*) gene on chromosome 22 with the v-abl Abelson murine leukemia viral oncogene homolog (*ABL1*) gene on chromosome 9 (*11*). The resulting *BCR-ABL1* gene fusion represents one of the first cancer specific molecular alterations to be identified and technological advances soon demonstrated that *BCR-ABL1* is also an active kinase critical for disease development. In 1990, Lugo *et al.* demonstrated that RAT1 cells transfected with either *BCR-ABL1* or *v-src*, but not *v-ras* or *v-myc*, had a significant increase in total phosphotyrosine content (*12*). Ensuing studies led to the identification of a small molecule Bcr-Abl1 inhibitor called Imatinib (*13, 14*), which has been used to great initial clinical success in patients with chronic-phase CML (*15, 16*). As such, advances in technology led to the rationale design of targeted therapeutics for the treatment of CMLs harboring the *BCR-ABL1* translocation. The example represents the first true case of genetically defined modern medicine. Until this point, few other molecularly-directed therapies have shared the success of

Imatinib and some of these will be described below. In the future, the hope is that new technologies will enable the development of targeted therapies tailored for additional classes of disruptive events.

**Discovery of *TMPRSS2-ETS* gene fusions in prostate cancer.**

In 2005, a bioinformatics approach led to the discovery of class of gene fusions that occurs in about 50% of all prostate cancers (PCas) (*17*). In this study, Tomlins *et. al.* analyzed gene expression data by cancer outlier profile analysis (COPA) – a method that ranks genes by normalizing expression values based on median absolute deviation of gene expression to accentuate outlier profiles (reviewed in (*18*)) – and identified several cancer related outlier genes including those of the ETS family, *ERG* and *ETV1* (*17*). Given that homologues of this gene family were known to be rearranged in Ewing's sarcoma, the group postulated that the mechanism of ETS gene overexpression in prostate cancer was similarly due to genetic rearrangement. Subsequent exon-walking quantitative PCR demonstrated loss of the 5' exons in both *ERG* and *ETV1* and 5'-RNA ligase-mediated rapid amplification of cDNA ends (5'-RACE) was used to identify the 5' fusion partner, *TMPRSS2* (transmembrane protease, serine 2) (*17*). Following the discovery, much excitement was generated about the possibility of molecularly targeted therapy for such a large population of patients. However, in contrast to *BCR-ABL1* gene fusions that encode a chimeric kinase, ETS gene fusions encode a nearly full length ETS transcription factor (*19*) and, as a class of proteins, transcription factors have largely been considered "undruggable" (*20*).

The challenge of this thesis was to break that paradigm in a clinically relevant manner using diseases driven by ETS transcription factors as the primary model systems. Furthermore, I also aimed to create a generalized methodology to develop targeted therapies for similar classes of disruptive events in the future.

## THE ETS FAMILY OF TRANSCRIPTION FACTORS

In order to therapeutically target ETS transcription factors it is critical to understand divergence amongst the protein family, its conserved protein domains, as well as the role of ETS genes in development, normal biology and disease pathogenesis. In humans, the ETS (v-ets erythroblastosis virus E26 oncogene) family of transcription factors consists of 28 evolutionarily related genes in humans that drive unique transcriptional programs by binding to the specific DNA recognition sequence 5'-GGA(A/T)-3' and recruiting transcriptional machinery (reviewed in (*21*)). Here, I will review what is known about ETS transcription factors with a focus on the ETS genes most commonly rearranged in prostate cancer, *ERG*, and in Ewing's sarcoma, *FLI1*.

### Avian Erythroleukemia Virus, E26, and the first ETS gene

In 1962, a group of Bulgarian scientists isolated a retrovirus from an avian leukosis that had the unusual property of inducing both myeloid and erythroid leukemias (*22*). The retroviral isolate was named E26. Over time much interest was generated in the properties of the E26 retrovirus because it was found to induce a more aggressive phenotype than most other retroviruses including the

Avian Erythroblastosis virus (AEV), which only induced erythroid leukemia (*23*). In fact, the E26 virus contained properties that had previously been thought to be specific to other viruses such as the ability to transform myeloblasts *in vivo* and *in vitro*, a property that was previously specific to Avian Myeloblastosis virus (AMV) (*24-26*). As such, accumulating phenotypic data suggested that the E26 virus harbored a unique and more aggressive isolate than many other avian retroviruses.

With advances in molecular cloning technology in the late 1970s, analysis of genes from the E26 virus was performed and revealed the presence of two oncogenes *myb*, which had been identified in AMV(*26*), and an unidentified sequence which was named *v-ETS* for ***E***26 ***T***ransformation-specific ***S***equence (*27-29*). The viral sequence was then used as a probe to identify homologs of *v-ETS* in chickens (*27*) and other species (*30, 31*). Consequently studies of avian leukosis led to the discovery of the first ETS genes, *v-ETS* and chicken *ETS1*.

**Evolution of the ETS gene family**

After the initial gene discovery, sequence analysis and Southern blot techniques led to a rapid expansion in the number of ETS genes as defined by the presence of a conserved "ETS domain". Subsequent phylogenetic analysis suggested that genes containing an ETS domain can be consistently traced back from the 28 human genes to 8 and 10 genes in *Drosophila melanogaster* and *Caenorhabditus elegans*, respectively, to a single gene in *metazoan phyla* (*32, 33*). Beyond the ETS domain, this gene family frequently contains a second

conserved domain called PNT (discussed below) and some individual genes contain additional unique domains to provide functional diversity. For example, *GABPA* contains a CBP/EP300 binding domain in its N-terminus that is not contained in other ETS genes (*34*) and *ELK1*, *ELK3* and *ELK4* contain serum response factor (SRF) recruitment domains called B-box domains that are unique to this subfamily of ETS genes (*35*). The remaining regions of the ETS genes encode for divergent stretches of protein, which may be uniquely acetylated, glycosylated, phosphorylated, sumoylated and/or ubiquitinated to provide unique properties to the ETS protein. These unique sequences are thought to facilitate the recruitment of either unique co-factors to the ETS DNA recognition site or factors that modulate subtle changes to the distribution of the ETS transcription factor in the genome. The structures of the 28 unique ETS proteins can be seen schematically in **Figure 1.1A**.

Given the subtle differences between ETS protein sequences and structures, and the fact that in any given tissue contains 14-25 expressed ETS genes (*36, 37*), it has been postulated that ETS genes all have unique functional roles at normal expression levels. Interestingly, it was outlier overexpression of *ETV1* and *ERG* that led to the identification of ETS fusions in prostate cancer. As such, it has been further postulated that relative levels of ETS gene expression may have additional unique functional consequences for this class of transcription factors (*38*).

**ETS DNA binding domain**

All 28 human ETS genes and all prostate cancer and Ewing's sarcoma derived ETS gene fusions contain an ETS DNA binding domain. Up to this point, information about the ETS DNA binding domain has been derived from X-ray crystal structures of 10 ETS proteins all complexed with DNA (reviewed in (*38*)). Prior to this thesis, in which the 11$^{th}$ ETS DNA binding domain was crystallized from ERG, a crystal structure of an ETS transcription factor without DNA had not been reported. Thus, the analysis presented here finally allows for discrimination of the ETS domain structure before and after DNA binding. However, the first 10 structures did give detailed information about the DNA binding domain. The main structure of this domain is referred to as a winged-helix architecture and contains three conserved α-helices and a 4-stranded, anti-parallel β-sheet. The protein interacts directly with DNA through two ARG residues that fit in the major groove of a 12-15 base pair sequence, focused around the 5'-GGA(A/T)-3' sequence motif. Interaction with this DNA sequence is mediated by hydrogen bonding with the two guanine residues. Surprisingly, in any of the solved structures there are no other direct bonding interactions with base pairs.

This observation is surprising because some ETS genes can bind to subtly different DNA sequences. Several groups have used SELEX (systematic evolution of ligands by exponential enrichment), protein binding microarrays or microwell binding assays to identify the unique binding "logos" – conserved DNA binding sequences – for each ETS protein (reviewed in (*38*)). These studies concluded that all ETS proteins bind to the consensus 5'-GGA(A/T)-3'; however,

some proteins may alter their specificity based on a longer recognition sequence. For example, murine PU.1 was shown to have a 400-fold variation in binding affinity based on variations in DNA sequence flanking the consensus motif (*39*). To support this observation, it has been postulated that both the "turn" residues and the "wing" residues between strands H2 and H3 as well as between S3, S4 and H1, respectively, provide water-mediated hydrogen bonding, hydrophobic and electrostatic interactions that impart individualized DNA sequence specificity for each ETS protein (*39*). Taken together, some of the uniqueness among ETS transcription factors is derived from minor sequence and structural variation within the conserved ETS domain.

**Pointed (PNT) domain**

In contrast to the ETS DNA binding domain which was defined following interest in the oncogenic capacity of *v-ETS*, the pointed (PNT) domain is the second most conserved domain found in ETS genes and was discovered independently. In fact, the pointed domain received its name from a mutational screen following the embryonic development of *drosophila melanogaster* in which a larval mutant formed a "pointed" head rather than the normal rounded head (*40*). Several years after the phenotypic characterization, the mutant gene was cloned and called pointed (*pnt*) (*41*); ironically, it turned out to be an ETS gene.

PNT domains (also known as SPM, HLH or SAM domains) have been found in more than 250 regulatory protein classes including receptor tyrosine

kinases, serine/threonine kinases, adapter proteins and transcription factors (*42-44*). Of the 28 human ETS genes, only 11 contain a PNT domain and of the ETS gene fusions identified in prostate cancer only 2 (*ERG* and *FLI1*) of 6 contain the PNT domain (*ETV1*, *ETV4*, *ETV5* and *ELK4* do not have a PNT domain). Likewise, while the majority of *TMPRSS2-ERG* gene fusions encode nearly full length ERG protein, a few are missing the PNT domain (*45*), yet retain the ability to induce proliferation, motility and invasion (*46*). Therefore, the genetic evidence from fusion variants found in prostate cancer strongly suggests that the PNT domain is at least partially dispensable for oncogenic activity of ETS transcription factors in this disease. Similarly, the PNT domain of FLI1 is lost during rearrangement with EWS in Ewing's sarcoma.

Despite the fact that a few ETS fusions lose the PNT domain in prostate cancer, evidence exists that the presence of the PNT domain may alter ETS transcription factor specificity in general. This domain forms a helix-loop-helix (HLH) structure which can dimerize with other proteins and in the case of the ETS PNT domain appears to specifically self-associate (*47*). Additionally, the MAP kinase ERK can bind directly to the pointed domain of some ETS transcription factors and phosphorylate threonine and serine residues in the PNT domain (*41, 48*). As such, it is possible that some ETS proteins and fusions can modulate or perhaps redistribute MAPK activity. However, this has not been explored in cancers with ETS fusions. Of note, we also recently characterized *ETS2* as a tumor suppressor in prostate cancer frequently lost through interstitial deletion during *TMPRSS2-ERG* rearrangement or through loss-of-function

mutations. Because *ETS2* is one of the ETS genes to contain a PNT domain, one hypothesis for its tumor suppressive effects is that it disrupts ERG activity through dimerization with ERG, but this hypothesis has not yet been explored in detail.

**ETS genes in endothelial cell development and angiogenesis**

As ETS genes appear to segregate into unique functional subclasses based on minor sequence and structural variations, studies of normal ETS gene function have focused on de-convoluting the complex interplay between the many genes. Indeed, a screen of ETS gene expression demonstrated that at least 19 different genes are simultaneously expressed in normal endothelial cells (*37*) and the effects of modulating the expression of many of these ETS genes has been previously reviewed (*49*). Here, I will focus on the roles of ERG and FLI1.

Both of these proteins are highly expressed in both adult and embryonic endothelial cells (*50*) and have been postulated to be involved in several aspects of vascular development. In human umbilical vein endothelial cells, which have high ERG expression, ERG siRNA results in downregulation of several genes (including *THBS1*, *VWF*, *RHOA* and *ICAM2*) involved in various vascularization processes as well as loss of cell-to-cell contact and subsequent endothelial cell death (*51, 52*). In mice harboring a germline mutation in ERG that causes defective hematopoiesis, early vascular development appeared normal, but blood vessels were dilated (*53*). In contrast, knockdown of ERG using morpholinos in zebrafish caused hemorrhaging (*54*) indicating that ERG may have a more critical

role in zebrafish than in mice. Alternatively, the mouse model retained expression of the mutant ERG allele indicating that this specific mutation may not affect ERG's role in vascularization. Although these results are conflicting, they indicate that other ETS genes cannot fully compensate for loss of ERG expression to rescue vascularization in some models.

As with ERG, FLI1 is also important for endothelial cell development. FLI1 knockdowns show reduced expression of the vascular markers Kdr and Aplnr in *Xenopus* experiments (*54*), FLI1 morpholinos display loss of trunk circulation and hemorrhaging in the head in zebrafish (*54, 55*), and FLI1 knockouts are embryonic lethal (E12.5) due to hemorrhaging (*56*). Despite these phenotypes, in the knockout mice, specification of endothelial lineage and vascular patterning were reported to be normal (*56*) suggesting that vascular development can be partially compensated by other ETS transcription factors in mice. However, the role of FLI1 in adult angiogenesis remains to be demonstrated. Given the data for these two factors, anti-ERG and anti-FLI1 therapies need to be aware of potential effects on adult vasculature.

**The role of ERG and FLI1 in Hematopoiesis**

As mentioned, a mutational screen for defective hematopoiesis recently found that homozygous mutations in ERG are embryonic lethal between E10.5 and E13.5 (*53, 57*). Analysis of heterozygotes from this screen found that single *ERG* mutants had few committed hematopoietic cells, which were unable to compete with wild type cells in transplantation assays. Interestingly, subsequent

analysis demonstrated that *ERG* is not essential for hematopoietic specification or differentiation, but rather *ERG* is only required for fetal hematopoietic stem cell maintenance (*58, 59*), a critical point for potential anti-ERG therapies in adults. Despite this observation, the role of *ERG* in hematopoiesis has not been fully explored and remains an area of active research.

In contrast to the effect of *ERG* disruption in transgenic mouse models, *FLI1* knockout mice have defective megakaryocytic development (*60, 61*). Consistent with this observation, patients that have Paris-Tousseau Syndrome (also Jacobsen Syndrome), which frequently have dysmegakaryopoiesis, have hemizygous loss of *FLI1* (*60, 62*) suggesting that *FLI1* is critical for megakaryocytic development. Indeed, retroviral mediated overexpression of *FLI1* in K562 cells has been shown to induce megakaryocytic differentiation (*63*) and to drive expression of many megakaryocytic-specific genes such as *GPIX* and *GPIbα* (*64*).

A recent study assessed the role of both *ERG* and *FLI1* in hematopoiesis (*57*) to assess potential complementation effects of these two closely related ETS genes. The group found that mice with disruptive heterozygous mutations in both *ERG* and *FLI1* had more pronounced phenotypes with both loss of megakaryocyte number and hematopoietic stem cell homeostasis than single mutants (*57*). Consequently, ETS targeted therapies that disrupt both ERG and FLI1 transcriptional activity will need to monitor potential effects on the hematopoietic compartment.

**ERG in Leukemia**

In addition to the normal roles of ERG in vascularization and hematopoiesis, ERG has also been indicated to have critical roles in several different cancers. In T-cell adult lymphoblastic leukemia (T-ALL), high level ERG expression has been associated with poor prognosis (*65*). Amplification of the *ERG* locus has also been observed in a mouse model of T-cell lymphoma (*66*). Intriguingly, *ERG* deletions have also been reported in T-ALL suggesting unique molecular subtypes of the disease (*67*). Interestingly, most acute myeloid leukemias (AMLs) have high *ERG* expression (*68, 69*) including those that appear to be cytogenetically normal (*70, 71*). This can occur by a variety of mechanisms. In some cases this is due to loss of the *ERG*-regulating microRNAs, miR-196a/b (*72*), and in others it is due to a gene fusion that puts the *ERF4* 5'UTR and a minimal *ERF4* coding sequence in frame with nearly full length ERG coding sequence (*73*). Likewise, some low frequency ETS fusions observed in AML put the RNA binding domain of the *FUS* gene in frame with the ETS DNA binding domain of *ERG* (*74*).

Given the high expression of ERG in T-ALL and AML as well as its role in fetal hematopoiesis, ERG could be important for either maintenance of leukemogenesis along the myeloid lineage. Indeed, shRNA experiments have demonstrated that loss of ERG expression causes a loss in proliferation and (*75, 76*) ectopic expression of ERG in fetal hematopoietic progenitors was shown to promote megakaryopoiesis and to induce rapid-onset leukemia when transplanted into sub-lethally irradiated syngeneic mice (*76*). Similarly, forced expression of

ERG in adult bone marrow cells caused expansion of T-cell and erythroid cell lineages and induction of spontaneous *NOTCH1* mutations leading to T-cell ALL in a transplantation model (*75*). Taken together, emerging evidence suggests that ERG overexpression in hematopoietic cells can initiate both T-ALL and AML and may be required for cell maintenance even in the presence of other disruptive events such as *NOTCH* mutations.

**ETS gene fusions in prostate cancer**

In PSA (Prostate Specific Antigen) screened cohorts, approximately half of all prostate cancers harbor recurrent gene fusions of the androgen-regulated gene, *TMPRSS2*, to the oncogenic *ETS* transcription factor *ERG* (*19*). While *ERG* contributes to approximately 90% of ETS family gene rearrangements in prostate cancer, other *ETS* transcription factors including *ETV1* (*17*), *ETV4* (*77*), *ETV5* (*78*), *ELK4* (*79, 80*) and *FLI1* (*81*) also participate in prostate cancer rearrangements albeit at a lower frequency. The structures of the various protein products encoded by ETS fusions in prostate cancer are shown in **Figure 1.1B**.

In contrast to the ETS gene family which constitutes the most frequent 3' partners, several different 5' partners have been identified including those with androgen-responsive promoters (*TMPRSS2*, *SLC45A3* , *KLK2*, *HERV-K_22q11.23* and *CANT1)*, one with an androgen-insensitive promoter (*DDX5*), one with a constitutively active promoter (*HNRPA2B1)*, and, in a single case, one with an androgen-repressed promoter (*C15orf21)* (*17, 77, 78, 82-84*). In a recent IHC expression analysis study, we demonstrated that ERG protein expression was extremely high regardless of the 5' fusion partner as *TMPRSS2*, *SLC45A3* and

*NDRG1* to *ERG* gene fusions all expressed a similar amount of ERG protein (*85*). This is also supported by our observation that *TMPRSS2-ERG* overexpression drives feed-forward regulation of wild type *ERG* in human PCa (*86*). The most common *TMPRSS2-ERG* gene fusion variants involve *TMPRSS2* exon 1 or 2 fused to *ERG* exon 2, 3, 4, or 5 (*17, 87-95*) and less frequently rearrangements of *TMPRSS2* exon 4 or 5 fused to *ERG* exon 4 or 5 (*93*). As such, the most prevalent *TMPRSS2-ERG* gene fusions encode for an Erg protein product is missing only 4 N-terminal amino acids as compared to wild type protein and, like *TMPRSS2* (*96-98*), the *TMPRSS2-ERG* gene fusion is androgen regulated (*17*). As with the 5' fusion partner, ERG protein expression is also independent of the *TMPRSS2-ERG* fusion mRNA variant (*85*).

*ETS* gene fusions occur early in prostate cancer progression during the transition from high grade prostatic intraepithelial neoplasia (PIN) lesions to invasive carcinoma (*99*). Consistent with this observation, we demonstrated that ERG IHC staining was observed exclusively in prostate cancer and adjacent HG PIN (*85*). Two rearrangement mechanisms have been described for their formation (reviewed in (*100*)). Because the *TMPRSS2* and *ERG* genes are located 3MB apart on chromosome 21q, rearrangement can occur by either interstial deletion (*90, 94, 101, 102*) or by interchromosomal insertion (*103*). Recently, several groups demonstrated that androgen stimulation causes an induced chromosomal proximity of the *TMPRSS2* and *ERG* genomic loci in cell lines (*86, 104, 105*), which can lead to the formation of *TMPRSS2-ERG* rearrangements at low frequency (*104, 105*). Importantly, with the addition of genomic stress such

as ionizing radiation, the frequency of fusion formation increases dramatically (*86, 105*). Lin *et. al.* postulated that endogenous endonucleases including activation induced cytidine deaminase (AID) and the protein encoded from LINE-1 ORF2 are responsible for creating AR-directed DNA double strand breaks that are required for site-specific genomic rearrangements (*105*). In line with this hypothesis, studies of translocations following site-directed DNA double strand breaks have indicated that gene fusions preferentially utilize actively transcribed genes (*106, 107*).

Subsequent to gene fusion formation, ETS gene fusions have a clear role in prostate cancer pathogenesis. Genetically-engineered mice expressing *ERG* or *ETV1* under androgen regulation exhibit PIN-like lesions, but do not develop frank carcinoma (*18, 108-112*) and overexpression of *ERG* leads to accelerated carcinogenesis in mouse prostates with deletion of the tumor suppressor *PTEN* (*110, 111*). Additionally, in a transplant model, mouse prostate epithelial cells that are forced to overexpress both *ERG* and the androgen receptor (*AR*) gene form invasive prostate cancer (*112*). This suggests that *ERG* accelerates prostate carcinogenesis. Furthermore, because these models do not develop androgen-dependent disease, the data suggests that either a fundamental difference between mice and men exists – especially considering that wild type mice do not develop prostate cancer – or that we do not yet have a complete understanding of all of the disruptive events that occur prior to metastatic progression in ETS positive tumors.

Functionally, ETS gene expression drives several malignant phenotypes and is critical for the survival of cancer cells. In VCaP cells, which were derived from a vertebral metastatic PCa lesion and naturally harbor the *TMPRSS2-ERG* rearrangement, *TMPRSS2-ERG* targeting siRNA inhibits the ETS transcriptional program, cell growth, invasion, metastasis and xenograft tumor growth (*46, 113*). Likewise, forced overexpression of truncated ERG or other ETS genes commonly rearranged in PCa such as *ETV1*, *ETV4* and *ETV5* in benign immortalized or non-transformed prostate epithelial cells drives a unique ETS transcriptional program, which in turn drives cell invasion and metastasis (*17, 77, 78*). Taken together, ETS gene fusions represent a prostate cancer-specific genetic lesion that has promise as a target for molecularly tailored therapy.

**ETS fusions in Ewing's Sarcoma**

Ewing's sarcoma family of tumors (ESFTs) are highly malignant tumors of the bone and soft tissue that occur in children and adults (*114*). ESFTs are characterized by the presence of EWS-FLI1 (95%) or EWS-ERG (5%) gene fusions. Protein structure is shown schematically in **Figure 1.1B**. The resulting chimeric proteins are active ETS transcription factors that drive downstream signaling pathways. In these rearrangements, the pointed domain of the ETS protein is replaced by the n-terminal portion of an RNA binding domain from the *EWS* gene that is thought to enhance post-transcriptional splicing of ETS target genes (*115*).

*EWS-FLI1* or *EWS-ERG* expression in ESFTs disrupts a number of pathways that drive a variety of oncogenic phenotypes and some of these have been suggested to depend on a direct interaction with RNA helicase A (RHA) (reviewed in (*116*)). Consequently, ESFT gene fusion and RHA expression is important for the maintenance of a unique transcriptional program in ESFT (*117*), some components of which are required for transformation such as *NKX2.2* (*118*). Importantly, disruption of *EWS-FLI1* expression using shRNA inhibits ESFT cell line growth *in vitro* and *in vivo* xenograft models suggesting that *EWS-FLI1* is critical for the maintenance of ESFT growth (*119-121*) similar to *TMPRSS2-ERG* in PCa.

In contrast to PCa, which has a clear cell-of-origin, the cells from which Ewing's sarcoma derives are still unknown. Together with the observation that *EWS-FLI1* only transforms a small subset of primary cells, transgenic mouse models of Ewing's sarcoma have been limited and are difficult to interpret. For example, *EWS-FLI1* overexpression has been shown to induce leukemic phenotypes when expressed in hematoepoetic stem cells (*122, 123*) and to induce skeletal disruption when expressed in mesenchymal progenitors using a *PRX1* promoter (*124*). Importantly, however, *EWS-FLI1* expression accelerates tumor formation in conjunction with *TP53* deletion (*124*). As such *EWS-ETS* fusions may require additional collaborating mutations to initiate transformation, but this will remain unknown until a cell-of-origin is defined for the disease. In the future, high throughput sequencing of ESFTs will help determine the prevalence of

additional disruptive events and subsequent mechanistic studies will determine the contribution of these events to ESFT pathogenesis.

**ETV6 gene fusions**

A final class of ETS gene fusion that has been discovered also represented the first recurrent epithelial rearrangement to be discovered (reviewed in (*125*)). These fusions occur in secretory breast carcinoma, a rare subtype of infiltrating ductal carcinoma that affects both children and adults (*126*). The fusion gene contains the 5' portion of an ETS gene called *ETV6* (also *TEL*) containing the PNT domain, but not the ETS DNA binding domain, fused to a receptor tyrosine kinase called *NTRK3* (*127*). Interestingly, *ETV6* is also rearranged in several hematologic malignancies with a transcription factor called *AML1* (*128*) as well as to protein tyrosine kinase domains from *ABL* (*129, 130*), *JAK2* (*131-133*), *ARG* (*134, 135*), *PDGFRβ* (*136*) and *FGFR3* (*137*) each of which define unique sub-types of leukemia (reviewed in (*138*)). However, as these gene fusions all appear to use the PNT dimerization domain to primarily activate kinase domain activity, they will not be discussed further as a model of therapeutic inhibition of transcription factors.

**Conclusions for the ETS family of transcription factors**

As a model for the development of transcription factor based therapeutics, the ETS family of proteins is one of the most promising as it is frequently deregulated in cancer. In all cases studied thus far, ETS fusion gene

overexpression is essential for maintaining many of the malignant properties of cancers harboring an ETS aberration. ETS genes were among the first viral oncogenes to be discovered. In cancer, ETS gene fusions come in two classes: ones that utilize the PNT dimerization domain to drive constitutive fusion partner kinase activity and those that retain the ETS DNA binding domain to drive a unique transcriptional program. Among the second class of cancer fusions, ERG is the most frequently rearranged in human cancer. ERG is known to have a role in early hematopoietic development and endothelial cell maintenance as demonstrated in both mouse knockout and zebrafish morpholino studies. Finally, the ETS genes rearranged in prostate cancer and ESFTs contain a highly conserved ETS DNA binding sequence that is slightly divergent from other ETS proteins (**Figure 1.1C**).

## WHEN TO TARGET ETS TRANSCRIPTION FACTORS

While ETS gene fusions represent an obvious molecular target in several cancers, in the quest for "personalized medicine" – in which therapies are selected for a patient based on their unique profile of disruptive events, it is important to remember that cancers are driven by multiple disruptive events. Consequently, even directed therapies need to be administered under consideration of the global mutational state of the patient's cancer cell population. With the recent advances in high throughput sequencing, we have entered an era when disruptive events can be rapidly profiled. Here, I will discuss additional abnormalities that have been

described in prostate cancer as well as the implications of these events for personalized medicine and directed molecular therapies.

**The emerging mutational landscape of PCa**

The following was adapted from (*9*).

The number of genomic changes observed in cancer is rapidly expanding due to the emergence of enabling technologies. In prostate cancer, the most common somatic alterations are ETS rearrangements (*17*). However, another well-established set of molecular alterations in this disease are mutations in the tumor suppressor gene *PTEN*, which lead to activation of phosphoinositide 3-kinase (PI3K)/AKT pathways (*139*) and cooperate with *ETS* fusions in prostate carcinogenesis (*111, 140*). While many prostate cancers can be characterized by *ETS* and/or *PTEN* status, additional disruptive events are being identified in smaller subsets of disease which exemplify the enormous genomic complexity of prostate cancer.

With the goal of identifying novel somatic events in prostate cancer, Berger *et. al.* recently completed whole genome sequencing of seven primary prostate cancers (three harboring *ETS* rearrangements) and their matched normal controls (*141*). This led to the identification (and in some cases re-confirmation) of mutations in several genes including *SPTA1*, *SPOP*, *ZNF407*, *CHD1, CHD5, HDAC9, DICER* and *PTEN*. As is often the case with genomic changes that drive cancer progression, it is common to find functionally recurrent mutations which disrupt multiple genes in a pathway. Specifically, rearrangements disrupting both

*PTEN* and its interacting protein *MAGI2* were identified. In the future, knockdown experiments may confirm that loss of MAGI2 expression drives AKT phosphorylation, suggesting that susceptibility pathways can be mutated at different points, expanding the number of mutations known to disrupt PTEN signaling in prostate cancer.

Similarly, Ding and colleagues recently compared the gene expression profiles of *PTEN* deletion-induced PIN to wild type prostate epithelium (*142*). This led to the identification of a difference in expression of genes in the TGFβ signaling pathway. Subsequent prioritization of targets in this pathway led to the identification of SMAD4 as a key regulator of TGFβ signaling that is downregulated in human prostate cancer metastasis as compared to localized prostate cancer. Importantly, prostate-specific deletion of both *SMAD4* and *PTEN* led to faster occurring prostate cancer with a high propensity for metastasis while *SMAD4* deletion alone had no effect. This suggested that, before *PTEN*-impaired prostate tumors become metastatic, they must first develop mechanisms to disrupt the tumor suppressive effects of SMAD4-mediated canonical TGFβ signaling. It will be interesting to see if genomic sequencing of metastatic prostate cancer reveals evidence of TGFβ pathway disruption.

Although Berger *et. al.* did not identify events disrupting TGFβ signaling, their study provides insight into the mechanism of how gene fusions are formed, which has implications for detecting disruptive events in personalized medicine. For example, many of the rearrangements occurred in a balanced manner such that reciprocal genomic rearrangements are generated – creating a series of many

different gene fusions in which no DNA copy number changes were identified (*141*). This study also found that a single gene could be disrupted by a different mechanism of rearrangement in each tumor. For example, they identified rearrangements that occur in approximately 6% of prostate cancers in a cell adhesion gene called *CADM2* and each rearrangement occurred by a different combination of genomic deletions, duplications and inversions. Because of this, traditional assays such as FISH and IHC may not be appropriate for clinical detection of alterations in this gene if it is to be used for determining clinical course. Furthermore, this suggests that while the type of rearrangement may be different, a conserved mechanism must be responsible for creating a rearrangement "hotspot".

By overlapping the breakpoint locations with available genome-wide location analyses for androgen receptor (AR), ERG and histone marks (*143*), Berger et al. demonstrated that the breakpoints correlated with open chromatin marks as well as AR binding in tumors with ETS rearrangements. This is surprising as the most common ETS gene fusion product, ERG, functions to disrupt AR signaling (*143*). It will be interesting to see if other genomic events correlated with *ETS* status, such as chromosome 3p14 deletion, are also correlated with enrichment of these factors (*144*). Nonetheless, this observation supports recent mechanistic data suggesting that activated AR facilitates genomic rearrangements by bringing linearly distant genomic loci together in a process termed induced-proximity (Multiple studies reviewed in (*100*)). The fact that

androgen receptor was enriched at genomic breakpoints also suggests that these "hotspots" may be tissue type specific.

Based on the success of whole genome sequencing, we recently performed exome based profiling to further decipher the mutational landscape of castrate resistant prostate cancer. The study represents a large cohort of heavily pre-treated patients with lethal metastatic disease and provides insights into treatment options to overcome resistance mechanisms that evolve in hormone refractory tumors. This study is unique from previous sequencing studies that focus on localized disease and is critical for understanding how ERG inhibitors may function in the castrate resistant setting. For example, most patients with CRPC in our cohort have high level *AR* amplification and continued activation of AR signaling through *TMPRSS2-ERG*, suggesting therapy with novel anti-androgens may continue to be effective (androgen signaling is often maintained in androgen independent cancers).

Interestingly, one mechanism leading to androgen independence in our cohort is the activation of AR transcriptional co-activators, where we show that interacting AR cofactors, including MLL2, MLL, ASXL1, UTX and FOXA1, are somatically mutated in CRPC. As our data suggests that the MLL methyltransferase core complex is altered in approximately 10% of CRPC and required for AR signaling, this may represent a novel epigenetic target for CRPC, for which drugs are in early development to treat acute leukemias (*145*). Likewise, through targeted resequencing of localized prostate cancer and CRPC, Taylor *et al*. recently identified recurrent aberrations in AR activators, *EP300*,

*NCOA1*, *NCOA2*, *PXN*, *SOX9*, and *TNK2*, and AR inhibitors, *NRIP1*, *NCOR1*, and *NCOR2 (144)*. Detailed studies need to be performed to determine the functional consequences to AR signaling of the large set of newly identified mutant/amplified co-factors, and it will be important to understand how second and third generation AR antagonists such as MVD3100 function in these settings. In some cases, it is possible that AR cofactor mutation will bypass the effects of AR antagonists, in which case downstream AR target genes such as the *TMPRSS2-ERG* gene fusion product would serve as a more rational therapeutic target.

Given the enormous complexity of disruptive events in a prostate cancer genome, an important question is how can this information be used clinically? Consistent with the importance of androgen signaling in prostate cancer, we identified unexpected deregulation of a number of chromatin/histone modifiers in CRPC that physically interact with AR, but have also been identified as being significantly mutated in multiple tumor types (*146-152*)). This is of note because it supports the call for sequencing-based therapy in modern medicine. As some mutations may or may not be derived independent of histological subtype ("hotspot" hypothesis), in some cases, targeted therapies will be utilized for non-exclusive histological subtypes. Despite the early success of personalized medicine-based strategies, future studies will be needed to understand the temporal development of individual aberrations during cancer development and the implications to clinical course (including therapy and resistance mechanisms).

Taken together, the recent reports by Berger, Ding, and colleagues as well as our recent exome sequencing studies highlight the complex nature of disruptive events in the life of prostate cancer. As advanced sequencing approaches become more widely implemented, it is certain that additional genetic alterations along key progression pathways will be identified. Understanding the genesis and effect of these events, relative to existing lesions such as *PTEN* inactivation and *ETS* fusions, will be critical to efforts to develop better biomarker-based predictors of progression and to identify potential targets for prostate cancer therapy. While the discovery of key genetic alterations, such as *PTEN* inactivation, *ETS* and *RAF* fusions, *SPOP* mutations, *MAGI2* rearrangements, AR co-factor mutation and *SMAD4* silencing have represented significant strides toward better understanding of prostate cancer progression, it is clear that we are seeing just the beginning of cancer-defining disruptive events, but are also at the beginning of an exciting period of discovery that will help to define personalized medicine.

## CLINICAL CONSIDERATIONS

While the future of cancer therapy will be intimately linked with sequencing-based detection of disruptive events, it will also be integrated with the current standard of care regimens regardless of whether or not these therapies include targeted anti-neoplastic agents. As such, approaches to develop novel therapeutics should be mindful of current treatment paradigms for the model disease. Likewise, it is important to understand how potential biomarkers can be detected if they are to direct clinical treatment protocols. Here, I will review

methods of detection for ETS gene fusions in prostate cancer as well as the current treatment protocols for both prostate cancer (localized and castrate resistant) and Ewing's sarcoma as they are the primary disease models used for this thesis. Therapeutics discussion will include the implications on ETS status if known.

**Clinical detection of *TMPRSS2-ERG* positive cancers**

Given the importance of *TMPRSS2-ERG* expression as a potential determinant of therapy, it is important to understand the molecular methods that can be used to type patients for potential clinical trials. Thus far, four screening methods have been established to type patients for ETS status: 1) Fluorescence in situ hybridization (FISH) for *ERG* split on tissue sections (*17*), 2) IHC analysis of ERG protein expression in tissue sections (*85*), 3) *TMPRSS2-ERG* mRNA detection from urine (*153*) and, 4) sequencing of DNA/mRNA from patient tissue (*154*). These methods each have limitations, but are all viable strategies for integration with clinical trials for real-time biomarker-driven patient selection.

Of note, it was demonstrated that urine detection of *TMPRSS2-ERG* in combination with prostate cancer antigen 3 (PSA3) enhances the utility of serum PSA for identifying clinically relevant cancer on biopsy (*153*). Given its clinical utility, ETS status may be available for all prostate cancer patients at the time of diagnosis making it the ideal biomarker for determining therapeutic course.

**Prostate Cancer Therapy**

For prostate cancer, treatment course is a complex decision that relies on multiple variables ranging from patient preference, to clinical stage and tumor grade, to serum PSA (prostate specific antigen) level. Likewise, as the disease usually occurs in older men, co-morbidities can be a major issue in determining treatment course. Generally, however, PCa therapy can be divided into three categories: 1) clinically localized cases, 2) advanced and recurrent disease, and finally 3) castrate-resistant disease. From these three categories, I will assess the standard chemotherapies that are utilized to understand how anti-ETS therapies may be best designed for the different "treatment stages" of disease.

PSA screen-detected prostate localized cancers can be small, well-differentiated tumors and are generally thought to have a low risk of progression. For these cancers, active surveillance – or a postponement of immediate therapy – is often recommended, especially in men with high associated co-morbidities. Clinical trials are ongoing to determine the survival benefit. In some low risk disease, as well as intermediate risk disease, the recommended treatments include radical prostatectomy (removal of the prostate), radiation therapy through either external beam or brachytherapy (implants) (*155*). Other experimental approaches focus on killing tumor cells by freezing (cryotherapy) or by focused heating using high intensity ultrasound. Anti-androgen therapy (or androgen ablation) is not currently recommended for low or intermediate risk localized disease by either the AUA or the NCCN following a study of 19,271 men (*156*) that suggested only subsets of men benefited from the therapy. In fact, the therapy was associated

with deleterious effects on quality of life without an improvement in overall mortality (*157-159*). Indeed, these studies highlight a need to determine early biomarkers of progression and response, especially for patients with intermediate risk disease.

High risk localized disease as well as metastatic disease requires more aggressive chemotherapeutic approaches in addition to both radical prostatectomy and radiation therapy to manage the primary tumor. In this setting, androgen ablation is highly recommended and often performed by orchiectomy and medical castration using a gonadotropin releasing hormone agonist. Upon PSA recurrence or disease progression, secondary therapies include additional anti-androgens (discussed below), docetaxel plus prednisone (*160, 161*), or cabazitaxel (a second generation taxane derivative) (*162*). Importantly, the majority of these therapies are not targeted. However, anti-ETS therapies intended to be administered in this disease setting should be tested in combination with and against androgen ablation or anti-androgens that would be expected to reduce androgen-regulated ETS fusion expression.

**Anti-androgen therapy in castrate resistant prostate cancer**

While 80-90% of men with high risk or metastatic prostate cancer receive and generally respond well to androgen deprivation therapy, nearly all men eventually relapse with castrate-resistant disease. The nomenclature is confusing because androgen signaling is frequently restored in these cancers, which also restores the *TMPRSS2-ERG* axis (*163*). Though it should be noted that recent

debates have suggested *TMPRSS2-ERG* expression is restored by mechanisms independent of the androgen axis in castrate resistant disease (AACR 2012 Prostate cancer meeting). Consequently, therapeutic inhibition of the androgen signaling axis may remain a viable option for some ETS positive CR-PCa patients and several second and third generation AR antagonists have been developed for this purpose. The three most promising second generation anti-androgen therapies are abiraterone, MVD3100 and ARN-509. Thus, because *TMPRSS2-ERG* expression may still be modulated by androgen signaling in this setting, it is important to understand how ETS-targeted therapies work in the context of these anti-androgen therapies.

Abiraterone acetate was discovered through small molecule screening of derivative compounds to identify small molecules capable of irreversible binding to cytochrome P450 17 alpha-hydroxysteroid dehydrogenase (CYP17), an enzyme that can produce testosterone and dihydrotestosterone in the absence of androgen (*164-166*). Clinically, abiraterone in combination with prednisone was shown increase mean overall survival by approximately 4 months in a phase III trial CR-PCa patients previously treated with a docetaxel-containing regimen as compared to prednisone and placebo. Statistically significant improvements were also reported for time to PSA progression, progression-free survival and PSA response (*167*) as well as palliation of pain due to bone metastasis (*168, 169*). Recently, a study was completed to determine if *TMRSS2-ERG* status can be used as a biomarker to determine sensitivity to abiraterone, however, only 41 patients were evaluated and a statistical outcome could not be made (*170*).

Another AR antagonist, MVD3100, is also being developed as an androgen receptor inhibitor. In contrast to abiraterone, MVD3100 was not designed to inhibit the synthesis of androgen by Cyp17. Instead, it inhibits cells with amplified androgen receptor by binding to the surface of the receptor and sequestering it to the cytoplasm, thereby, preventing its ability to activate transcription. Two phase III trials are currently comparing the effects of MVD3100 to placebo in CR-PCa who also received prior docetaxel-based regimens and preliminary data has demonstrated a survival advantage for these patients (*171*). The second trial is in men with chemotherapy naïve disease and is still in the process of accruing patients (*172*).

More recently, ARN-509 has emerged as an even more potent agonist of AR than MVD3100 as it was able to achieve a reduction in xenograft growth at approximately 1/3 the dose of MVD3100 (*173*). As with the previous drugs, however, the effects on *TMPRSS2-ERG* expression need to be further evaluated in the castrate resistant setting.

**Clinical treatment of ESFTs**

In contrast to prostate cancers, which occur mainly in older men and can remain latent for many years, a majority of ESFT cases arise between the ages of 10-20 and rapidly spread to the lungs and bone marrow (reviewed by (*174*)). While less than 25 percent of patients present with metastasis, almost all patients are thought to have sub-clinical metastatic disease due to the high relapse rate of

patients undergoing localized disease therapy. Consequently, ESFT therapy is much more aggressive than prostate cancer therapy.

Therapy for localized disease consists of an intensive five drug regimen in which four to six cycles of vincristine, doxorubicin and cyclophosphamide are alternated with a combination of ifosfamide and etoposide (*175*).Therapy can also include de-bulking surgery and/or adjuvant radiation in combination with adjuvant or neoadjuvant administration of the five drug regimen. In patients with disseminated disease at presentation, despite the overt toxicity of the combinations, similar five drug strategies are employed with even higher doses of drug. Upon disease relapse, the choice of second line therapy can depend on the initial treatment course but can include radiation and surgery to remove metastatic lesions or even amputation of extremities (*176*). Second line chemotherapies are only beginning to emerge but two different approaches are currently in clinical trials 1) inhibitors of insulin-like growth factor-I receptor (*177, 178*) or 2) combination of two alkylating agents irinotecan and temozolomide (*179*). Assuming that novel chemotherapeutic protocols for ESFTs would enter the clinic in the salvage setting before being compared to frontline therapy, molecular approaches that disrupt EWS-FLI1 activity while also potentiating the effects of either IGF1 inhibitors or alkylating agents would be ideal for evaluation in relapsed ESFTs.

### Conclusion of Clinical Considerations

Taken together, ETS gene fusions are rapidly detectable events that can be used for cancer diagnosis and to stratify patients for targeted therapy. Beyond the emerging therapeutics for androgen signaling axis in PCa, current treatment protocols generally lack a specific molecular target in these diseases. In both cases, potential therapeutics would be evaluated in cancer cells that have been heavily pre-treated.

## APPROACHES TO INHIBIT TRANSCRIPTION FACTORS

Given the critical role of ETS transcription factors in a diverse array of cancers, these proteins represent an obvious target for the rationale implementation of molecularly tailored therapy. In recent years, several approaches have been taken to inhibit transcription factor activity both directly and indirectly. Here, I will review several of these approaches.

### Disruption of transcription factor protein expression

One way to inhibit transcription factor activity is to inhibit expression of the transcription factor itself. Investigations to therapeutically inhibit protein expression have primarily used one of two methods. The first method, and the only one being investigated clinically, is to deliver target specific siRNA to cancer cells. This has been an area of intense biomedical investigation (reviewed in (*180*)) and provides a promising approach for all cancer targets, especially those that harbor tumor-specific target mRNA sequences such as gene fusion

junctions. As such, in the long term, this approach may provide an excellent therapeutic window for many cancer-specific targets.

In line with approaches to selectively degrade target mRNAs, an alternative approach is to target proteins for degradation in the proteasome (reviewed in (*181*)). The idea is simple – to identify small molecules or peptides that bind tightly to the protein interest and to target these molecules (and therefore their tight binding partners) to the proteasome with targeting peptide sequences. However, this remains an area of active research and it is unclear if this will be a feasible approach to cancer therapeutics in the long term. Additionally, as this approach requires surface binding molecules, it is unclear as to whether or not this approach provides advantages over surface binding molecules that disrupt critical protein:protein interactions. Regardless, these approaches will not be explored in this thesis.

**Indirect inhibition**

Transcription factors are often thought of as proteins that bind to specific genomic loci and recruit co-factors in a combinatorial fashion to regulate the transcription of pathways involved in cellular responses (*182*). In this simplistic view of transcription factor activity, and given the fact that most clinically viable drugs inhibit enzymes, the most rationale targets to make an immediate impact for a transcription factor-driven malignancy are enzymes encoded by susceptibility genes – or genes that are critical for maintenance of the malignancy. In this way, approaches can leverage clinically available inhibitors to bypass the years of

clinical development that would be required for novel small molecules. With this idea, three concepts may be viable. First is to identify downstream enzymes that are critical for the transcription factor's oncogenic activities. Unfortunately, while several groups have established ETS transcriptional pathways in prostate cancer, a critical downstream enzyme has not yet been identified. Second is to inhibit enzymes critical for transcription factor activity and while this approach has been demonstrated to work for some transcription factors *in vitro*, it has not yet been successfully implemented in the clinic. Importantly, these two concepts can be approached using a combination of either hypothesis-based or screening-based methods and several groups have attempted a variety of approaches to identify critical co-factors or downstream enzymes. Finally, screening methods may identify susceptibility genes with an unknown connection to the transcription factor, thus, providing the third conceptual category to indirectly inhibit a transcription factor.

Given these three concepts, it is important to implement both hypothesis-driven and screening-based approaches. Of these, several groups have used both high throughput small molecule and shRNA/siRNA screens in attempts to identify both indirect and direct inhibitors of transcription factors. These screens have achieved variable levels success and have been previously reviewed elsewhere. Because I restricted my approaches to hypothesis-based methods for this thesis, I will not discuss these screening strategies further.

Finally, an alternative approach to performing in house screening is to leverage the high throughput, unbiased screening data that is rapidly becoming

publically available. For example, in the future, drug screens will be run on cell line panels consisting of 500 or more cell lines derived from various histological subtypes to determine the $IC_{50}$ values for each cell line. Such approaches will have great utility for identifying drug sensitivities for genetic lesions that are well represented in the cell line panel. Additionally, bioinformatics methods could be used to rank order the cell lines based on gene expression profiles (ie. by expression of a transcription factor and its target genes) to then identify drugs that are most effective in specific subsets of the panel. As such, while cell lines containing ETS genomic rearrangements are generally rare, bioinformatic clustering cell lines based on a "pathogenic" ETS signature may be an appropriate method to identify inhibitors of ETS transcription factors from publically available data.

**Direct physical disruption of a transcription factor**

The approaches outlined thus far will all be useful in identifying indirect inhibitors of transcription factors, some of which may be clinically available. To directly inhibit the transcription factor, one of a few approaches can be taken: 1) Disrupt the transcription factor:DNA interaction, 2) disrupt transcription factor:co-factor interactions. Importantly, however, these approaches will likely identify molecules that require long term clinical development and therefore will not make an impact on patients for many years. Despite the time commitment that is required to develop such an inhibitor, it is important to pursue these approaches as alternative strategies to treat transcription factor-driven

malignancies. One problem that will likely significantly impede the progress of this class of drug is that traditional small molecules are largely engulfed by enzymes and kinases such that they fit into hydrophobic pocket, therefore, surface binding molecules will have unique chemical properties by necessity.

Nonetheless, the first approach to directly inhibit a transcription factor is to disrupt its ability to bind to the genome. The problem with this approach is that transcription factor:DNA interactions are generally very high affinity and thus, require even higher affinity small molecules to disrupt the interaction. To bypass the affinity issue, oligonucleotide decoys have been introduced to sequester the transcription factor away from genomic DNA. This method has been used to inhibit the STAT3 transcription factor in EGFR amplified HNSCC models. In these experiments, the oligonucleotide is introduced to cells, can penetrate the nuclease, bind to STAT3 and disrupt STAT3 transcription (reviewed in (*183*)). Indeed, using next generation ChIP-sequencing experiments, we identified the consensus ERG binding site in prostate cancer cell lines and tissues (*143*) and approaches to introduce ERG binding site oligonucleotide decoys would be expected to inhibit tumor growth. However, as with siRNA delivery, several challenges still exist for introducing oligonucleotide decoys clinically.

An alternative approach to directly inhibit transcription factor:DNA interactions is to identify or design small molecules with a higher affinity for DNA target sequence of the transcription factor than the transcription factor itself. N-methylpyrrole (Py)-N-methylimidazole (Im) polyamides are a class of programmable DNA-binding ligands capable of binding specific DNA sequences

with very high affinity (*184, 185*). These molecules have been used to disrupt transcriptional activation of several factors including hypoxia-inducible factor (*186, 187*), androgen receptor (*188*) and glucocorticoid receptor (*189*). To date, these programmable DNA binding molecules have not been studied in xenograft models, thus, there therapeutic potential remains unclear. However, this strategy will be important to at least identify tool compounds capable of disrupting the ERG:DNA interaction and may also provide a structural basis for the design of more traditional small molecule like compounds.

Beyond the potential of blocking protein:DNA interactions to disrupt transcription factor activity, several groups have also aimed to block protein:co-factor interactions. Possibly the most successful approach to disrupt this type of interaction utilized a rationally designed a peptide inhibitor to disrupt the interaction between the repressive BCL6 transcription factor and its co-repressors SMRT, NCoR and BCoR. The peptide inhibitor disrupted BCL6-mediated transcriptional repression of reactivate BCL6 target genes in cell lines and xenograft models (*190*). The peptide was designed using detailed crystallographic data of the SMRT co-repressor in complex with the BCL6 BTB (Bric-a-brac, Tramtrack, Broad complex) domain (*191*) such that wtSMRT BBD motif was used to out-compete recruitment of the full co-repressor complex. Given the initial success of the competitive SMRT peptides, Cerchietti *et. al.* worked to shorten the binding peptide and converted it into the D-amino acid, retro configuration (*192*). This created a peptide that was resistant to serum proteases, more stable in xenograft models and inhibited xenograft growth of BCL6-

dependent cells at a much lower dose than the L-form SMRT peptide. Unfortunately, with a limited number of known interacting partners in prostate cancer, none of which were known to be functionally relevant, this method was not feasible for ETS transcription factors at the onset of this thesis.

Complementary to the idea of using peptidomimetics to block protein:co-factor interactions, Erkizan *et. al.* recently focused on disrupting the EWS-FLI1:RHA interaction using a surface plasmon resonance screen to identify small molecules that bind to the surface of EWS-FLI1 (*193*). This approach led to the identification of a lead compound called YK-4-279, which blocked the EWS-FLI1:RHA interaction as well as EWS-FLI1 mediated phenotypes such as colony formation, transcription and Cyclin D1 at 10μM dose. Likewise, 50μM YK-4-279 led to the induction of apoptosis in ESFT cell lines and administration of 1.5mg of drug delayed ESFT cell line xenograft growth (CHP-100 cells). Unfortunately, many of these experiments lacked controls and it is unclear whether the compound has off target effects or disrupts the activity of other ETS transcription factors.

Recently, several groups have successfully modulated cMyc transcriptional activity using both peptides and small molecules to disrupt the critical cMyc:Max interaction or cMYC:Max:DNA interactions. A paper by Lu *et. al.* developed a peptidomimetic that served as the basis for some of these studies that was able to repress reporter gene activity (*194*). While these approaches are leading to candidate small molecules, many utilize advanced biophysical techniques to model structural details of the cMyc:Max:DNA interactions and are

based on improving the peptidomimetic inhibitor of cMyc:Max and these approaches have been recently reviewed (*195*). At the initiation of this thesis, ETS co-factors were almost completely unknown and the structure of ERG had not been solved. Given the infancy of our ETS transcription factor knowledge compared to cMyc, many of the approaches that have been recently used to develop inhibitors cMyc were simply too complex to adapt to ETS proteins at the initiation of this thesis.

Finally, in addition to these approaches, groups have also generally screened purified transcription factors to identify protein binding molecules that disrupt either transcription factor:DNA or transcription factor:co-factor interactions. One problem with this approach is that many DNA binding protein domains are thought to be intrinsically disordered in the absence of interacting partners, and possible in the absence of DNA (*196, 197*). This was unknown for ETS proteins prior to this thesis. Furthermore, it is difficult to grossly predict which protein surfaces are most important for co-factor interactions. Despite this fact, in recent work by Ng *et. al.*, a new diversity-based chemical library was developed based on a lactam carboxamide backbone (*198*). In screening this library, the group found that one of these compounds could disrupt the interaction of the HoxA13 transcription factor with DNA and this molecule was demonstrated to inhibit HoxA13 transcription factor activity *in vitro*.

Taken together, several potential approaches exist to disrupt ETS transcription factors in cancer. While many of these approaches have been attempted for other transcription factors as highlighted above, none have achieved

significant clinical activity. **Figure 1.2** demonstrates some of the potential methods to specifically disrupt ERG transcription factor activity in prostate cancer.

## CONCLUSION OF INTRODUCTION

Despite the clear pathogenic role of aberrant transcription factor activity in a wide array of diseases, these proteins are currently considered to be "undruggable" molecular targets (*20*). Consequently, clinical protocols for these diseases often utilize non-targeted agents such as vincristine, doxorubicin and cyclophosphamide alternated with a combination of ifosfamide and etoposide in ESFTs (*175*). These protocols have achieved some clinical success, but are very toxic to patients and can lead to even more aggressive disease upon recurrence. In the future, the hope is that molecular lesions will dictate therapeutic course and any approach – whether direct or indirect – that can disrupt transcription factor activity in a clinically relevant manner would represent a major breakthrough for the field of medicine. Of all the approaches outlined above that have been attempted to inhibit transcription factors and because transcription factors are dependent on the ability to assemble protein interaction networks at various genomic loci (*182*), perhaps the most promising long term strategy is to develop inhibitors of transcription factor:co-factor interactions. To date, however, these interactions are currently intractable and the quest to identify inhibitors of any protein:protein interactions is impeding a move to personalized medicine for transcription factor-driven malignancies. Undeniably, this problem was brought

into the limelight in a highly publicized challenge to all researchers issued by James Darnell ten years ago: "*What is the benefit to medicine in all of the twenty-first century promise of proteomics if we can not selectively inhibit protein:protein interactions*?" (*20*).

Therefore, in this thesis, I will challenge the dogma that transcription factors are "undruggable" molecular targets using both prostate cancer and ESFT as the primary model systems to develop inhibitors of ETS transcription factors with hypothesis driven approaches. I will aim to expand upon the concepts and approaches to inhibit transcription factors outlined above in an innovative and systematic fashion. And because this thesis's overall impact on cancer patients will be determined by the ability to translate therapies into the clinic, I will focus on increasing the indication of targeted small molecules already in clinical development to include ETS-driven malignancies. In this way, I may be able to leverage existing compounds to make an immediate impact for patients. Finally, I will also aim to build up the structural knowledge of ETS transcription factors in prostate cancer such that long term approaches to directly inhibit ETS factors can be initiated. In doing so, I will aim to develop a proof-of-concept inhibitor to challenge the notion that ETS transcription factor:co-factor interactions are therapeutically intractable.

## FIGURES

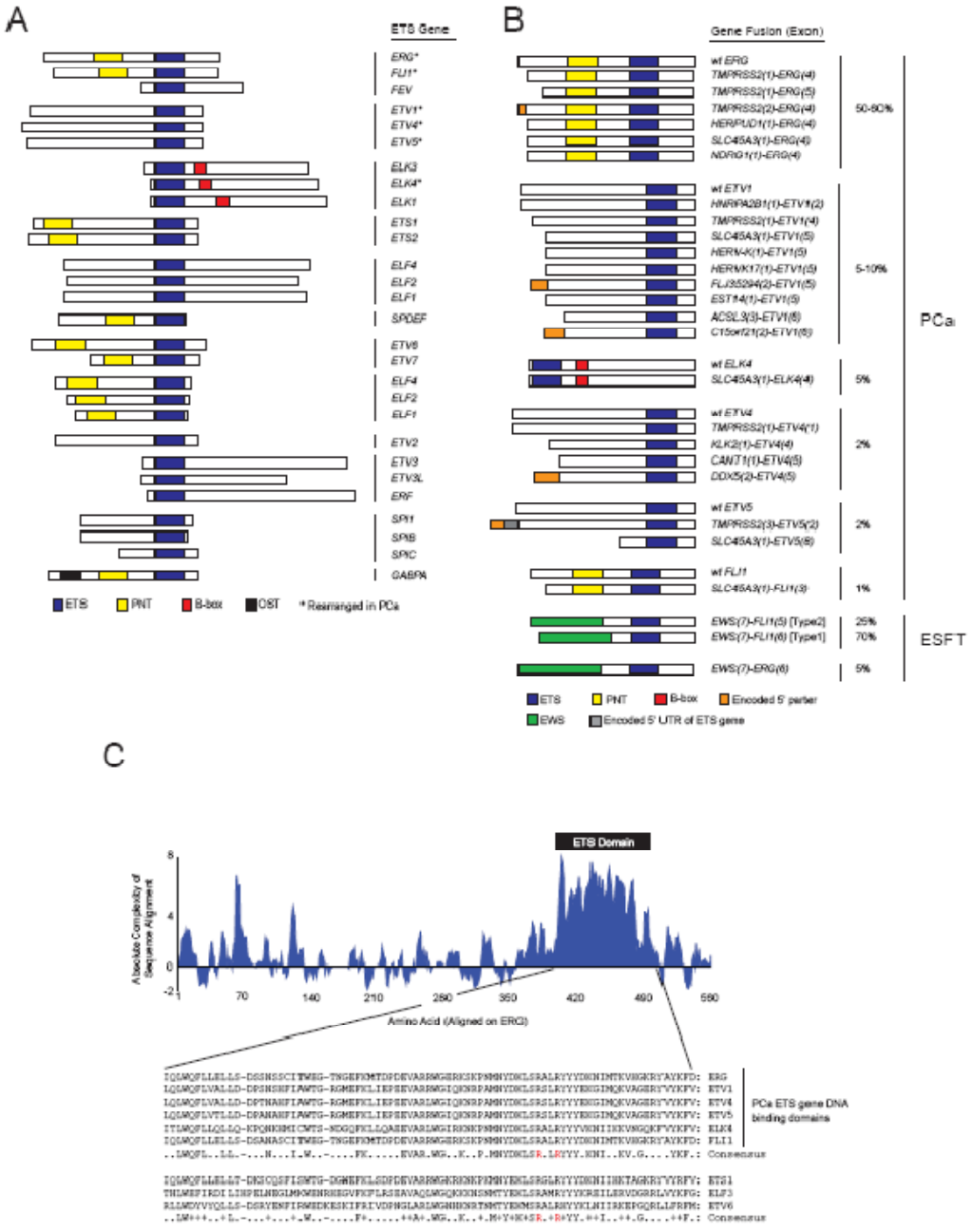



**Figure 1.1: Schematic representation of conserved domains in ETS proteins and ETS gene fusion protein products.**

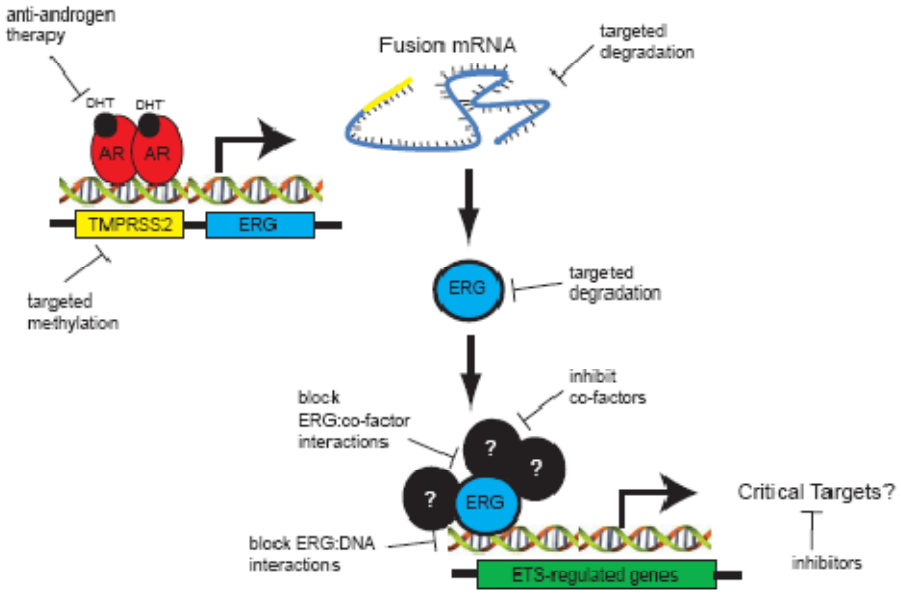



**Figure 1.2: Schematic representation of strategies to inhibit the *TMPRSS2-ERG* transcription factor signaling axis.** These including blocking AR-mediated transcription (AR antagonists, targets DNA silencing), anti-sense RNA approaches, induction of protein degradation, co-factor inhibitors, molecules that block either ERG:co-factor and/or ERG:DNA interactions as well as inhibitors of critical downstream components of the signaling axis.

## REFERENCES


1. D. Von Hansemann, *Virchows Arch. A Pathol. Anat.* 119, 299 (1890).
2. T. Boveri, *Zur frage der entstehung maligner tumoren*. G. Fischer, Ed. (Jena, Germany, 1914), pp. 64.
3. J. Lejeune, M. Gautier, R. Turpin, *C R Hebd Seances Acad Sci* 248, 1721 (Mar 16, 1959).
4. J. Lejeune, R. Turpin, M. Gautier, *Bull Acad Natl Med* 143, 256 (Apr 7-14, 1959).
5. K. H. Rothfels, L. Siminovitch, *Stain Technol* 33, 73 (Mar, 1958).
6. P. Nowell, D. Hungerford, *Science* 132, 1497 (1960).
7. P. C. Nowell, D. A. Hungerford, *J Natl Cancer Inst* 27, 1013 (Nov, 1961).
8. I. M. Tough *et al.*, *Lancet* 1, 411 (Feb 25, 1961).
9. J. C. Brenner, A. M. Chinnaiyan, *Cancer Cell* 19, 301 (Mar 8, 2011).
10. J. D. Rowley, *Clin Haematol* 9, 55 (Feb, 1980).
11. J. Groffen *et al.*, *Cell* 36, 93 (Jan, 1984).
12. T. G. Lugo, A. M. Pendergast, A. J. Muller, O. N. Witte, *Science* 247, 1079 (Mar 2, 1990).
13. B. J. Druker *et al.*, *N Engl J Med* 344, 1038 (Apr 5, 2001).
14. B. J. Druker *et al.*, *N Engl J Med* 344, 1031 (Apr 5, 2001).
15. H. Kantarjian *et al.*, *N Engl J Med* 346, 645 (Feb 28, 2002).
16. M. Talpaz *et al.*, *Blood* 99, 1928 (Mar 15, 2002).
17. S. A. Tomlins *et al.*, *Science* 310, 644 (Oct 28, 2005).
18. C. Kumar-Sinha, S. A. Tomlins, A. M. Chinnaiyan, *Nat Rev Cancer* 8, 497 (Jul, 2008).
19. S. A. Tomlins *et al.*, *Eur Urol* 56, 275 (Aug, 2009).
20. J. E. Darnell, Jr., *Nat Rev Cancer* 2, 740 (Oct, 2002).
21. A. D. Sharrocks, *Nat Rev Mol Cell Biol* 2, 827 (Nov, 2001).
22. M. Z. Ivanov X, Nedyalk S and Todorov TG., *Bugarian Acad. Sci. Bull. Inst. Pathol. Comp. Animaux Domestiques.* 9, 5 (1962).
23. N. Sotirov, *J Natl Cancer Inst* 66, 1143 (Jun, 1981).
24. T. Graf, N. Oker-Blom, T. G. Todorov, H. Beug, *Virology* 99, 431 (Dec, 1979).
25. M. G. Moscovici *et al.*, *Virology* 129, 65 (Aug, 1983).
26. K. Radke, H. Beug, S. Kornfeld, T. Graf, *Cell* 31, 643 (Dec, 1982).
27. D. Leprince *et al.*, *Nature* 306, 395 (Nov 24-30, 1983).
28. M. Nunn, H. Weiher, P. Bullock, P. Duesberg, *Virology* 139, 330 (Dec, 1984).
29. M. F. Nunn, P. H. Seeburg, C. Moscovici, P. H. Duesberg, *Nature* 306, 391 (Nov 24-30, 1983).
30. C. de Taisne, A. Gegonne, D. Stehelin, A. Bernheim, R. Berger, *Nature* 310, 581 (Aug 16-22, 1984).
31. D. K. Watson *et al.*, *Proc Natl Acad Sci U S A* 82, 7294 (Nov, 1985).
32. B. M. Degnan, S. M. Degnan, T. Naganuma, D. E. Morse, *Nucleic Acids Res* 21, 3479 (Jul 25, 1993).
33. Z. Wang, Q. Zhang, *Evol Bioinform Online* 5, 119 (2009).
34. H. S. Kang *et al.*, *J Mol Biol* 377, 636 (Mar 28, 2008).
35. M. Hassler, T. J. Richmond, *Embo J* 20, 3018 (Jun 15, 2001).
36. C. K. Galang, W. J. Muller, G. Foos, R. G. Oshima, C. A. Hauser, *J Biol Chem* 279, 11281 (Mar 19, 2004).
37. P. C. Hollenhorst, D. A. Jones, B. J. Graves, *Nucleic Acids Res* 32, 5693 (2004).

38. P. C. Hollenhorst, L. P. McIntosh, B. J. Graves, *Annu Rev Biochem* 80, 437 (Jun 7, 2011).
39. B. R. Szymczyna, C. H. Arrowsmith, *J Biol Chem* 275, 28363 (Sep 15, 2000).
40. W. E. Jurgens G., Nusslein-Volhard C., Kluding H., *Roux's Arch Dev Biol* 193, 283 (1984).
41. C. Klambt, *Development* 117, 163 (Jan, 1993).
42. C. P. Ponting, *Protein Sci* 4, 1928 (Sep, 1995).
43. J. Schultz, C. P. Ponting, K. Hofmann, P. Bork, *Protein Sci* 6, 249 (Jan, 1997).
44. M. Kyba, H. W. Brock, *Dev Genet* 22, 74 (1998).
45. Y. Hu *et al.*, *Clin Cancer Res* 14, 4719 (Aug 1, 2008).
46. J. Wang *et al.*, *Cancer Res* 68, 8516 (Oct 15, 2008).
47. C. A. Kim *et al.*, *Embo J* 20, 4173 (Aug 1, 2001).
48. C. Wasylyk, J. P. Kerckaert, B. Wasylyk, *Genes Dev* 6, 965 (Jun, 1992).
49. S. M. Meadows, C. T. Myers, P. A. Krieg, *Semin Cell Dev Biol* 22, 976 (Dec, 2011).
50. A. M. Randi, A. Sperone, N. H. Dryden, G. M. Birdsey, *Biochem Soc Trans* 37, 1248 (Dec, 2009).
51. G. M. Birdsey *et al.*, *Blood* 111, 3498 (Apr 1, 2008).
52. F. McLaughlin *et al.*, *Blood* 98, 3332 (Dec 1, 2001).
53. S. J. Loughran *et al.*, *Nat Immunol* 9, 810 (Jul, 2008).
54. F. Liu, R. Patient, *Circ Res* 103, 1147 (Nov 7, 2008).
55. V. N. Pham *et al.*, *Dev Biol* 303, 772 (Mar 15, 2007).
56. D. D. Spyropoulos *et al.*, *Mol Cell Biol* 20, 5643 (Aug, 2000).
57. E. A. Kruse *et al.*, *Proc Natl Acad Sci U S A* 106, 13814 (Aug 18, 2009).
58. S. A. Lacadie, L. I. Zon, *Genes Dev* 25, 289 (Feb 15, 2011).
59. S. Taoudi *et al.*, *Genes Dev* 25, 251 (Feb 1, 2011).
60. A. Hart *et al.*, *Immunity* 13, 167 (Aug, 2000).
61. H. Kawada *et al.*, *Int J Hematol* 73, 463 (Jun, 2001).
62. H. Raslova *et al.*, *J Clin Invest* 114, 77 (Jul, 2004).
63. M. Athanasiou *et al.*, *Cell Growth Differ* 7, 1525 (Nov, 1996).
64. M. Eisbacher *et al.*, *Mol Cell Biol* 23, 3427 (May, 2003).
65. C. D. Baldus *et al.*, *J Clin Oncol* 24, 4714 (Oct 10, 2006).
66. R. S. Maser *et al.*, *Nature* 447, 966 (Jun 21, 2007).
67. C. G. Mullighan *et al.*, *Nature* 446, 758 (Apr 12, 2007).
68. P. J. Valk *et al.*, *N Engl J Med* 350, 1617 (Apr 15, 2004).
69. C. D. Baldus *et al.*, *Proc Natl Acad Sci U S A* 101, 3915 (Mar 16, 2004).
70. G. Marcucci *et al.*, *J Clin Oncol* 23, 9234 (Dec 20, 2005).
71. K. H. Metzeler *et al.*, *J Clin Oncol* 27, 5031 (Oct 20, 2009).
72. E. Coskun *et al.*, *Leuk Res* 35, 208 (Feb, 2011).
73. S. D. Moore *et al.*, *Leuk Res* 30, 1037 (Aug, 2006).
74. J. Kim *et al.*, *Cancer Genet Cytogenet* 194, 111 (Oct 15, 2009).
75. S. Tsuzuki, O. Taguchi, M. Seto, *Blood* 117, 3858 (Apr 7, 2011).
76. S. Salek-Ardakani *et al.*, *Cancer Res* 69, 4665 (Jun 1, 2009).
77. S. A. Tomlins *et al.*, *Cancer Res* 66, 3396 (Apr 1, 2006).
78. B. E. Helgeson *et al.*, *Cancer Res* 68, 73 (Jan 1, 2008).
79. C. A. Maher *et al.*, *Nature* 458, 97 (Mar 5, 2009).
80. D. S. Rickman *et al.*, *Cancer Res* 69, 2734 (Apr 1, 2009).
81. P. Paulo *et al.*, *Genes Chromosomes Cancer* 51, 240 (Mar, 2011).

82. K. G. Hermans *et al.*, *Cancer Res* 68, 3094 (May 1, 2008).
83. R. Mehra *et al.*, *Cancer Res* 68, 3584 (May 15, 2008).
84. B. Han *et al.*, *Cancer Res* 68, 7629 (Sep 15, 2008).
85. K. Park *et al.*, *Neoplasia* 12, 590 (Jul, 2011).
86. R. S. Mani *et al.*, *Science* 326, 1230 (Nov 27, 2009).
87. N. Cerveira *et al.*, *Neoplasia* 8, 826 (Oct, 2006).
88. J. Clark *et al.*, *Oncogene* 26, 2667 (Apr 19, 2007).
89. K. G. Hermans *et al.*, *Cancer Res* 66, 10658 (Nov 15, 2006).
90. K. Iljin *et al.*, *Cancer Res* 66, 10242 (Nov 1, 2006).
91. J. Lapointe *et al.*, *Mod Pathol* 20, 467 (Apr, 2007).
92. A. B. Rajput *et al.*, *J Clin Pathol* 60, 1238 (Nov, 2007).
93. M. J. Soller *et al.*, *Genes Chromosomes Cancer* 45, 717 (Jul, 2006).
94. M. Yoshimoto *et al.*, *Neoplasia* 8, 465 (Jun, 2006).
95. J. Wang, Y. Cai, C. Ren, M. Ittmann, *Cancer Res* 66, 8347 (Sep 1, 2006).
96. B. Lin *et al.*, *Cancer Res* 59, 4180 (Sep 1, 1999).
97. D. E. Afar *et al.*, *Cancer Res* 61, 1686 (Feb 15, 2001).
98. M. H. Vaarala, K. Porvari, A. Kyllonen, O. Lukkarinen, P. Vihko, *Int J Cancer* 94, 705 (Dec 1, 2001).
99. S. Perner *et al.*, *Am J Surg Pathol* 31, 882 (Jun, 2007).
100. R. S. Mani, A. M. Chinnaiyan, *Nat Rev Genet* 11, 819 (Dec, 2010).
101. S. Perner *et al.*, *Cancer Res* 66, 8337 (Sep 1, 2006).
102. R. Mehra *et al.*, *Mod Pathol* 20, 538 (May, 2007).
103. M. R. Teixeira, *Am J Surg Pathol* 32, 642 (Apr, 2008).
104. M. C. Haffner *et al.*, *Nat Genet* 42, 668 (Aug, 2010).
105. C. Lin *et al.*, *Cell* 139, 1069 (Dec 11, 2009).
106. I. A. Klein *et al.*, *Cell* 147, 95 (Sep 30, 2011).
107. R. Chiarle *et al.*, *Cell* 147, 107 (Sep 30, 2011).
108. S. A. Tomlins *et al.*, *Nature* 448, 595 (Aug 2, 2007).
109. O. Klezovitch *et al.*, *Proc Natl Acad Sci U S A* 105, 2105 (Feb 12, 2008).
110. B. S. Carver *et al.*, *Nat Genet* 41, 619 (May, 2009).
111. J. C. King *et al.*, *Nat Genet* 41, 524 (May, 2009).
112. Y. Zong *et al.*, *Proc Natl Acad Sci U S A* (Jul 10, 2009).
113. S. A. Tomlins *et al.*, *Neoplasia* 10, 177 (Feb, 2008).
114. C. A. Arndt, W. M. Crist, *N Engl J Med* 341, 342 (Jul 29, 1999).
115. L. L. Knoop, S. J. Baker, *J Biol Chem* 276, 22317 (Jun 22, 2001).
116. A. Uren, J. A. Toretsky, *Future Oncol* 1, 521 (Aug, 2005).
117. J. A. Toretsky *et al.*, *Cancer Res* 66, 5574 (Jun 1, 2006).
118. R. Smith *et al.*, *Cancer Cell* 9, 405 (May, 2006).
119. H. Kovar *et al.*, *Cell Growth Differ* 7, 429 (Apr, 1996).
120. A. Maksimenko *et al.*, *Ann N Y Acad Sci* 1002, 72 (Dec, 2003).
121. M. Ouchida, T. Ohno, Y. Fujimura, V. N. Rao, E. S. Reddy, *Oncogene* 11, 1049 (Sep 21, 1995).
122. S. S. Myatt, S. A. Burchill, *Oncogene* 27, 985 (Feb 7, 2008).
123. E. C. Torchia, K. Boyd, J. E. Rehg, C. Qu, S. J. Baker, *Mol Cell Biol* 27, 7918 (Nov, 2007).
124. P. P. Lin *et al.*, *Cancer Res* 68, 8968 (Nov 1, 2008).
125. J. C. Brenner, A. M. Chinnaiyan, *Biochim Biophys Acta* 1796, 201 (Dec, 2009).
126. H. A. Oberman, *Am J Surg Pathol* 4, 465 (Oct, 1980).

127. C. Tognon *et al.*, *Cancer Cell* 2, 367 (Nov, 2002).
128. T. R. Golub *et al.*, *Proc Natl Acad Sci U S A* 92, 4917 (May 23, 1995).
129. P. Papadopoulos, S. A. Ridge, C. A. Boucher, C. Stocking, L. M. Wiedemann, *Cancer Res* 55, 34 (Jan 1, 1995).
130. T. R. Golub *et al.*, *Mol Cell Biol* 16, 4107 (Aug, 1996).
131. V. Lacronique *et al.*, *Science* 278, 1309 (Nov 14, 1997).
132. P. Peeters *et al.*, *Blood* 90, 2535 (Oct 1, 1997).
133. J. Schwaller *et al.*, *Embo J* 17, 5321 (Sep 15, 1998).
134. G. Cazzaniga *et al.*, *Blood* 94, 4370 (Dec 15, 1999).
135. Y. Iijima *et al.*, *Blood* 95, 2126 (Mar 15, 2000).
136. T. R. Golub, G. F. Barker, M. Lovett, D. G. Gilliland, *Cell* 77, 307 (Apr 22, 1994).
137. F. Yagasaki *et al.*, *Cancer Res* 61, 8371 (Dec 1, 2001).
138. S. K. Bohlander, *Semin Cancer Biol* 15, 162 (Jun, 2005).
139. J. Li *et al.*, *Science* 275, 1943 (Mar 28, 1997).
140. B. S. Carver *et al.*, *Nature* 457, E1; discussion E2 (Feb 12, 2009).
141. M. F. Berger *et al.*, *Nature* 470, 214 (Feb 10, 2011).
142. Z. Ding *et al.*, *Nature* 470, 269 (Feb 10, 2011).
143. J. Yu *et al.*, *Cancer Cell* 17, 443 (May 18, 2010).
144. B. S. Taylor *et al.*, *Cancer Cell* 18, 11 (Jul 13, 2010).
145. M. Liedtke, M. L. Cleary, *Blood* 113, 6061 (Jun 11, 2009).
146. I. Varela *et al.*, *Nature* 469, 539 (Jan 27, 2011).
147. D. W. Parsons *et al.*, *Science* 331, 435 (Jan 28, 2011).
148. R. D. Morin *et al.*, *Nature* 476, 298 (Aug 18, 2011).
149. Y. Jiao *et al.*, *Science* 331, 1199 (Mar 4, 2011).
150. Y. Gui *et al.*, *Nat Genet* 43, 875 (Sep, 2011).
151. K. C. Wiegand *et al.*, *N Engl J Med* 363, 1532 (Oct 14, 2010).
152. S. Jones *et al.*, *Science* 330, 228 (Oct 8, 2010).
153. S. A. Tomlins *et al.*, *Sci Transl Med* 3, 94ra72 (Aug 3, 2011).
154. S. Roychowdhury *et al.*, *Sci Transl Med* 3, 111ra121 (Nov 30, 2011).
155. I. Thompson *et al.*, *J Urol* 177, 2106 (Jun, 2007).
156. G. L. Lu-Yao *et al.*, *Jama* 300, 173 (Jul 9, 2008).
157. U. E. Studer *et al.*, *J Clin Oncol* 24, 1868 (Apr 20, 2006).
158. H. K. Tsai, A. V. D'Amico, N. Sadetsky, M. H. Chen, P. R. Carroll, *J Natl Cancer Inst* 99, 1516 (Oct 17, 2007).
159. A. L. Potosky *et al.*, *J Natl Cancer Inst* 94, 430 (Mar 20, 2002).
160. I. F. Tannock *et al.*, *N Engl J Med* 351, 1502 (Oct 7, 2004).
161. D. R. Berthold *et al.*, *J Clin Oncol* 26, 242 (Jan 10, 2008).
162. J. S. de Bono *et al.*, *Lancet* 376, 1147 (Oct 2, 2010).
163. G. Attard *et al.*, *Cancer Res* 69, 2912 (Apr 1, 2009).
164. S. E. Barrie *et al.*, *J Steroid Biochem Mol Biol* 50, 267 (Sep, 1994).
165. G. Attard *et al.*, *J Clin Oncol* 27, 3742 (Aug 10, 2009).
166. G. A. Potter, S. E. Barrie, M. Jarman, M. G. Rowlands, *J Med Chem* 38, 2463 (Jun 23, 1995).
167. J. S. de Bono *et al.*, *N Engl J Med* 364, 1995 (May 26, 2011).
168. C. J. Ryan *et al.*, *Clin Cancer Res* 17, 4854 (Jul 15, 2011).
169. V. Tzelepi *et al.*, *J Clin Oncol* 29, 2574 (Jun 20, 2011).
170. D. C. Danila *et al.*, *Eur Urol* 60, 897 (Nov, 2011).
171. H. I. Scher *et al.*, *Lancet* 375, 1437 (Apr 24, 2010).

172. C. Schmidt, *J Natl Cancer Inst* 103, 175 (Feb 2, 2011).
173. N. J. Clegg *et al.*, *Cancer Res* (Jan 20, 2012).
174. H. D. Suit *et al.*, *J Clin Oncol* 6, 854 (May, 1988).
175. H. E. Grier *et al.*, *N Engl J Med* 348, 694 (Feb 20, 2003).
176. G. Bacci, A. Briccoli, P. Picci, S. Ferrari, *Eur J Cancer* 31A, 999 (Jun, 1995).
177. H. Juergens *et al.*, *J Clin Oncol* 29, 4534 (Dec 1, 2011).
178. A. S. Pappo *et al.*, *J Clin Oncol* 29, 4541 (Dec 1, 2011).
179. D. A. Casey *et al.*, *Pediatr Blood Cancer* 53, 1029 (Dec, 2009).
180. K. A. Whitehead, R. Langer, D. G. Anderson, *Nat Rev Drug Discov* 8, 129 (Feb, 2009).
181. K. D. Wilkinson, *Semin Cell Dev Biol* 11, 141 (Jun, 2000).
182. M. T. Marr, 2nd, Y. Isogai, K. J. Wright, R. Tjian, *Genes Dev* 20, 1458 (Jun 1, 2006).
183. A. L. Boehm *et al.*, *Mol Pharmacol* 73, 1632 (Jun, 2008).
184. P. B. Dervan, B. S. Edelson, *Curr Opin Struct Biol* 13, 284 (Jun, 2003).
185. C. F. Hsu *et al.*, *Tetrahedron* 63, 6146 (Jul 2, 2007).
186. B. Z. Olenyuk *et al.*, *Proc Natl Acad Sci U S A* 101, 16768 (Nov 30, 2004).
187. N. G. Nickols, C. S. Jacobs, M. E. Farkas, P. B. Dervan, *ACS Chem Biol* 2, 561 (Aug 17, 2007).
188. N. G. Nickols, P. B. Dervan, *Proc Natl Acad Sci U S A* 104, 10418 (Jun 19, 2007).
189. K. A. Muzikar, N. G. Nickols, P. B. Dervan, *Proc Natl Acad Sci U S A* 106, 16598 (Sep 29, 2009).
190. J. M. Polo *et al.*, *Nat Med* 10, 1329 (Dec, 2004).
191. K. F. Ahmad *et al.*, *Mol Cell* 12, 1551 (Dec, 2003).
192. L. C. Cerchietti *et al.*, *Blood* 113, 3397 (Apr 9, 2009).
193. H. V. Erkizan *et al.*, *Nat Med* 15, 750 (Jul, 2009).
194. X. Lu, P. K. Vogt, D. L. Boger, J. Lunec, *Oncol Rep* 19, 825 (Mar, 2008).
195. A. N. Koehler, *Curr Opin Chem Biol* 14, 331 (Jun, 2010).
196. T. Berg, *Curr Opin Chem Biol* 12, 464 (Aug, 2008).
197. S. K. Nair, S. K. Burley, *Cell* 112, 193 (Jan 24, 2003).
198. P. Y. Ng, Y. Tang, W. M. Knosp, H. S. Stadler, J. T. Shaw, *Angew Chem Int Ed Engl* 46, 5352 (2007).

# Chapter II

# PARP1 and DNA-PKcs are required for ETS-mediated invasion and metastasis in Prostate Cancer


## SUMMARY

Recurrent fusions of ETS genes are considered driving mutations in a diverse array of cancers, including Ewing's sarcoma, acute myeloid leukemia, and prostate cancer. We investigate the mechanisms by which ETS fusions mediate their effects, and find that the product of the predominant ETS gene fusion, *TMPRSS2:ERG*, interacts in a DNA-independent manner with the enzyme poly(ADP-ribose) polymerase 1 (PARP1) and the catalytic subunit of DNA protein kinase (DNA-PKcs). ETS gene-mediated transcription and cell invasion require PARP1 and DNA-PKcs expression and activity. Importantly, pharmacological inhibition of PARP1 inhibits ETS positive, but not ETS negative, prostate cancer xenograft growth to a similar extent as a *BRCA1* deficient cell line xenograft.[1]


[1] This chapter was published in *Cancer Cell* and completed in collaboration with the following authors: Bushra Ateeq, Yong Li, Anastasia K. Yocum, Qi Cao, Irfan A. Asangani, Sonam Patel, Xiaoju Wang, Hallie Liang, Jindan Yu, Nallasivam Palanisamy, Javed Siddiqui, Wei Yan, Xuhong Cao, Rohit Mehra, Aaron Sabolch, Venkatesha Basrur, Robert J. Lonigro, Jun Yang, Scott A Tomlins, Christopher A. Maher, Kojo S.J. Elenitoba-Johnson, Maha Hussain, Nora M. Navone, Kenneth J. Pienta, Sooryanarayana Varambally, Felix Y. Feng, Arul M. Chinnaiyan

## INTRODUCTION

ETS transcription factors are aberrantly expressed in a diverse array of cancers including prostate, breast, melanoma and Ewing's sarcoma (*1-7*). Given the functional consequences of ETS gene rearrangements in prostate cancer progression presented in chapter 1, a critical question remains unanswered: are ETS gene fusions therapeutic targets, either directly or indirectly? Understanding the difficulties in targeting nuclear transcription factors using conventional therapeutic strategies (*8*), we hypothesized that associated enzymes critical for ERG function may instead serve as viable therapeutic targets to inhibit ETS-positive prostate cancer cell growth.

## MATERIALS AND METHODS

### Cell Lines

PC3 prostate cancer cell lines were grown in RPMI 1640 (Invitrogen, Carlsbad, CA) and VCaP cells in DMEM with Glutamax (Invitrogen) both supplemented with 10% FBS (Invitrogen) in 5% $CO_2$ cell culture incubator. The immortalized prostate cell line RWPE-1 was grown in Keratinocyte media with L-glutamine (Invitrogen) supplemented with 2.5μg EGF (Invitrogen) and 25mg Bovine Pituitary Extract (Invitrogen). All cultures were also maintained with 50 units/ml of Penicillin/streptomycin (Invitrogen). The genetic identity of each cell line was confirmed by genotyping samples. Briefly, DNA samples were diluted to 0.10ng/μl and analyzed in the University of Michigan DNA sequencing Core using the Profiler Plus PCR Amplification Kit (Applied Biosystems, Foster City, CA) according to the manufacturer's protocol. The 9 loci D3S1358, D5S818,

D7S820, D8S1179, D13S317, D18S51, D21S11, FGA, vWA and the Amelogenin locus were analyzed and compared to ladder control samples as previously described (*9*). Lentiviruses were generated by the University of Michigan Vector Core. PC3 cells were infected with lentiviruses expressing pLentilox-CMV-ERG or pLentilox-CMV-LACZ control and stable cell lines were selected by sorting at the University of Michigan flow cytometry core. Stable infection was monitored by confirming GFP expression every three days. Likewise, PC3-luciferase cells were created by transduction of pLentilox-CMV-Luciferase available from the University of Michigan Vector Core. Stable RWPE-ERG, RWPE-ETV1, RWPE-LACZ and RWPE_SLC45A3-BRAF cells were created and described previously (*10, 11*).

**Mass Spectrometry**

IP eluate was resolved with SDS-PAGE and visualized with silver stain (PROTSIL-2, Sigma). Each gel lane, experimental and control was excised into 14 equal sized and corresponding pieces for in-gel trypsin digestion per the manufacturer's instructions, treated with 10 mM DTT followed immediately with 50 mM iodoacetaminde to reduce and alkylate cysteine residues, and then incubated with trypsin (1:20) enzyme:protein (w:w) overnight at 37° C (Promega, Madison, WI). Peptides were then extracted from the gel, lyophilized to dryness, and stored at -80 C until further analysis.

**Protein identification by LC-tandem MS**

Lyophilized peptides from each gel fraction were reconstituted in 1 % acetic acid in 5 % acetonitrile for reverse phase separation on-line to the nano-spray equipped LTQ-XL ion trap mass spectrometer (ThermoFisher, San Jose, CA). Each fraction was loaded using an autosampler (SPARK, Michrom Bioresources, CA) and separated on an Aquasil C18 Picofrit column, (15 cm x 75 μm i.d., 15 μm tip) (New Objectives, Woburn MA). Peptides were eluted over 60 minutes with an increasing linear gradient of acetonitrile at a flow rate of 300 nL / minute. To identify the eluting peptides, the mass spectrometer was operated in a data-dependent MS/MS mode (*m*/*z* 400-2000), in which the top five ions were subjected to MS/MS at 35 % of 1 V normalized collision induced disassociation. Dynamic mass exclusion was enabled with a repeat count of 2 for 2 minutes and a list size of 200 *m/z*.

**Database Searching**

Raw spectra files were converted to mzXML format using an in-house installation of ReAdW (version 4.0.2), and searched with X!Tandem (The Global Proteome Machine Organization; version 2007.07.01.1) on a decoy database that contained the forward human IPI sequences concatenated to the reversed human IPI sequences (version 3.41) plus cRAP database of common contaminants. The database search used trypsin enzyme specificity, a mass error of 3 Da on parent ion and 0.8 Da on fragment ions, a maximum of one missed cleavage, and variable modifications of oxidation on methionine and carbamidomethlyation on cysteines. The X!Tandem database search results were analyzed and validated

using Trans-Proteomic Pipeline (TPP) software (version 4.0.2). All entries with a peptide probability > 90 % and a protein probability error rate < 5 % with at least two unique peptides were considered high confidence identifications. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Only individual protein isoforms are reported and those proteins identified with high confidence unique to experimental when compared with matched control. As a control, epitope-tagged beta-galactosidase (*LACZ*) was run in parallel and immunoprecipitated in each replicate experiment. Non-specific interactions identified in the vector controls were manually removed from the experimental protein listing to generate a final list of genuine ERG interactions found in both HEK293 and VCaP cells over eight biological replicates.

**Immunoprecipitation**

Cells were lysed in Triton X-100 lysis buffer (20mM MOPS, pH 7.0, 2mM EGTA, 5mM EDTA, 30mM sodium fluoride, 60mM β-glycerophosphate, 20mM sodium pyrophosphate, 1mM sodium orthovanadate, 1% Triton X-100, 1mM DTT, protease inhibitor cocktail (Roche, #14309200)). Cell lysates (0.5-1.0mg) were then pre-cleaned with protein A/G agarose beads (Santa Cruz, # sc-2003) by incubation for 1 hour with shaking at 4°C followed by centrifugation at 2000 rpm for 3 minute. Lysates were incubated with ethidium bromide (Sigma) as previously described(*12*). Antibody coupling reactions were performed according to the Dynabeads Antibody Coupling Kit (Invitrogen, Cat# 143.11D). Briefly, 10mg Dynabeads M-270 were isolated, washed with buffer and mixed with

primary antibody as indicated. Reactions were then incubated on a roller at 37°C overnight (16-24 hours), washed with buffer and resuspended to a final concentration of 10mg antibody coupled beads/mL. Lysates were then incubated overnight with the coupled antibodies as indicated. In the case of IPs performed with uncoupled antibody, 2µg antibody was used per reaction and lysates were incubated at 4°C for 4 hours with shaking prior to addition of 20µL protein A/G agarose beads. The mixture was then incubated with shaking at 4°C for another 4 hours or overnight prior to washing the lysate-bead precipitate (centrifugation at 2000 rpm for 3 minutes) 4 times in Triton X-100 lysis buffer. Beads were finally precipitated by centrifugation, resuspended in 25µL of 2x loading buffer and boiled at 80°C for 10 minutes for separation the protein and beads. Samples were then analyzed by SDS-PAGE Western blot analysis as described below.

**Immunoblot Analysis**

The cell lines were plated in two wells of a 6-well plate at 250,000 cells/mL 24hours prior to harvesting by trypsinization. Pellets were then flash frozen, briefly sonicated and homogenized in NP40 lysis buffer (50 mM Tris-HCl, 1% NP40, pH 7.4, Sigma, St. Louis, MO), and complete proteinase inhibitor mixture (Roche, Indianapolis, IN). Ten micrograms of each protein extract were boiled in sample buffer, size fractionated by SDS-PAGE, and transferred onto Polyvinylidene Difluoride membrane (GE Healthcare, Piscataway, NJ). The membrane was then incubated overnight at 4ºC in blocking buffer [Tris-buffered saline, 0.1% Tween (TBS-T), 5% nonfat dry milk] and incubated for 4 hours at room temperature with the following: anti-DNA-PKcs mouse monoclonal (1:1000

in blocking buffer, BD Biosciences #610805, San Jose, CA), anti-Ku70 mouse monoclonal (1:1000 in blocking buffer, BD Biosciences #611892), anti-Ku80 rabbit polyclonal (1:1000 in blocking buffer, Cell Signaling Cat #2180S, Danvers, MA), anti-ATR rabbit polyclonal (1:1000 in blocking buffer, Cell Signaling Cat #2790), anti-ERG1/2/3 rabbit polyclonal (1:1000 in blocking buffer, Santa Cruz Biotech # sc-354, Santa Cruz, CA), anti-PARP-1 polyclonal (1:1000 in blocking buffer, Santa Cruz Biotech Cat #sc-8007, Santa Cruz, CA), anti-XRCC4 mouse polyclonal (1:1000 in blocking buffer, BD Biosciences Cat #611506), anti-LACZ (1:1000 in blocking buffer, GenwayBio Cat# 18-732-292258), anti-FLAG rabbit (1:1000 in blocking buffer, Sigma Cat #F7425), anti-V5 mouse monoclonal (1:1000 in blocking buffer, Invitrogen #R-960, Carlsbad, CA) and anti-β-Actin mouse monoclonal antibody (1:10000, Cell Signaling, Cat #: 4967). Following a wash with TBS-T, the blot was incubated with horseradish peroxidase-conjugated secondary antibody and the signals visualized by enhanced chemiluminescence system as described by the manufacturer (GE Healthcare).

**Expression and confirmation of HaloTag fusion proteins**

ETS genes (as described), ERG and sub-domains were cloned into pFN19A vector (Promega, Wisconsin) according to the manufacturer's instructions. Alanine scan mutations were created in the HALO-ETS pFN19A expression vector according to standard Quikchange XL site directed mutagenesis kit protocol (Stratagene). Fusion proteins were expressed in TNT® SP6 High-Yield Wheat Germ Reaction (Promega) based on the manufacturer's protocol. A total of 2.0µl of cell-free reaction containing the HaloTag® fusion protein was

mixed with 8µl HaloTag® Biotin Ligand (final concentration 1µM), and incubated at room temperature for 30 minutes. The biotin-labeled samples were separated on SDS gel and bloted using HRP-streptavidin.

**Interaction assays**

A total of 100µl of cell-free reaction containing ETS proteins or ERG and ERG sub-domain fusion proteins were incubated with 224U of DNA-PKcs (Promega) or PARP1 (Trevigen) in PBS-T (0.1% tween) overnight at 4°C. Ten microliters of HaloLink beads (Promega) were then blocked in BSA overnight at 4°C. After incubation the beads were washed 3 times with PBS, the beads were mixed with Halo-ERG:DNA-PKcs mixture and incubated at RT for 1 hr. Halolink beads were then washed with PBST for 4 times and eluted in SDS loading buffer. Proteins were separated on SDS gel and blotted with anti-DNA PKcs Ab (Santa Cruz). Halo-GUS fusion proteins were used as negative controls.

**Chromatin Immunoprecipitation**

VCaP or RWPE cells were grown in complete medium and ChIP was carried out as previously described(*13*) using antibodies against ERG (Santa Cruz, #sc-354), DNA-PKcs (BD Biosciences, #610805), pDNA-PKcs (T2609) (Santa Cruz, #sc-101664), Ku80 (Cell Signaling, # 2180), Ku70 (BD Biosciences, #611892), PARP-1 (Santa Cruz, #sc-8007), rabbit IgG (Santa Cruz, #sc-2027) and mouse IgG (Santa Cruz, #sc-2025). Briefly, cultured cells were crosslinked with 1% formaledehyde for 10 min and the crosslinking was inactivated by 0.125M glycine for 5 min at room temperature (RT). Cells were then rinsed with 1x PBS

twice and frozen in 1x PBS + 10µl/ml PMSF + phosphatase inhibitor cocktail (Calbiochem) for 30 min. The following steps were performed at 4°C. Cell pellets were resuspended and incubated in cell lysis buffer + 10ul/ml PMSF and protease inhibitor (Roche) for 10 min. Nuclei pellets were spun down at 5,000 x g for 5 min, resuspended in nuclear lysis buffer, and then incubated for another 10min. Chromatin was sonicated to an average length of 600bp and then centrifuged at 14,000 x g for 10min to remove the debris. Supernatants containing chromatin fragments were incubated with agaros/protein A or G beads (Upstate) for 15 min and centrifuged at 5,000xg for 5min to remove the nonspecific binding. To immunoprecipitate protein/chromatin complexes, the supernatants were incubated with 3-5µg of antibody or IgG overnight, then added 50ul of agarose/protein A or G beads and incubated for another 1 hour. Beads were washed twice with 1X dialysis buffer and four times with IP buffer. The antibody/protein/DNA complexes were eluted with 150µl IP elution buffer twice. To reverse the crosslinks, the complexes were incubated in elution buffer + 10µg RNase A and 0.3 M NaCl at 67°C for 4 hours. DNA/proteins were precipitated with ethanol, air-dried, and dissolved in 100µl of TE. Proteins were then digested by proteinase K at 45°C for 1 hour and DNA was purified with QIAGEN PCR column and eluted with 30µl EB. The final ChIP yield was 10-30ng for each antibody. QPCR is described below and primers were either previously described (*11*) or are in **Appendix 1**.

**small RNA interference**

Knockdowns of specific genes were accomplished by RNA interference using commercially available siRNA duplexes for DNA-PKcs, ATM, ATR, PARP-1, XRCC4 and DNA Ligase4 (Dharmacon, Lafayette, CO) or as previously described for ERG(*14*). At least 4 independent siRNAs were screened for knockdown efficiency against each target and the best siRNA was selected, in some cases only one siRNA was identified. Transfections were performed with OptiMEM (Invitrogen) and oligofectamine (Invitrogen) as previously described(*15*).

**Luciferase Reporter Assay**

Luciferase reporter assays were performed as previously descried (*16*). Briefly, RWPE cells were infected with siRNA as indicated 6 hours before the addition of either ERG or control LACZ adenovirus. The PLA1A promoter reporter construct was co-transfected along with pRL-TK (internal control). Twenty four hours post-infection, cells were harvested with passive lysis buffer and luciferase activity was monitored using dual luciferase assay system (Promega, Madison, WI) following manufacturer's instructions. The PLA1A promoter fragment was PCR amplified using a genomic BAC clone as template with the forward primer (5'- CCCCATTGACTTGCCTAGAA) and reverse primer (5'- GGCTTTTAGGGGATCTTCCA) and subcloned into pGL4.14 vector (Promega) using Xho1 and Hind3 enzymes.

**Quantitative Real-Time PCR Assays**

Total RNA was isolated from VCAP, RWPE or PrEC cells that were transfected with siRNA as indicated (Qiagen). Quantitative PCR (QPCR) was performed using SYBR Green dye on an Applied Biosystems 7300 Real Time PCR system (Applied Biosystems, Foster City, CA) as described (5). Briefly, 1μg of total RNA was reverse transcribed into cDNA using SuperScript III (Invitrogen, Carlsbad, CA) in the presence of random primers (Invitrogen). All reactions were performed in triplicate with SYBR Green Master Mix (Applied Biosystems) plus 25ng of both the forward and reverse primer according to the manufacturer's recommended thermocycling conditions, and then subjected to melt curve analysis. Threshold levels for each experiment were set during the exponential phase of the QPCR reaction using Sequence Detection Software version 1.2.2 (Applied Biosystems). The relative quantity of the target gene was completed for each sample using the ΔΔCt method by the comparing mean Ct of the gene to the average Ct of the housekeeping gene, β-Actin(*17*). All oligonucleotide primers were synthesized by Integrated DNA Technologies (Coralville, IA). The primer sequences for the transcript analyzed were either previously described (*11*) or are provided in **Appendix 1**. Targeted DNA damage pathway arrays were obtained from SA biosciences and run according to the manufacturer's protocol (Qiagen).

**Gene Expression Array Analysis**

Expression profiling was performed using the Agilent Whole Human Genome Oligo Microarray (Santa Clara, CA) according to the manufacturer's

protocol and described previously (*18*). For all hybridizations, the reference was VCaP treated with control siRNA for 48hrs. A total of 4 arrays were performed per sample such that all hybridizations were performed in duplicate with duplicate dye flips. Data was filtered to include only features with significant differential expression (Log ratio, $P < .01$) in all hybridizations and, after correction for the dye flip, two-fold average over- or under-expression (Log ratio). Over- and under-expressed signatures for the VCaP siRNA experiment were generated by identifying differential features common among both 2 independent DNA-PKcs siRNA sets and 2 independent PARP1 siRNA sets.

**Molecular concept maps**

The expression signatures in common between both DNA-PKcs and PARP1 siRNA were uploaded into the Oncomine Concepts Map (OCM, www.oncomine.org) as molecular concepts(*19*), using all features on the Agilent Whole Human Genome Oligo Microarray as the null set. Data are reported as RMA-normalized fluorescent intensities and analysis was completed as previously described(*11*).

**Basement Membrane Matrix Invasion Assays**

For invasion assays, the prostate cell lines RWPE-1 and VCaP were transfected with siRNA or negative controls as indicated. NU7026 (Sigma) and NU1025 (Calbiochem) were dissolved in DMSO and stored at -20C in the dark. Forty-eight hours post-transfection/transduction, cells were seeded onto the

basement membrane matrix (EC matrix, Chemicon, Temecula, CA) in the chamber insert with 8.0μM pores of a 24-well culture plate in serum free media. Cells were attracted to the lower chamber by the addition of complete media as a chemoattractant. After 48 hours incubation at 37˚C with 5% $CO_2$, the non-invading cells and EC matrix were gently removed with a cotton swab. Invasive cells, which were located on the lower side of the membrane, were stained with crystal violet, air dried and photographed. To quantify the relative number of invaded cells, colorimetric assays were performed by treating the inserts with 150μl of 10% acetic acid (v/v) and measuring absorbance of each condition at 560nm using a spectrophotometer (GE Healthcare).

**CAM assays**

The CAM assay was performed as described previously (*20*). Briefly, fertilized eggs were incubated in a rotary humidified incubator at 38°C with for 10 days. After releasing the CAM by applying mild amount of low pressure to the hole over the air sac and cutting a square 1-$cm^2$ window encompassing a second hole near the allantoic vein, cultured human cells that had been pre-treated with siRNA as indicated were detached by trypsinization and re-suspended in complete medium prior to implantation of 2 x $10^6$ cells adjacent to the mesenchyme in each egg. Cells used for invasion assays were labeled with green fluorescent microspheres as previously described (*21*). The windows were subsequently, sealed and the eggs were returned to a stationary incubator. For invasion and intravasation experiments, implanted eggs were treated 6 hours after cell

inoculation and either the upper or lower CAM was isolated after 72 hours, respectively. Invasion assays were processed, stained and scored as previously described (*21*). For metastasis and tumor growth assays, cells were then treated as indicated with 40mg/kg body weight Olaparib which was administered on day13, 15 and 17. Day 18 embryos were sacrificed. Tumors were then excised and weighed. No detectable metastatic cells were present in the livers of animals xenografted with either 22RV1 or VCaP cells suggesting that these cell lines are either not metastatic or that they do not metastasize to the liver. After incubation of either the metastasis or intravasation assays, the extra-embryonic tumor and CAM, as well as the embryonic liver were harvested and analyzed for the presence of tumor cells by quantitative human *alu*-specific PCR. Genomic DNA from lower CAM and lungs were prepared using Puregene DNA purification system (Qiagen). Quantification of human cells in the extracted DNA was done as described (*22*). Fluorogenic TaqMan qPCR probes were applied as above, and DNA copy numbers were quantified.

## RESULTS

### Identification of ERG interacting proteins by mass spectrometry

To identify ERG-interacting proteins that may serve as rational therapeutic targets and explore the mechanism by which ETS gene fusions mediate their effects, we performed mass spectrometric (MS) analysis of proteins that interact with the most prevalent ETS gene fusion product, ERG (encoded from *TMPRSS2* exon 1 fused to *ERG* exon 2 (*5*)). VCaP prostate

cancer cells (which harbor a *TMPRSS2*:*ERG* rearrangement) or human embryonic kidney cells (HEK) 293 cells were infected with either adenoviral V5- or FLAG- epitope-tagged ERG expression vectors, respectively. Immunoprecipitation (IP) was completed in 8 biological replicates to isolate protein-protein interactions as described by schematic (**Figure 2.1)**. As expected, the interaction bait, ERG, was the top scoring protein identified in the pull-down with 64.4% coverage with 17 tryptic peptides scanned over 500 times (**Figure 2.2A, Table 2.1**). Interestingly, three of the next four interacting proteins of high confidence and high sequence coverage identified were components of the DNA-dependent protein kinase complex and included the large catalytic subunit of a phosphatidylinositol 3/4 (PI3/4)-kinase called DNA-dependent protein kinase (DNA-PKcs) (10% coverage) and its known interacting subunits Ku70 (26% coverage) and Ku80 (34% coverage) (**Figure 2.2A, Table 2.1**). Interactions were confirmed with an independent antibody (**Figure 2.3A**) and IPs performed from VCaP cells demonstrated an endogenous association that occurs in the absence of ectopic overexpression (**Figure 2.2B and Figure 2.3B**).

To identify additional proteins participating in the ERG:DNAPK complex, we assessed our list of ERG interactors for other proteins known to interact with DNA-PKcs, Ku70, or Ku80 and identified two peptides for poly-(ADP-ribose) polymerase (PARP1): VVSEDFLQDVSASTK and QQVPSGESAILDR. Importantly, we demonstrated that PARP1 endogenously associated with ERG in VCaP cells (**Figure 2.2B**). We then

performed reverse IPs using antibodies against DNA-PKcs, PARP1 and Ku80 and showed that each antibody was able to detect ERG-V5 protein (**Figure 2.3C**). To detect the PARP1:ERG interaction with the endogenous *TMPRSS2*:*ERG* gene fusion product, we used agarose-coupled PARP1 antibody to perform the IP-Western, which confirmed that PARP1 interacts with the gene fusion product in an endogenous setting (**Figure 2.3D**).

As DNA-PKcs only binds with Ku70 and Ku80 in the presence of DNA (*23*), we tested the dependence of the ERG:PARP1 and ERG:DNA-PKcs interactions on intact DNA by performing the IP in the presence of 100μM ethidium bromide. This treatment disrupted the interaction between ERG, Ku70 and Ku80, but not the interaction between ERG and either PARP1 or DNA-PKcs, demonstrating that the ERG:PARP1 and ERG:DNA-PKcs interactions are DNA-independent (**Figure 2.2B**). As a control, we tested whether ERG would bind another PI3/4 kinase family member, ATR, or another protein known to interact with the DNAPK complex, XRCC4. Consistent with our IP-MS data, we were unable to detect an interaction between ATR or XRCC4 and ERG by IP-Western blot analysis (**Figure 2.3B and Figure 2.2B,** respectively).

We next assessed whether the ERG:PARP1 and ERG:DNA-PKcs interactions occur in human prostate cancer tissues. ERG-IP showed enrichment for DNA-PKcs, Ku70, Ku80 and PARP1 in ERG gene fusion-positive, but not in ETS gene fusion-negative prostate cancer tissues (**Figure 2.2C, Figure 2.3E**). Interestingly, the lack of detectable ERG:PARP1

interaction in tissue without ETS gene rearrangement is likely due to the near absent ERG expression in rearrangement-negative prostate cancer (*24*), as when overexpressed, wild type (WT) ERG interacts with PARP1 in cell lines that do not harbor the translocation (**Figure 2.3F**). Additional IP-Westerns were performed to test the dependence of the ERG:PARP1 interaction on DNA in human prostate cancer tissues. Importantly, the interaction occurred independent of DNA in all three independent human tissues (**Figure 2.3G**).

Next, we sought to map the interactions and created a series of flag tagged ERG expression vectors with tiling deletions including: the N-terminus (deletion of AA: 47-115, predicted molecular weight 44.6 kDa), pointed domain (AA: 115-197, 43.4 kDa), the middle amino acids (197-310, 41.6 kDa), the ETS domain (AA: 310-393, 43.7 kDa) or the C-terminus (AA: 393-479, 43.6 kDa) and labeled the constructs ΔN, ΔP, ΔM, ΔE, ΔC, respectively as depicted in **Figure 2.2D**. IP following transient transfection demonstrated that the interactions between ERG, DNA-PKcs, Ku70, Ku80 and PARP1 occurred in the C-terminal half of the ERG protein (**Figure 2.2E**). To further map the ERG:PARP1 and ERG:DNA-PKcs interactions and to confirm that both PARP1 and DNA-PKcs interact with other ETS family member proteins, we performed IP-Western blot analysis in HEK293 cells transfected with either *ERG-FLAG*, *ETS1-FLAG*, *SPI1-FLAG* or *ETV1-FLAG* expression vectors, which were selected for their sequence relationship to ERG (**Figure 2.4A**). In all four experiments, pull downs confirmed the interactions (**Figure 2.3F, 2.4B, 2.4C and 2.4D**). We then created N-terminal halo-tagged

expression vectors for *in vitro* purification of these ETS genes. Subsequent IP-westerns demonstrated that all 4 of these proteins bind directly to DNA-PKcs (**Figure 2.4E**). Given the sequence alignment of these 4 ETS proteins and the large tiling deletion data, our data suggested that the interactions occur through the ETS DNA binding domain.

To definitively map the ERG:DNA-PKcs and ERG:PARP1 interactions, we utilized HALO-tagged WT ERG and six individual HALO-tagged fragments spanning the entire ERG protein (*25*). As expected, IP-Western blot demonstrated that the ERG:DNA-PKcs interaction occurred through the ETS DNA binding domain. To further map the interaction between ERG and DNA-PKcs, we utilized a series of three HALO-tagged fragments that tiled the ETS domain, which localized the interaction to the final 28 amino acids of the ETS domain (**Figure 2.2F and 2.4F**). Importantly, although the crystal structure of the ETS domain from ERG has not yet been reported, the crystal structure of another ETS factor that we demonstrated interacts with DNA-PKcs, ETS1, has been published (*26*). Based on homology with other interacting ETS proteins and structural information, we predicted that the interaction was dependent on the amino acids, YYDKN. By site directed mutagenesis of each residue to alanine, we demonstrated that the Y373A mutant was unable to precipitate DNA-PKcs suggesting that this interaction is mediated by Tyrosine 373 (**Figure 2.2F and 2.4H**). Analysis of the ETS1 structure shows that Y373 is adjacent to the arginine residues that fit

into the DNA groove and that Y373 is accessible to potential interacting proteins (**Figure 2.2G**).

After demonstrating that ERG interacts with DNA-PKcs directly through amino acid Y373, we sought to map the ERG:PARP1 interaction. However, purified ERG was only able to interact with purified PARP1 in the absence of ethidium bromide (**Figure 2.4G**). Because the interaction occurred in cells independent of ethidium bromide, this suggests that the ERG:PARP1 interaction is mediated by other proteins. This is consistent with the results from our IP-MS experiment in which few PARP1 peptides were identified suggesting that the ERG:PARP1 interaction is mediated by an intermediate protein such as DNA-PKcs.

**PARP1 and DNA-PKcs are required for ERG-mediated transcription**

Given that the interaction of DNA-PKcs and PARP1 with ERG occurs through the ETS domain, we hypothesized that both PARP1 and DNA-PKcs function as co-regulators of ERG transcriptional activity. Thus, we performed chromatin immunoprecipitation (ChIP) assays in VCaP cells and assessed enrichment of known ERG-targets including the *PLA1A* promoter and the *FKBP5*, *PSA* and *TMPRSS2* enhancers. These experiments demonstrated that ERG, DNA-PKcs, activated DNA-PKcs (assessed by T2609 phosphorylation), Ku70, Ku80 and PARP1 bind to these sites, but not to the negative control gene *KIAA0066* (*11*) (**Figure 2.5A**). Interestingly, this enrichment was disrupted by ERG siRNA (**Figure 2.6A and Figure 2.5B**) and ChIP

performed in ERG-overexpressing RWPE cells supporting a model in which ERG recruits PARP1 and DNA-PKcs to specific genomic loci during transcription (**Figure 2.6B**). Consistent with this hypothesis, serial ChIP reactions (ERG, then PARP1 or DNA-PKcs) demonstrated that an ERG:PARP1 complex and an ERG:DNA-PKcs complex are both present at ERG-regulated loci (**Figure 2.5C**). Although it was not possible to perform re-ChIP experiments with the PARP1 and DNA-PKcs antibodies, IP-Western blot analysis confirmed that PARP1 and DNA-PKcs interact in a DNA-independent manner in VCaP cells (**Figure 2.5D**). Likewise, this experiment suggests that DNA-PKcs binding to ERG does not disrupt the ERG:DNA interaction.

To test whether DNA-PKcs and PARP1 are required for ERG-mediated transcriptional activation, we constructed a *PLA1A* promoter reporter. Transfection of the reporter into RWPE cells treated with either LACZ or ERG adenovirus and siRNA (**Figure 2.5E**) indicated that both DNA-PKcs ($p = 1.99 \times 10^{-3}$) and PARP1 ($p = 2.37 \times 10^{-3}$) are required for ERG-induced activation of *PLA1A* (**Figure 2.5E**) in RWPE cells. In contrast, inhibition of the related PI3/4-like kinase, ATM, had no significant effect on ERG activity.

While ATM and ATR repair DNA strand breaks through different pathways, DNA-PKcs is specifically required for NHEJ (*27*). In this process, DNA-PKcs, Ku70 and Ku80 form a complex on the broken DNA end which facilitates DNA end processing and re-joining in a multi-step procedure that

requires the XRCC4/DNA Ligase IV complex. In fact, XRCC4 and DNA Ligase IV are both independently required for execution of NHEJ in mammalian cells, as targeted inactivation of either gene leads to NHEJ defects in mouse cells (*28, 29*). Subsequently, we used siRNA to knockdown XRCC4 (**Figure 2.5E and 2.7A**) to evaluate the necessity of effective execution of NHEJ for ERG-induced transcriptional activation of the *PLA1A* promoter. Because knockdown of XRCC4 had no effect on ERG activity, the experiment further suggests a NHEJ-independent role for DNA-PKcs in ERG-mediated transcription (**Figure 2.5E**).

Given the importance of PARP1 and DNA-PKcs for ERG-mediated transcription, we sought to explore the global effects of inhibiting PARP1 and DNA-PKcs on the ERG-transcriptome. To do this, we used Agilent Whole Genome Oligo Expression Arrays to profile RNA from VCaP cells treated with either DNA-PKcs or PARP1 siRNA (knockdown confirmed in **Figure 2.7B**). Our analysis revealed 50 and 252 unique features that were greater than two-fold down- and upregulated, respectively, in both the PARP1 and DNA-PKcs siRNA-treated samples (**Appendix 2**). Venn diagram analysis was used to show the overlap of differential gene sets to genes regulated by ERG in VCaP cells (*11*) ($p < 0.0001$ for all interactions unless indicated, hypergeometric test) (**Figure 2.5F, 2.5G**). To then understand how this gene signature is related to existing signatures, we uploaded our expression signature into Oncomine Concepts Map (OCM) (*18, 19*) to identify human tissue gene signatures that are enriched for genes upregulated by DNA-PKcs and PARP1 siRNA in VCaP cells (genes repressed by PARP1 and DNA-PKcs). This provided unbiased validation that the

tissue-based gene signatures most highly enriched with our gene set were the genes repressed in ETS-positive as compared to ETS-negative prostate cancer; Xiaojun *et. al.* (*18*) [OR = 3.08, p =1.40 x $10^{-15}$], Yang *et. al.* (*18*) [OR = 2.91, p =3.30 x $10^{-10}$] and Lapointe *et. al.* (*30*) [OR = 3.33, p =2.30 x $10^{-6}$] (**Figure 2.6C**). Interestingly, the gene signature also showed significant overlap with the set of genes repressed in metastatic as compared to localized prostate cancer suggesting that repression of these genes is important for prostate cancer progression; (*31*) [OR = 2.99, p =1.5 x $10^{-10}$], (*32*) [OR = 3.31, p =1.50 x $10^{-6}$] (**Figure 2.6C**). Treatment of VCaP cells with siRNA confirmed gene expression changes as predicted by the gene expression arrays (**Figure 2.6D**) as did treatment with either the small molecule DNA-PKcs kinase inhibitor, NU7026, or the small molecule PARP1 inhibitor, Olaparib (**Figure 2.6E**). Analysis of siRNA treated RWPE-ETV1 cells (**Figure 2.8A**) confirmed that DNA-PKcs and PARP1 regulated ETV1 transcriptional activity as well (**Figure 2.6F**). Taken together these data suggest that PARP1 and DNA-PKcs play a role in modulating transcriptional activity of a number of ETS target genes, some of which are differentially expressed between localized and metastatic disease.

**PARP1 and DNA-PKcs are required for ERG-mediated cell invasion, intravasation and metastasis**

Inhibition of PARP1 and DNA-PKcs altered ERG-transcriptional activity of several progression-associated genes such as *EZH2*, here we tested the role of these enzymes in ERG-induced cell invasion. Both DNA-PKcs

siRNA (**Figure 2.7A and 2.7B**) and NU7026 attenuated invasion in RWPE cells transduced with ERG (**Figure 2.9A**) and VCaP cells (**Figure 2.9B**) [$p < 0.01$ for DNA-PKcs siRNA or NU7026 > 10μM]. Likewise, we found that treatment with either PARP1 siRNA (**Figure 2.7A and 2.7B**) or Olaparib led to a significant reduction in ERG-driven RWPE and VCaP cells invasion (**Figure 2.9A and 2.9B**) [$p < 0.05$ for all PARP1 siRNA or Olaparib treatments]. As with our analysis of ERG-mediated transcription, knockdown of either ATM or XRCC4 did not have an effect on ERG-mediated invasion (**Figure 2.9A and 2.9B**). Treatment of stable RWPE cells stably overexpressing ETV1 (*33*) with PARP1 or DNA-PKcs siRNA (**Figure 2.8A**) or small molecule inhibitors also led to a significant reduction in invasion ($p < 0.01$ for all PARP1 or DNA-PKcs treatments) (**Figure 2.9C**). Importantly, however, invasion of two negative control models, the ETS rearrangement-negative cell line PC3 and RWPE cells overexpressing an alternative prostate cancer gene fusion, *SLC45A3-BRAF*, was not affected by inhibiting either enzyme (**Figure 2.9D and Figure 2.8B**). To determine if the observed effect on cell invasion was due to cytotoxity, we performed chemosensitivity assays with both Olaparib and NU7026. Neither Olaparib nor NU7026 had an effect on the *in vitro* cell proliferation rate of any of the cell lines tested, suggesting that the reduction in cell invasion is not due to changes in cell proliferation (**Figure 2.10A**). In fact, the $EC_{50}$ for both drugs was well beyond the dose shown to block transcription and invasion (**Figure 2.10B and 2.10C**).

We next sought to define the role of PARP1 in ERG-mediated invasion and intravasation *in vivo*. To do this, we implanted cells onto the upper *chorioallantoic membrane* (CAM) of a fertilized chicken embryo and analyzed the relative number of cells that invade and intravasate into the vasculature of the lower CAM three days after implantation (*34*). In this assay, Olaparib blocked both ERG-mediated invasion and intravasation ($p < 0.1$) (**Figure 2.11A and 2.11B**). Because increased expression of *EZH2* alone is sufficient to drive metastasis in several different cell systems, we monitored *EZH2* mRNA expression and found that *EZH2* expression was downregulated following either PARP1 or DNA-PKcs inhibition (**Figure 2.12A, 2.12B, 2.12C and 2.12D**), suggesting that mechanistically, PARP inhibition disrupts ERG-mediated invasion and intravasation by inhibiting ERG-mediated transcriptional activation of invasion-associated genes such as *EZH2*. More importantly, however, our models suggested that therapeutic disruption of either ERG-interacting enzymes (PARP1 or DNA-PKcs) inhibits the metastatic spread of prostate cancers harboring ETS gene fusions.

To test this postulate, we analyzed the metastatic potential of an ETS positive (LNCaP) and an ETS negative (PC3) cell line. As shown in **Figure 2.11C**, Olaparib treatment blocked the formation of liver metastases from LNCaP ($p = 0.01$), but not PC3 cells. Importantly, we also noticed that over the extended treatment period the ETS-positive tumors were significantly smaller than the ETS-negative tumors (**Figure 2.11D**) with $p < 0.05$ for VCaP and $p < 0.01$ for LNCaP. This suggests that PARP1 could play a role in the

long term maintenance of ETS-positive cancer cell survival. Because it appeared that the long term survival of ETS overexpressing tumors can be diminished by treatment with Olaparib, we sought to compare the magnitude of effect to that of a clinically validated model. Therefore, we xenografted HCC1937 (BRCA1 mutant) and MDA-MB-231 (BRCA1/2 WT), following Olaparib treatment a significant effect was observed on the BRCA1 mutant HCC1937 tumors, while no measurable effect was observed in MDA-MB-231 tumors. Surprisingly, the magnitude of effect observed in the HCC1937 cells was equivalent to the magnitude of effect observed in the two ETS-positive cell line xenografts (**Figure 2.11D**).

## DISCUSSION

As my discovery of the ETS:DNA-PKcs and ETS:PARP1 interactions serves as the basis for the therapeutics work presented in the subsequent chapters, I will review the functional background for these enzymes before delving into a detailed analysis and discussion about their importance for ETS-mediated oncogenic activities in the subsequent chapters.

### Poly(ADP) ribose Polymerase 1

Poly(ADP) ribose Polymerases are a family of several enzymes with unique structure, function and cellular distribution which have the ability to append ADP-ribose units onto to either aspartic or glutamic acid residues in target proteins. ADP-ribose chains are built using NAD+ as a substrate (*35*) in a

dynamic process (*36*) and can extend from a few to over 200 units long and different PARPs are thought to add ADP-ribose groups in different branching/elongation ratios (*37*). Poly(ADP)-ribose side chains carry a highly negative charge, which is known to alter a number of cellular processes including DNA replication, transcription, and DNA repair (reviewed in (*38*)) through the modification of several proteins including topoisomerase I (*39, 40*), RNA polymerase II (*41, 42*), DNA polymerases (*43*), p53 (*42*) and DNA-PKcs (*44*). Once a poly(ADP)-ribose side chain is completed, one of two enzymes can catabolize the protein modification, poly(ADP)ribose gylcohydrolase or ADP-ribosyl protein lyase (*45*).

Of the family of PARP enzymes, PARP1 is a functionally complex molecule involved in both transcriptional regulation (reviewed in (*46*)) and multiple DNA repair pathways (reviewed in (*47*)). PARP1 was originally discovered as an enzyme that uses NAD+ as a substrate (*48*). PARP1 knockout mice are viable with almost no significant phenotypic alterations (*49, 50*), but give several key insights into the normal role of PARP1 in various cellular processes. For example, although PARP is specifically cleaved during apoptosis, knockout cells undergo apoptosis indicating that PARP1 is dispensable for this cell death pathway (*50*). More importantly, however, PARP1 mutant fibroblasts and splenocytes were observed to have an elevated frequency of spontaneous sister chromatid exchanges and elevated micronuclei formation after treatment with various genotoxic agents (*49, 50*), consistent with the role of PARP1 in the maintenance of genomic integrity.

In line with its diverse number of cellular functions, PARP1 contains many conserved domains. The N-terminus contains a DNA binding domain consisting of two zinc fingers. This domain allows PARP1 to bind both single strand DNA and double strand DNA lesions (*51*) and also contains a caspase cleavage site (DEVD) and nuclear localization sequence. The middle of PARP1 contains a breast cancer susceptibility protein C terminus motif (BRCT), which is responsible for many protein:protein interaction including with the Ku heterodimer complex (*52*). The PARP1 catalytic domain is in the c-terminal half of the protein (*53-55*) and contains the NAD+ binding site (*56*) and a 50 amino acid "PARP signature" domain contained in all vertebrate PARP homologues identified thus far (*55*).

As mentioned, PARP1 has many interactions with proteins involved in transcriptional regulation including RNA polymerase II (*41, 42*) and, using ChIP-ChIP experiments, PARP1 has been shown to be enriched in the promoters of as many as 90% of transcribed genes (*57*). The promoter enrichment further suggests a role of PARP1 in transcription and, interestingly, the enrichment is highly associated with loss of histone H1 and is most enriched -250 base pairs from the transcription start sites (*57*). Early microarray based studies of PARP1 knockout mice indicated that as much as 3.5% of the transcriptome is regulated by the enzyme with about 60-70% being positively regulated (*58*). The discordance between the percent of differentially transcribed genes with promoter occupation is an area of active investigation, and suggests that some compensatory

mechanisms may play a role in the maintenance of transcription in the absence of PARP1.

In addition to its role in transcriptional regulation, PARP1 is often noted in the literature as being essential for the execution of base excision repair (BER), however, recent arguments have suggested that it is only required for efficient BER (*47*). In the presence of DNA nicks, PARP1 binds to DNA through its zinc finger domain, forms homodimers and poly(ADP)-ribosylates several BER factors including endonucleases (*59-61*). PARP1 can also interact with and recruit other BER factors to DNA nicks including DNA ligase III, XRCC1 and DNA polymerase β (*62-66*). Consistent with the mechanistic role of PARP1 in BER, PARP1 mutant, null and dominant negative treated cells are more sensitive to alykylating-agent, UV radiation and topoisomerase inhibitor-induced cytotoxicity than controls (*49, 50, 67-69*). However, careful kinetic studies have shown that PARP1 accelerates BER (*70, 71*). Consistent with this observation is the fact that DNA polymerase β (*72*) and XRCC1 knockouts are embryonic lethal (*73*), while PARP1 knockout mice are viable (*49, 50*).

Beyond its kinetic role in the base excision repair pathway, PARP1 has also been implicated to play a critical role in alternative end-joining (AEJ) (*74*), a backup pathway enacted in the absence of functional NHEJ. Evidence for the AEJ pathway came from analysis of Ku-/- cells which are unable to initiate NHEJ. A reporter plasmid, which contains a single I-SCEI restriction digest site, was integrated into the genomes of the Ku deficient cells such that transient transfection of the I-SCEI enzyme could create a single DNA double strand break.

Using this model, it was shown that Ku deficient cell lines have slow DNA double strand break repair that results in abnormally large amounts of sequence loss surrounding the break point as compared to WT cells with intact NHEJ – indicating the presence of an alternative DNA DSB repair pathway. Using this model system and a first generation PARP inhibitor that "freezes" inactive PARP proteins to DNA (*75*), Mansour *et. al.* demonstrated that PARP1 activity is required for DNA resection at the breakpoint (*74*). Consequently, AEJ was completely disrupted by PARP inhibition. The importance of the DNA-PKcs:PARP1 complex for the execution of AEJ is unknown.

**PARP inhibitors in BRCA mutant cancer cells**

In 2005, two groups demonstrated that *BRCA1/2* mutant cancer cells, which are unable to effectively execute a DNA double strand repair process called homologous recombination, are highly sensitive to PARP inhibition (*76, 77*). Much excitement was generated over these studies, and synthetic lethality – a combination of disruptive events in two or more proteins that leads to cell death – was revitalized as a paradigm in cancer research. In this case, the synthetic lethality is between a mutation in either *BRCA1* or *BRCA2* and induced by chemical disruption of a second protein, PARP1. The initial clinical trials using a PARP inhibitor called Olaparib demonstrated early efficacy in a phase I clinical trial, with limited side effects (*78*). However, the precise mechanism of how PARP inhibition leads to cell death in *BRCA1/2* mutant cancer cells remains unknown (questions reviewed in (*47*)). Given the complexity and diversity of

cellular processes in which PARP enzymes participate, the original model of *BRCA1/2* mutant-mediated PARP susceptibility is most likely overly simplistic and ongoing research is focused on identifying additional mechanistic details of the induced-sensitivity.

**DNA-dependent Protein Kinase (DNA-PKcs)**

As with the functionally complex protein, PARP1, DNA-PKcs functions in both transcriptional regulation and DNA repair. DNA-PKcs functions both independently and as part of a holoenzyme complex called the "DNAPK" complex, which consists of the DNA-PKcs (catalytic subunit), Ku70 and K80. This complex only assembles on DNA (*79*). The gene encoding DNA-PKcs resides on 8q11 and is transcribed and translated into a protein 4128 amino acids long (~429kDa). DNA-PKcs belongs to the phosphatidylinositol-3-OH kinase (PI(3)K) family of protein kinases, which also includes the DNA repair enzymes Ataxia Telangiectasia Mutated (ATM) and Ataxia Telangiectasia and Rad3-related (ATR) (*80*). DNA-PKcs has several conserved domains including two phosphorylation clusters called ABCDE and PQR, a kinase domain, FAT domain, FATC domain, many HEAT repeat elements and a Ku interaction domain. Because of the high molecular weight of the protein, it was only recently crystallized (*81*). This crystal structure demonstrated that HEAT repeats facilitate a subtle bending that allows the enzyme to form a hollow circular structure, providing a flexible "cradle" to promote DNA double strand break repair. The

kinase domain is located on the outside of the circle (*81*), and has been shown to phosphorylate a number of targets including itself (reviewed in (*82*)).

From the many roles of DNA-PKcs, perhaps the best characterized is its role in Non-homologous End Joining (NHEJ), a process which can repair DNA double strand breaks in all phases of the cell cycle. Consistent with this, small molecule inhibitors or inactivating point mutations in DNA-PKcs lead to radiation sensitivity and defective DNA double strand break repair (*83-85*). NHEJ is thought two occur in approximately three gross steps: recognition of the DNA double strand break, processing of the DNA ends to create ligatable termini, and ligation of the DNA ends. After a DNA double strand break is formed, the two severed DNA ends are first recognized by the Ku heterodimer, which then recruits DNA-PKcs to the each DNA end (*86, 87*). Once the holoenzyme is formed and positioned over the end, the DNA-PK complex phosphorylates histone H2A.X on Ser139, which then leads to the recruitment of several DNA end processing enzymes including artemis, polynucleotide kinase 3'-phosphatase, DNA polymerases μ and λ and/or the Mre11-Rad50-Nbs1 (MRN) endo/exonuclease complex (reviewed in (*88*)). After processing, the DNA is ligated together by the XRCC4/DNA Ligase 4 complex (*89*) bringing the two DNA-PKcs enzymes close enough to phosphorylate one another, which causes a release from the now ligated DNA ends. Importantly, this process can lead to DNA sequence changes at the sites of DNA damage as a result of end processing.

In addition to its role in NHEJ, a recent report demonstrated that DNA-PKcs also plays a role in transcription (*90*). In this study, Wong *et. al.*

demonstrated that DNA-PKcs associates with and phosphorylates the transcription factor USF-1 following insulin stimulation *in vitro* or animal feeding *in vivo* (*90*). Interestingly, activation of the *FAS* promoter by the USF-1:DNA-PKcs complex was associated with transient DNA damage (*90*). Likewise, DNA-PKcs had previously been shown to associate with promoters regulated by several nuclear hormone receptors including the estrogen receptor (*91*). Despite these initial observations that DNA-PKcs plays a role in some transcriptional activation pathways, the precise mechanism of this role has not been established.

## CONCLUSION

Taken together, this chapter demonstrates that DNA-PKcs and PARP1 expression and activity are critical for ETS-mediated transcription, cell invasion, intravasation, and *in vivo* metastasis. While both enzymes are functionally complex, they have both been indicated to have roles in both transcription and DNA repair. Moving forward, at the time of this research, several PARP1 inhibitors were in development and only one had been successful in a phase I clinical trial. In contrast, DNA-PKcs inhibitors were still in their infancy and effective *in vitro* at only high micromolar doses. Consistent with this observation, we ran a pilot study with NU7026 to determine xenograft efficacy, but found general toxicity and off-target effects in ETS negative cell lines (Data not shown). Because of drug efficacy, we therefore focused the next two chapters of this thesis on studies with PARP inhibitors.

**FIGURES**

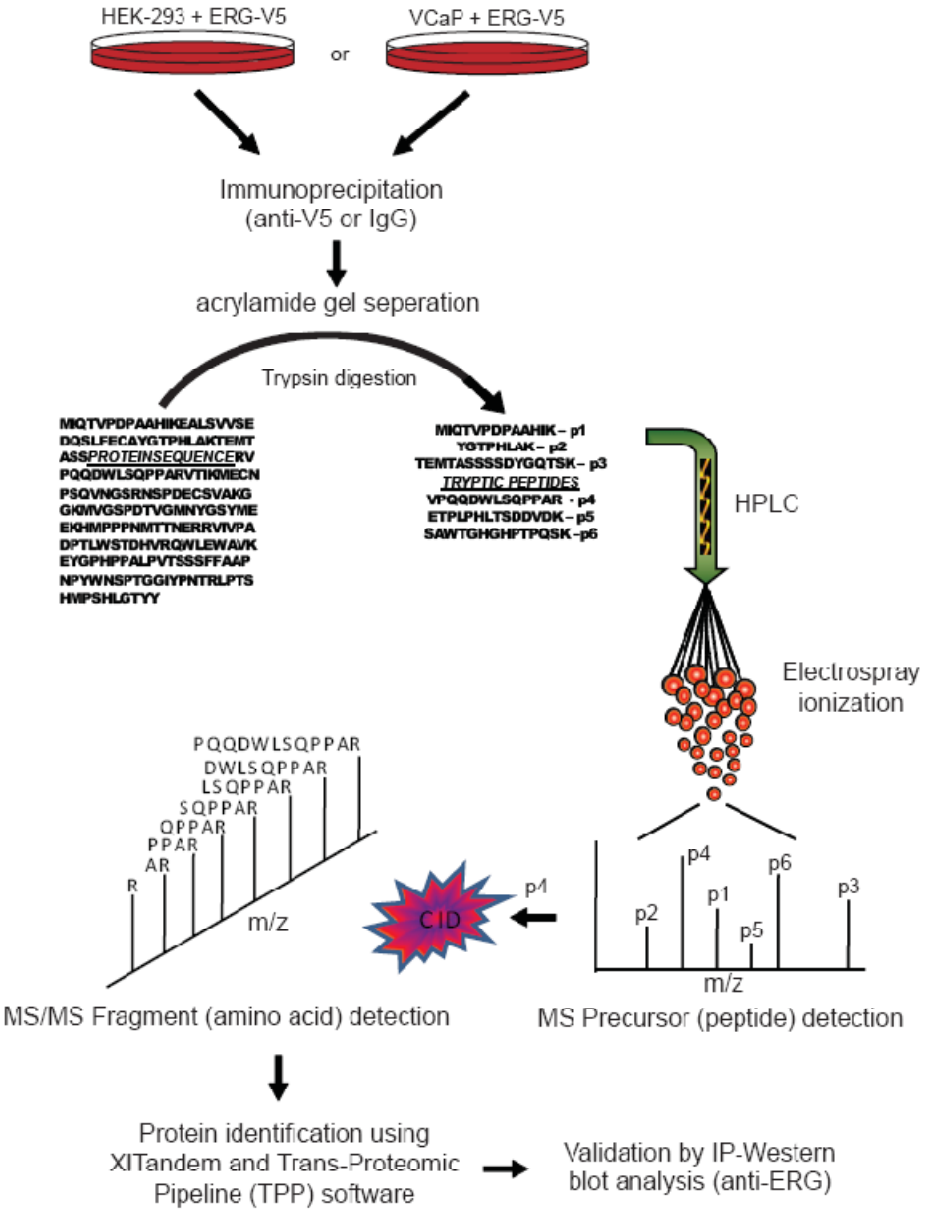



**Figure 2.1: Schematic representation of mass spectrometry workflow used for identification of ERG interacting proteins.**

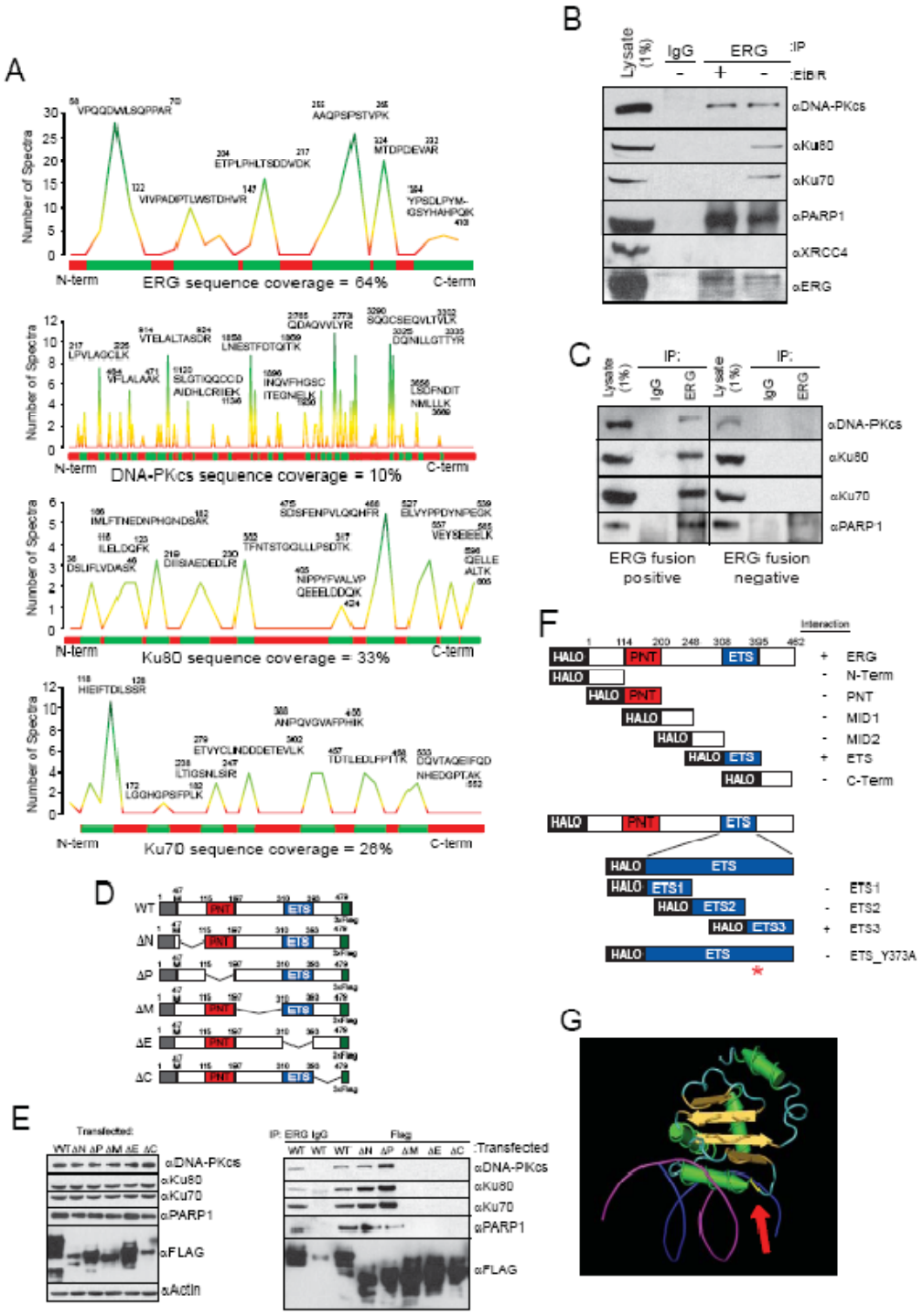



**Figure 2.2: The *TMPRSS-ERG* gene fusion product interacts with PARP1 and the DNA-PK complex.**
**(A)** MS analysis of proteins interacting with ERG. Histograms show peptide coverage of ERG, DNA-PKcs, Ku70 and Ku80.

**(B)** ERG, DNA-PKcs, PARP1 but not Ku70 or Ku80, interact independent of DNA. IP performed from VCaP cells which naturally harbor the ERG translocation.
**(C)** ERG, DNA-PKcs, PARP1, Ku70 and Ku80 associate in ERG gene fusion positive human prostate cancer tissues. Representative ERG positive and negative prostate cancers shown of three pairs of tissues.
**(D)** Schematic of *TMPRSS2-ERG* gene fusion tiling deletion expression vectors.
**(E)** IP of DNA-PKcs, Ku70, Ku80 and PARP1 from HEK293 cells transfected with ERG expression vectors depicted in (D). Input Western is shown on the left, IP-Western shown on the right. All IPs were performed with FLAG antibody unless otherwise indicated.
**(F)** Schematic representation of halo-tagged ERG fragment vectors. The construct were transcribed using wheat germ extracts and halo-tagged protein was purified. Proteins were then incubated with purified DNA-PKcs and IP-westerns performed. Fragments able to IP DNA-PKcs are indicated with a '+'.
**(G)** ETS1:DNA crystal from (*26*) used to demonstrate physical location of Tyrosine373 (from ERG) relative to the DNA binding residues.
1% of the total cell lysate used for IP was added to the input lane.
Representative experiments are shown.

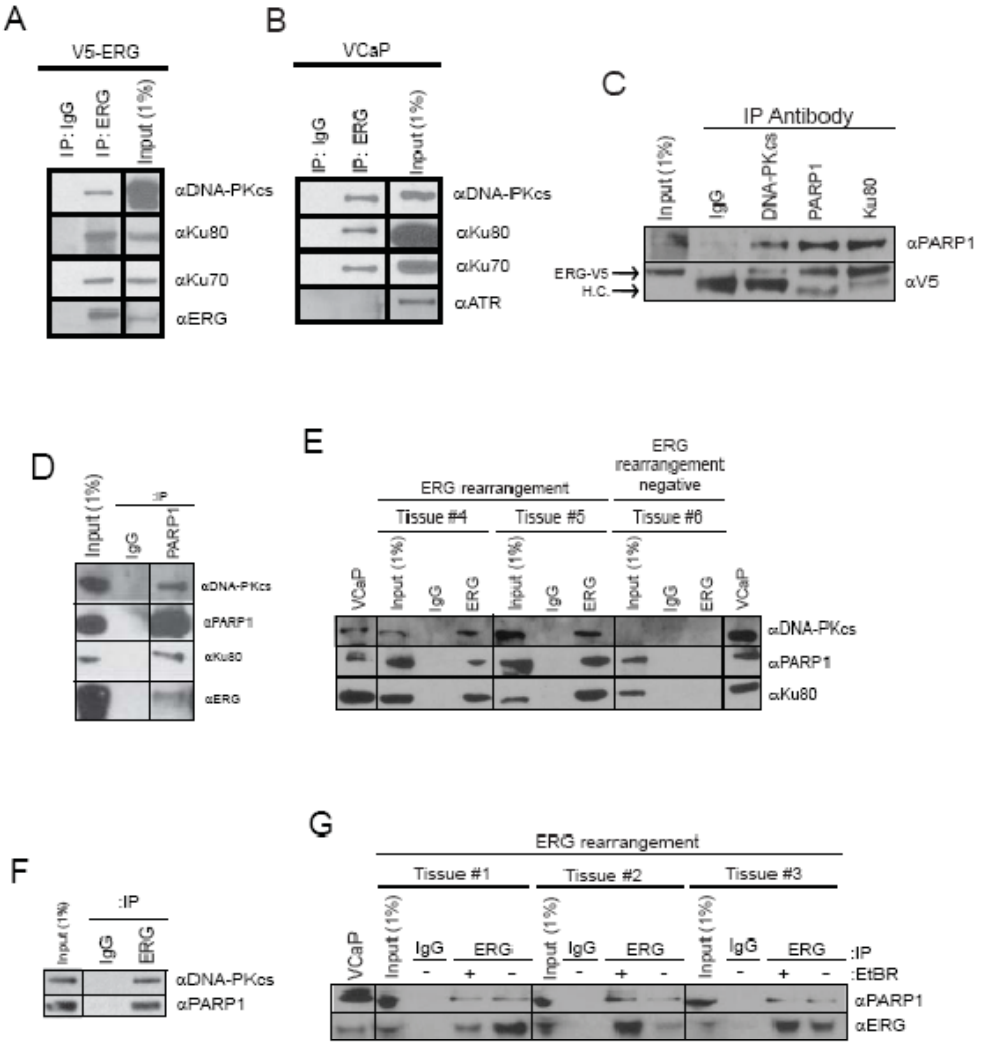



**Figure 2.3: ERG interacts with DNA-PKcs, PARP1, Ku70 and Ku80.**
**(A)** Co-immunoprecipitation of V5-ERG, DNA-PKcs, Ku70 and Ku80 confirmed by immunoblot. VCaP cells were infected with adenovirus encoding V5-ERG.
**(B)** Untreated VCaP cells were used to show that the endogenous gene fusion product interacts with DNA-PKcs, Ku70 and Ku80, but not ATR. Immunoprecipitations were run in triplicate. H.C. stands for heavy chain.
**(C)** IP-Western blot analysis of RWPE cells (low endogenous ERG) transduced with ERG-V5 adenovirus 48hours prior to harvesting total cell lysate. Immunoprecipitations were performed with different antibodies as indicated.
**(D)** IP-Western performed on VCaP cells using agarose coupled-PARP1 or IgG antibodies.
**(E)** Human prostate cancer tissue samples were collected from the UM warm autopsy program. Two human prostate cancer tissue samples with ERG

rearrangement and one without ERG translocation were used for additional IP-Western blot analysis of the ERG:DNA-PKcs and ERG:PARP1 interactions.
**(F)** HEK293 cells were transfected with full length wild type ERG and used for IP-Western blot analysis.
**(G)** Three tissues with ERG rearrangement were used to perform IP-Western blot analysis using an agarose coupled ERG antibody in the presence or absence of ethidium bromide as indicated. Membranes were blot for PARP1 and ERG expression.

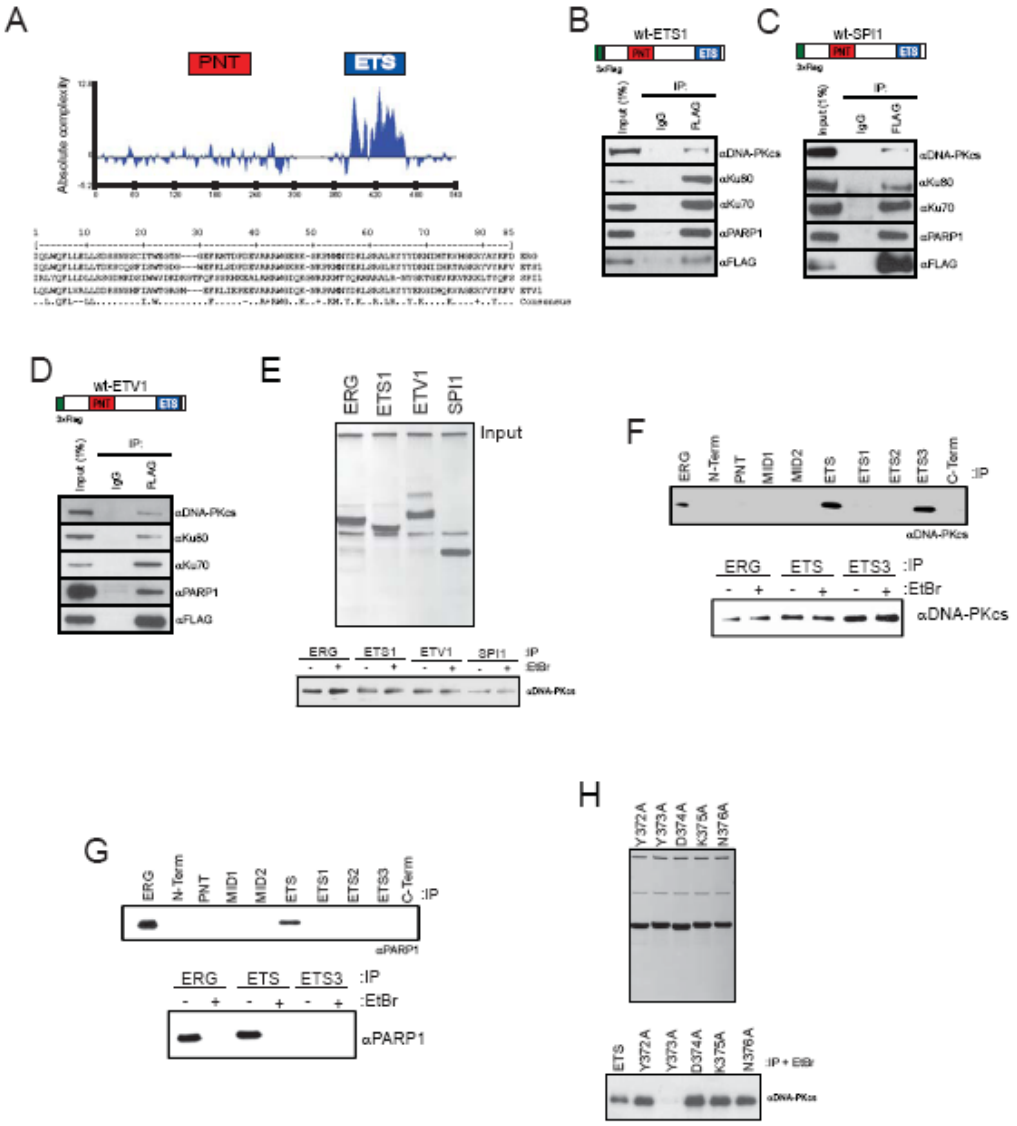



**Figure 2.4: Several ETS proteins interact with DNA-PKcs directly, ERG requires Y373.**

**(A)** Absolute complexity of the three additional ETS gene paralogues selected for immunoprecipitation experiments. Outside of the conserved pointed and ETS DNA binding domains, this plot shows low amino acid sequence homology with ERG.

**(B)**, **(C)** and **(D)**, IP-Western blot analysis of HEK293 cells transfected with an FLAG-ETS1, FLAG-SPI1 or FLAG-ETV1 expression vector, respectively.

**(E)** IP-Western performed using purified halo-tagged ERG, ETS1, ETV1 and SPI1 as well as purified DNA-PKcs.

**(F)** and **(G)** IP-Western blots performed using HALO ligand linked bead to map DNA-PKcs:ERG and PARP1:ERG interactions, respectively. HALO-tagged ERG

fragments were overexpressed and purified as previously described (Yu et al., 2010). IPs with tiling fragments (top blot) performed without ethidium bromide. **(H)** IP-Western using purified HALO-ETS constructs with different alanine mutations as indicated and purified DNA-PKcs. All IP-Westerns were performed in triplicate.

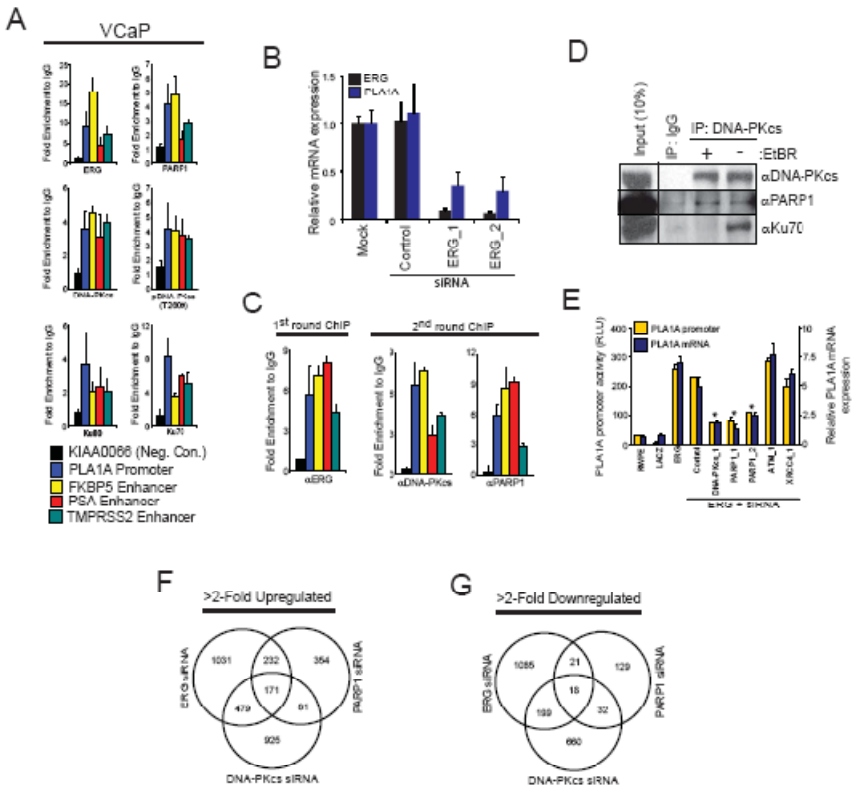



**Figure 2.5: DNA-PKcs and PARP1 associate with ERG regulated genomic loci in VCaP cells and are required for ERG-mediated transcription.**
**(A)** Chromatin immunoprecipitation assays performed on untreated VCaP cells. **(B)** VCaP cells were treated with either scrambled control or ERG siRNA and mRNA expression changes were analyzed by qRT-PCR. Experiment was run in triplicate. Standard error of the mean is shown in all plots.
**(C)** Co-recruitment as assessed by re-ChIP assays. First round ChIP was completed with rabbit ERG antibody compared to a rabbit IgG control while second round chip was completed with either PARP1 or DNA-PKcs mouse antibodies and compared to a mouse IgG control. All reactions were run three times and in quadruplicate.
**(D)** IP-Western blot analysis of a DNA-PKcs pulldown from VCaP cells. Representative images are shown. **(E)** RWPE cells transduced with either LACZ or ERG adenovirus were treated with siRNA for 48 hours prior to analysis for ETS target gene expression (*PLA1A*) by qPCR or promoter activity by luminescence assay. All experiments were completed in triplicate.
**(F)** and **(G)**, Analysis of genes greater than 2-fold up- or down-regulated, respectively, following siRNA treatments as indicated. All overlaps were statistically significant with $p < 0.0001$ except for >2-fold downregulated genes ERG and PARP1 siRNA treatments.

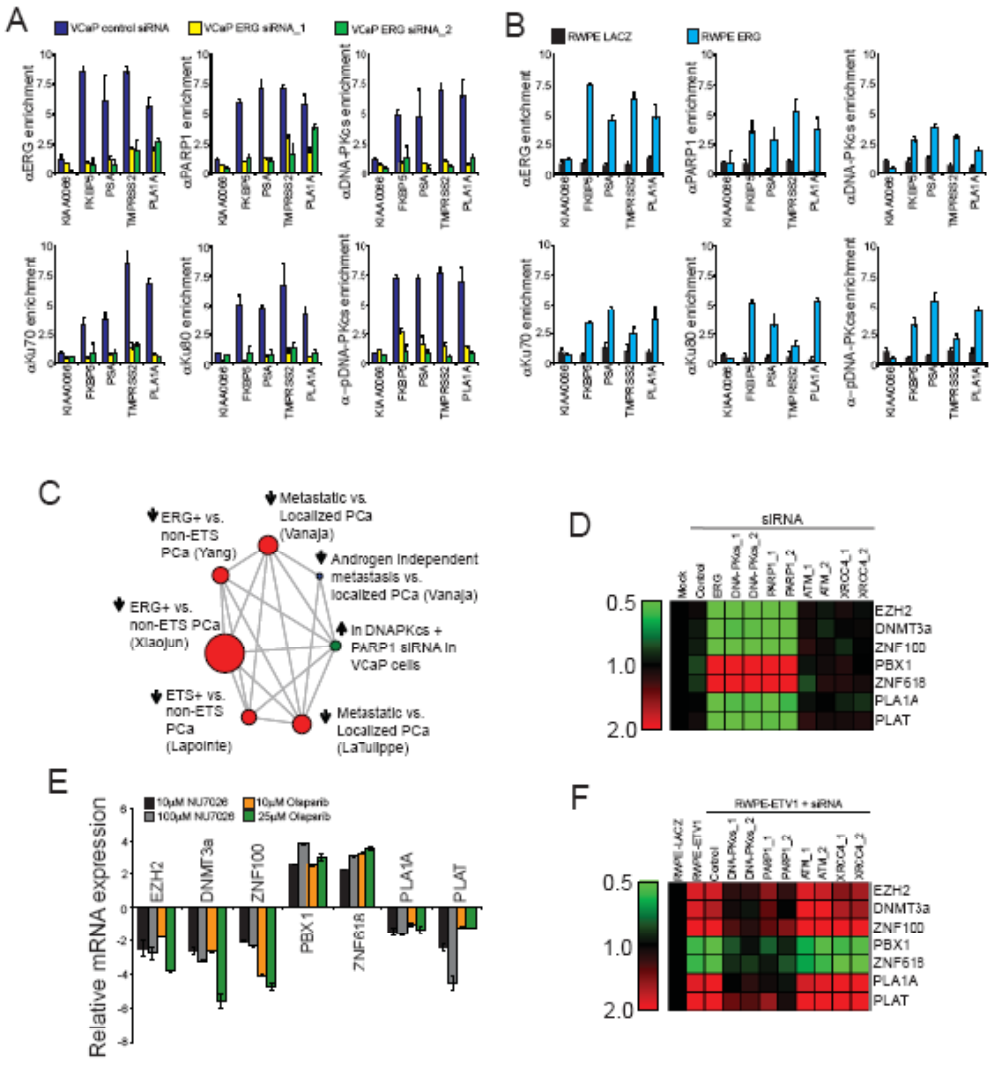



**Figure 2.6: PARP1 and DNA-PKcs are required for ERG-regulated transcription.**

**(A)** ChIP of PARP1 and the DNAPK complex shows an association with ERG-regulated targets including the *PLA1A* promoter as well as *FKBP5*, *PSA* and *TMPRSS2* enhancers, but not the negative control gene *KIAA0066*. ChIPs were performed in VCaP cells treated with control or one of two independent ERG siRNAs for 48 hours prior to cross-linking.

**(B)** ChIP performed as in (A), but with stable RWPE-ERG or -LACZ cells.

**(C)** Data from gene expression arrays was analyzed by molecular concept mapping. The gene set analyzed is the set of genes that were greater than 2-fold differential in all three siRNA treatments relative to control. This gene set was used to determine the correlation of genes regulated by ERG, DNA-PKcs and PARP1 in VCaP cells with published microarray data. Node size is proportional to the number of genes in the set and edges represent statistically

significant associations ($p < 0.01$). Arrow directionality represents gene sets either being induced or repressed.
**(D)** VCaP cells were treated with siRNA as indicated 48 hours prior to RNA isolation. qPCR was then run to confirm gene expression changes identified in the microarray experiment. Data is shown as a heat map with siRNA treatments along the x-axis and genes whose expression was analyzed by qPCR along the y-axis. Shades of green represent down-regulation of gene expression while shades of red represent up-regulation.
**(E)** VCaP cells were treated with either NU7026 or Olaparib for 48 hours as indicated and qPCR analysis of ERG-target genes identified from gene expression microarray experiment was performed.
**(F)** As in (D), except stable RWPE-ETV1 cells were used. All qPCR experiments were run three times in quadruplicate.
All bar graphs are shown with +/- SEM unless otherwise indicated.

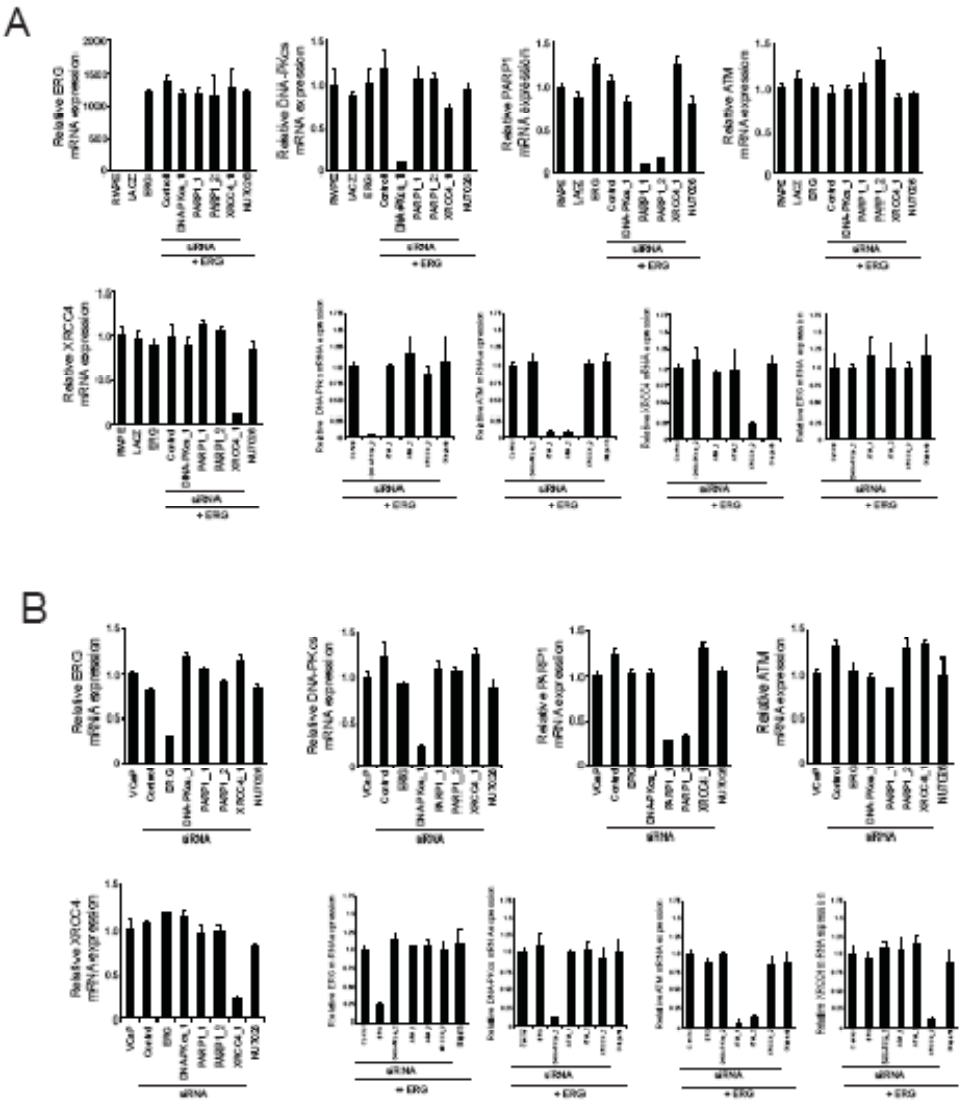



**Figure 2.7: Quantification of knockdown in RWPE and VCaP models.**
**(A)** qPCR analysis of RNA isolated in parallel to the invasion and promoter reporter assays forty eight hours after adenoviral transduction and siRNA treatment. As expected, the small molecule inhibitors had no effect on total levels of PARP1 or DNA-PKcs mRNA expression.
**(B)** as in (A), with untreated VCaP cells. Additional siRNAs were run as additional controls. These were DNA-PKcs siRNA2, ATM siRNA 2 and 3 and Olaparib treatments. All reactions were run three times and in quadruplicate.

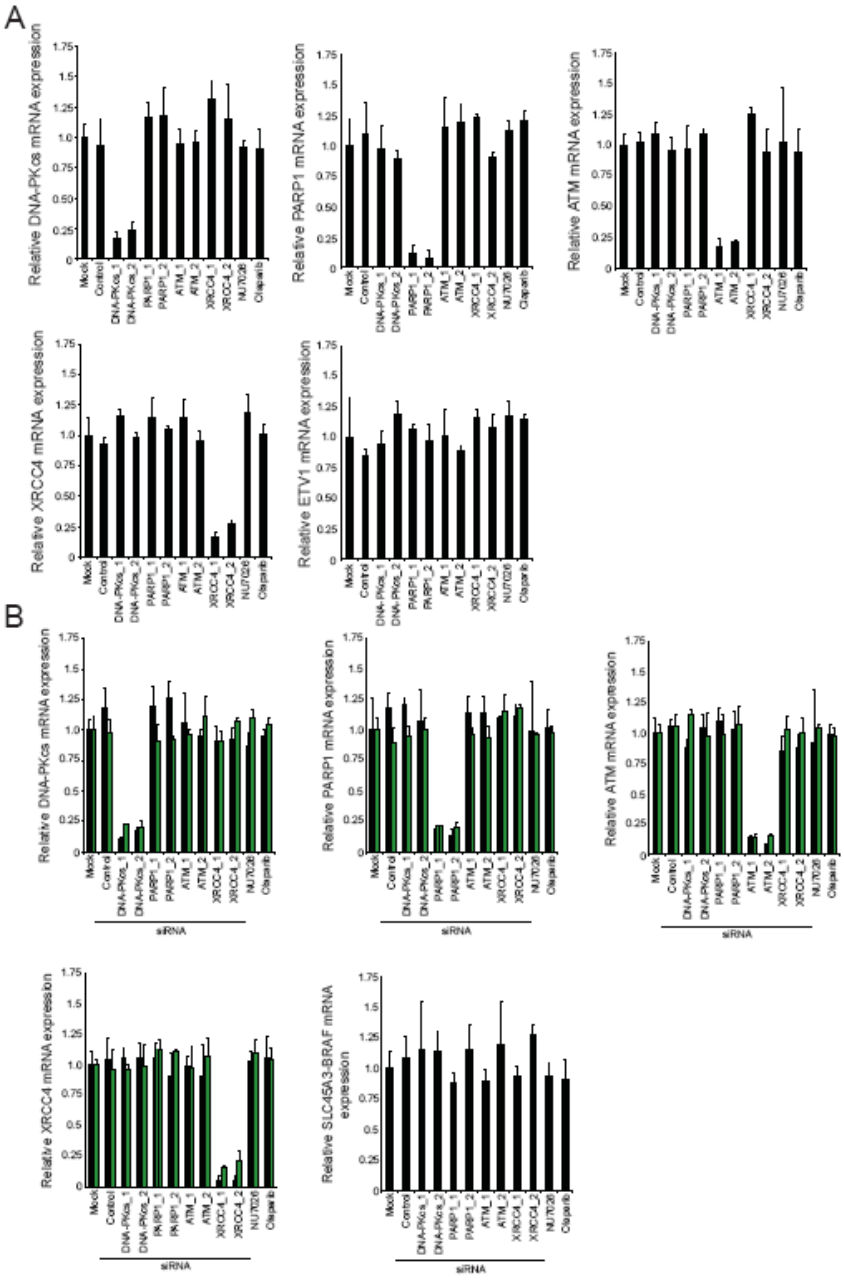



**Figure 2.8: Quantification of knockdown in RWPE-ETV1, PC3 and RWPE-SLC45A3-BRAF models.**
**(A)** qRT-PCR analysis of RWPE-ETV1 cells treated with siRNA, 100μM NU7026 or 25μM Olaparib as indicated.
**(B)** As in (A) except with PC3 and RWPE_SLC45A3-BRAF models. All reactions were run three times and in quadruplicate.

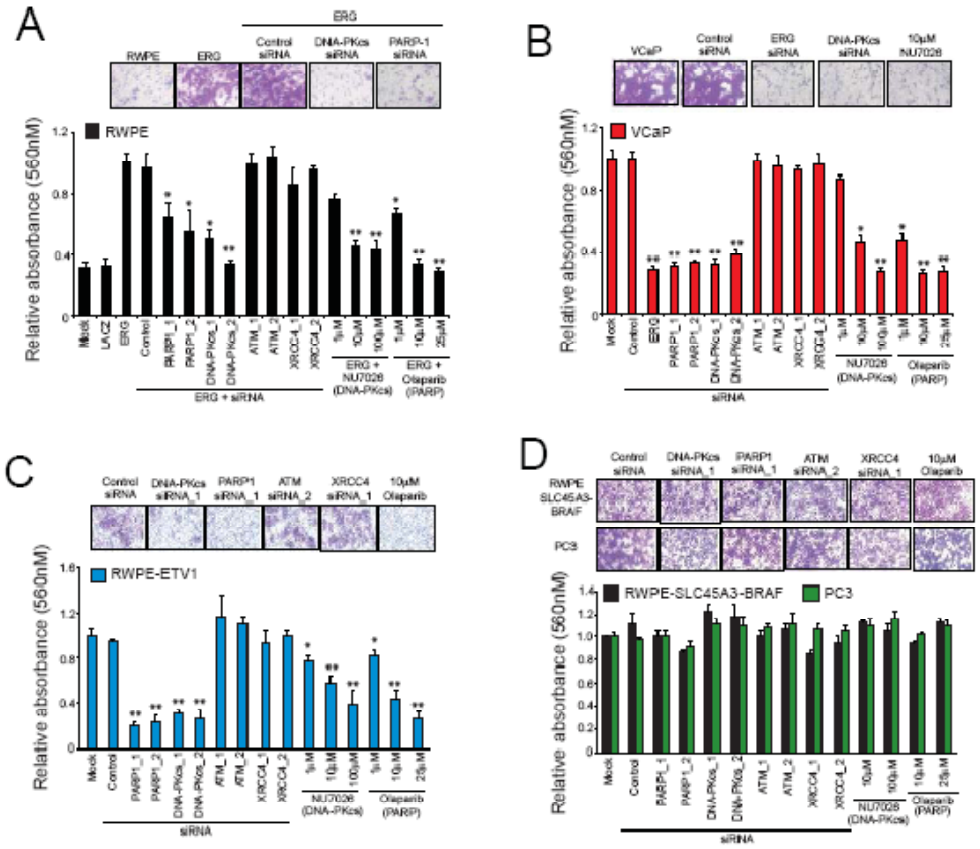



**Figure 2.9: ERG-mediated invasion requires engagement of PARP1 and DNA-PKcs.**

**(A)** RWPE cells were infected with ERG adenovirus and treated with indicated siRNAs or different doses of the DNA-PKcs inhibitor, NU7026, or the small molecule PARP1 inhibitor, Olaparib, for 48 hours prior to plating cells in matrigel-coated Boyden chambers. After another 48 hours, cell invasion quantified.

**(B)** As in (A), except VCaP cells.

**(C)** As in (A), except stable RWPE cells transduced with ETV1 lentivirus.

**(D)** As in (A), except PC3 or RWPE-SLC45A3-BRAF cells.

Representative of three independent experiments.

Representative photomicrographs of invaded cells shown (lower Boyden chamber stained with crystal violet). For all experiments mean +/- SEM shown, * $p < 0.05$, ** $p < 0.01$.

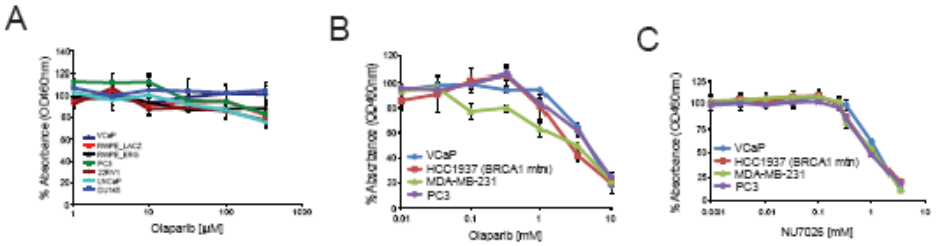



**Figure 2.10: WST chemosensitivity curves for the prostate cancer cell lines used for both *in vitro* and *in vivo* experiments.**
**(A)** Panel of cell lines treated with Olaparib. Doses were centered on the concentrations used for in vitro studies (10-25µM).
**(B)** Olaparib concentrations were extended for several cell lines in order to determine the short term (72 hour) $EC_{50}$ value.
**(C)** NU7026 chemosensitivity assay. Percent absorbance was assessed 72 hours after drug treatment. Experiment was run three times in triplicate with 5 independent samples analyzed per condition.

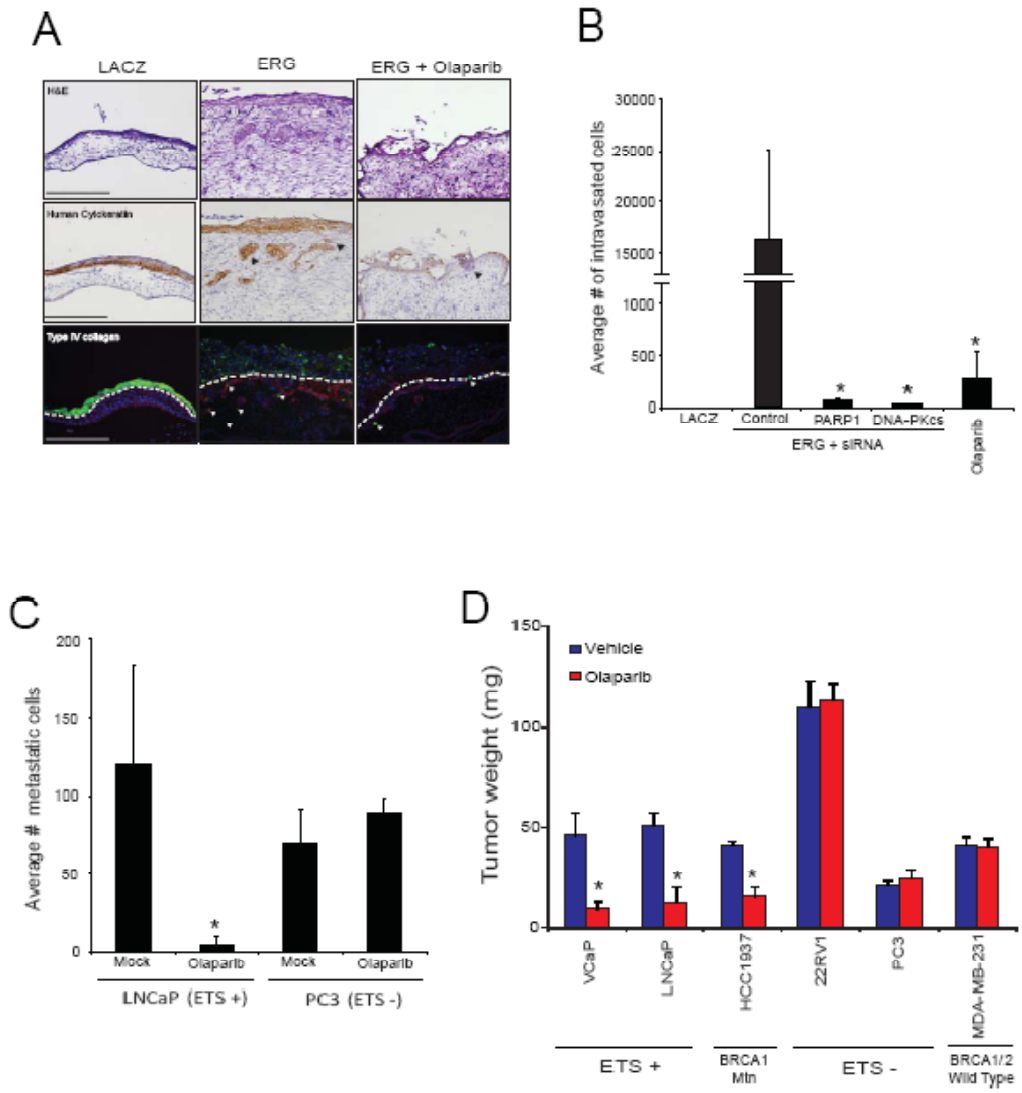



**Figure 2.11: ERG-mediated invasion, intravasation and metastasis requires PARP1 activity *in vivo*.**
**(A)** CAM invasion assay performed using stable RWPE-LACZ, or RWPE-ERG cells labeled with microspheres (green fluorescence emission) and treated with or without a single dose of Olaparib [40mg/kg] as indicated. Seventy two hours after implantation, the upper CAM was harvested. Frozen sections were created and stained for hematoxylin and eosin (top row), human-specific cytokeratin (immunohistochemistry, middle row) or chicken-specific Type IV collagen (red immunofluorescence, bottom row). Arrows indicate cells invaded through the upper CAM. Representative images are shown.
**(B)** CAM intravasation assay performed using stable RWPE-ERG cells pre-treated with siRNA as indicated. Alternatively, RWPE-ERG cells were implanted and treated with a single dose of Olaparib immediately after implantation [40mg/kg]. Seventy two hours after implantation, the lower

CAM was harvested. Total DNA was isolated from the lower CAM and qPCR was performed using human-specific *ALU* PCR primers. Total cell number was determined by comparing to a standard curve created using varying amounts of RWPE cells as input.
**(C)** Liver metastasis in chicken embryos was assessed 8 days following implantation of either LNCaP (ETV1 rearrangement) or PC3 (no ETS rearrangement) cells onto the upper CAM. Animals were injected every other day with Olaparib [40mg/kg] prior to harvesting chicken livers. Total cell number was then quantified by qPCR as in (**B**).
**(D)** ETS positive (VCaP and LNCaP), ETS negative (PC3 and 22RV1) prostate cancer cells as well as BRCA1 mutant (HCC1937) and BRCA1/2 WT (MDA-MB-231) breast cancer cells were implanted onto the upper CAM. These cell line xenografts were then treated with 40mg/kg Olaparib every other day for 8 days. Tumors (non-invaded cells remaining on the upper CAM) were collected and weighed. Average tumor weight is shown.
For all experiments mean +/- SEM shown, * $p < 0.05$, ** $p < 0.01$.

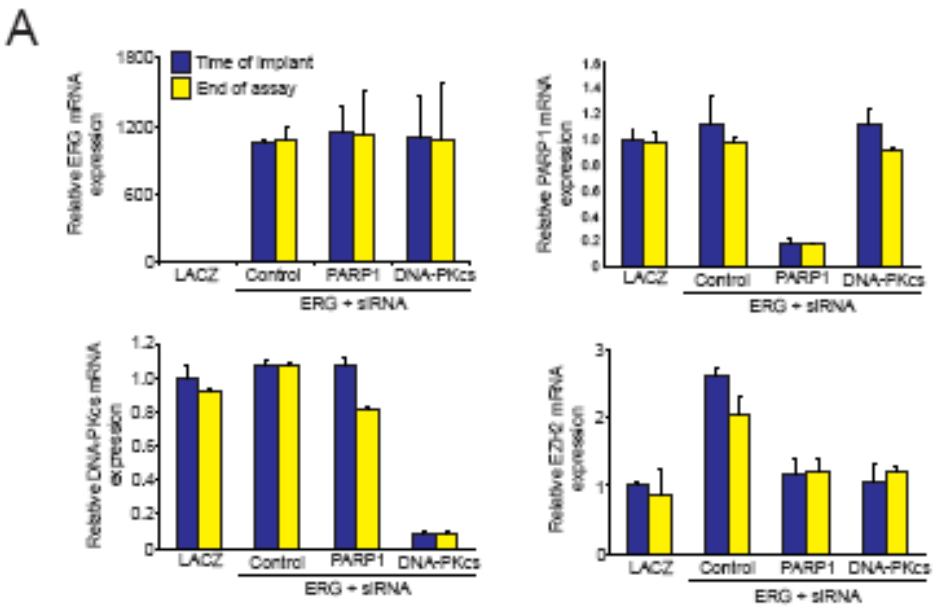



**Figure 2.12: PARP1 and DNA-PKcs expression are required for ERG-mediated transcription of *EZH2 in vivo*.**
**(A-D)** qPCR analysis of gene expression changes following siRNA treatment of cells at the time of implantation into the upper CAM and at the time of harvest from the CAM.
(Bar graphs are shown with +/- SEM unless otherwise indicated).

**Table 2.1: Summarized immunoprecipitation mass spectometry data.**

| Protein ID | Protein Probability | Unique Peptides | % Coverage | Total Peptides | Description |
|---|---|---|---|---|---|
| IPI00005012,IPI00217544,IPI00549267,IPI00786647,IPI00797389,IPI00872796 | 1 | 65 | 64.4 | 573 | ERG 51 kDa protein |
| IPI00296337 | 1 | 37 | 9.9 | 71 | DNAPK Isoform 1 of DNA-dependent protein kinase catalytic subunit |
| IPI00843765,IPI00844215,IPI00871535 | 1 | 36 | 20.4 | 95 | SPTAN1 Isoform 1, 2, or 3 of Spectrin alpha chain, brain |
| IPI00644712 | 1 | 18 | 33.5 | 32 | Ku70 ATP-dependent DNA helicase 2 subunit 1 |
| IPI00220834 | 1 | 17 | 26 | 46 | Ku80 ATP-dependent DNA helicase 2 subunit 2 |
| IPI00479217,IPI00644079,IPI00883857 | 1 | 13 | 19 | 34 | HNRNPU Isoform Short or Long of Heterogeneous nuclear ribonucleoprotein U,isoform a |
| IPI00013296 | 1 | 9 | 42.8 | 29 | RPS18 40S ribosomal protein S18 |
| IPI00440493 | 1 | 9 | 21.9 | 22 | ATP5A1 ATP synthase subunit alpha, mitochondrial precurs |
| IPI00021439,IPI00021440 | 1 | 9 | 21.6 | 27 | ACTB Actin, cytoplasmic 1 or 2 |
| IPI00301263 | 1 | 9 | 7.1 | 15 | CAD protein |
| IPI00011654,IPI00645452 | 1 | 7 | 55.9 | 15 | TUBB Tubulin beta chain,TUBB Tubulin, beta polypeptide |
| IPI00011274,IPI00045498,IPI00845282 | 1 | 7 | 36.5 | 18 | HNRPDL Isoform 1, 2, or 3 of Heterogeneous nuclear ribonucleoprotein D-like |
| IPI00299573 | 1 | 7 | 30.4 | 43 | RPL7A 60S ribosomal protein L7a |
| IPI00007928 | 1 | 7 | 6.1 | 12 | PRPF8 Pre-mRNA-processing-splicing factor 8 |
| IPI00008530 | 1 | 5 | 31.4 | 8 | RPLP0 60S acidic ribosomal protein P0 |
| IPI00219678 | 1 | 5 | 14.6 | 6 | EIF2S1 Eukaryotic translation initiation factor 2 subunit 1 |
| IPI00171903,IPI00383296 | 1 | 5 | 7.5 | 14 | HNRPM Isoform 1 or 2 of Heterogeneous nuclear ribonucleoprotein M |
| IPI00216134,IPI00384369 | 1 | 3 | 13.3 | 10 | TPM1 tropomyosin 1 alpha chain isoform 7,TPM1 Tropomyosi alpha variant 8 |
| IPI00018971 | 1 | 3 | 8.8 | 6 | TRIM21 52 kDa Ro protein |
| IPI00017617 | 0.9999 | 4 | 16.1 | 5 | DDX5 Probable ATP-dependent RNA helicase DDX5 |
| IPI00396378,IPI00414696 | 0.9999 | 3 | 12 | 11 | HNRNPA2B1 Isoform A2 or B1 of Heterogeneous nuclear ribonucleoproteins A2/B1 |
| IPI00016610 | 0.9999 | 2 | 10.1 | 6 | PCBP1 Poly(rC)-binding protein 1 |
| IPI00003865,IPI00007702,IPI00037070,IPI00795040 | 0.9998 | 19 | 33.4 | 35 | HSPA8 Isoform 1 or 2 of Heat shock cognate 71 kDa prote HSPA2 Heat shock-related 70 kDa protein 2 |
| IPI00215637 | 0.9994 | 3 | 9.7 | 6 | DDX3X ATP-dependent RNA helicase DDX3X |
| IPI00012048,IPI00026260,IPI00375531 | 0.9732 | 2 | 14.6 | 2 | NME1; NME2 Nucleoside diphosphate kinase A or B, |

**Table 2.1: Summarized immunoprecipitation mass spectrometry data.** List of highest probability interactors identified from multiple IP-MS experiments as described in the text. Only proteins identified with probability higher than 0.95 are shown.

## REFERENCES


1. I. S. Jeon *et al.*, *Oncogene* **10**, 1229 (Mar 16, 1995).
2. P. H. Sorensen *et al.*, *Nat Genet* **6**, 146 (Feb, 1994).
3. C. Tognon *et al.*, *Cancer Cell* **2**, 367 (Nov, 2002).
4. S. A. Shurtleff *et al.*, *Leukemia* **9**, 1985 (Dec, 1995).
5. S. A. Tomlins *et al.*, *Science* **310**, 644 (Oct 28, 2005).
6. J. Jane-Valbuena *et al.*, *Cancer Res* **70**, 2075 (Mar 1).
7. J. Jane-Valbuena *et al.*, *Cancer Res* **70**, 2075 (Mar 1, 2010).
8. J. E. Darnell, Jr., *Nat Rev Cancer* **2**, 740 (Oct, 2002).
9. J. C. Brenner *et al.*, *Head Neck* **32**, 417 (Sep 16, 2009).
10. N. Palanisamy *et al.*, *Nat Med* **16**, 793 (Jul, 2010).
11. S. A. Tomlins *et al.*, *Neoplasia* **10**, 177 (Feb, 2008).
12. K. Yano *et al.*, *EMBO Rep* **9**, 91 (Jan, 2008).
13. J. Yu *et al.*, *Cancer Cell* **12**, 419 (Nov, 2007).
14. B. E. Helgeson *et al.*, *Cancer Res* **68**, 73 (Jan 1, 2008).
15. S. Varambally *et al.*, *Science* **322**, 1695 (Dec 12, 2008).
16. Q. Cao *et al.*, *Oncogene* **27**, 7274 (Dec 11, 2008).
17. K. J. Livak, T. D. Schmittgen, *Methods* **25**, 402 (Dec, 2001).
18. S. A. Tomlins *et al.*, *Nat Genet* **39**, 41 (Jan, 2007).
19. D. R. Rhodes *et al.*, *Neoplasia* **9**, 443 (May, 2007).
20. A. Zijlstra *et al.*, *Cancer Res* **62**, 7083 (Dec 1, 2002).
21. J. I. Yook *et al.*, *Nat Cell Biol* **8**, 1398 (Dec, 2006).
22. E. H. van der Horst, J. H. Leupold, R. Schubbert, A. Ullrich, H. Allgayer, *Biotechniques* **37**, 940 (Dec, 2004).
23. L. Spagnolo, A. Rivera-Calzada, L. H. Pearl, O. Llorca, *Mol Cell* **22**, 511 (May 19, 2006).
24. K. Park *et al.*, *Neoplasia* **12**, 590 (Jul, 2010).
25. J. Yu *et al.*, *Cancer Cell* **17**, 443 (May 18, 2010).
26. C. W. Garvie, J. Hagman, C. Wolberger, *Mol Cell* **8**, 1267 (Dec, 2001).
27. E. Weterings, D. J. Chen, *J Cell Biol* **179**, 183 (Oct 22, 2007).
28. D. E. Barnes, G. Stamp, I. Rosewell, A. Denzel, T. Lindahl, *Curr Biol* **8**, 1395 (Dec 17-31, 1998).
29. K. M. Frank *et al.*, *Nature* **396**, 173 (Nov 12, 1998).
30. J. Lapointe *et al.*, *Proc Natl Acad Sci U S A* **101**, 811 (Jan 20, 2004).
31. D. K. Vanaja, J. C. Cheville, S. J. Iturria, C. Y. Young, *Cancer Res* **63**, 3877 (Jul 15, 2003).
32. E. LaTulippe *et al.*, *Cancer Res* **62**, 4499 (Aug 1, 2002).
33. S. A. Tomlins *et al.*, *Nature* **448**, 595 (Aug 2, 2007).
34. J. Kim, W. Yu, K. Kovalski, L. Ossowski, *Cell* **94**, 353 (Aug 7, 1998).
35. J. A. Brown, R. B. Marala, *J Pharmacol Toxicol Methods* **47**, 137 (May-Jun, 2002).
36. C. M. Whitacre *et al.*, *Cancer Res* **55**, 3697 (Sep 1, 1995).
37. G. Keith, J. Desgres, G. de Murcia, *Anal Biochem* **191**, 309 (Dec, 1990).
38. P. A. Nguewa, M. A. Fuertes, B. Valladares, C. Alonso, J. M. Perez, *Prog Biophys Mol Biol* **88**, 143 (May, 2005).

39. U. N. Kasid, B. Halligan, L. F. Liu, A. Dritschilo, M. Smulson, *J Biol Chem* **264**, 18687 (Nov 5, 1989).
40. A. M. Ferro, B. M. Olivera, *J Biol Chem* **259**, 547 (Jan 10, 1984).
41. P. C. Hanawalt, B. A. Donahue, K. S. Sweder, *Curr Biol* **4**, 518 (Jun 1, 1994).
42. S. L. Oei, J. Griesenbeck, M. Schweiger, M. Ziegler, *J Biol Chem* **273**, 31644 (Nov 27, 1998).
43. C. M. Simbulan *et al.*, *J Biol Chem* **268**, 93 (Jan 5, 1993).
44. T. Ruscetti *et al.*, *J Biol Chem* **273**, 14461 (Jun 5, 1998).
45. L. Davidovic, M. Vodenicharov, E. B. Affar, G. G. Poirier, *Exp Cell Res* **268**, 7 (Aug 1, 2001).
46. W. L. Kraus, *Curr Opin Cell Biol* **20**, 294 (Jun, 2008).
47. T. Helleday, *Mol Oncol* **5**, 387 (Aug, 2011).
48. P. Chambon, J. D. Weill, P. Mandel, *Biochem Biophys Res Commun* **11**, 39 (Apr 2, 1963).
49. J. M. de Murcia *et al.*, *Proc Natl Acad Sci U S A* **94**, 7303 (Jul 8, 1997).
50. Z. Q. Wang *et al.*, *Genes Dev* **11**, 2347 (Sep 15, 1997).
51. M. F. Langelier, J. L. Planck, S. Roy, J. M. Pascal, *J Biol Chem* **286**, 10690 (Mar 25, 2011).
52. M. N. Paddock *et al.*, *DNA Repair (Amst)* **10**, 338 (Mar 7, 2011).
53. G. de Murcia, J. Menissier de Murcia, *Trends Biochem Sci* **19**, 172 (Apr, 1994).
54. G. de Murcia *et al.*, *Mol Cell Biochem* **138**, 15 (Sep, 1994).
55. L. Virag, C. Szabo, *Pharmacol Rev* **54**, 375 (Sep, 2002).
56. A. Ruf, J. Mennissier de Murcia, G. de Murcia, G. E. Schulz, *Proc Natl Acad Sci U S A* **93**, 7481 (Jul 23, 1996).
57. R. Krishnakumar *et al.*, *Science* **319**, 819 (Feb 8, 2008).
58. H. Ogino *et al.*, *BMC Genomics* **8**, 41 (2007).
59. A. Burkle, *Bioessays* **23**, 795 (Sep, 2001).
60. S. Smith, *Trends Biochem Sci* **26**, 174 (Mar, 2001).
61. M. E. Smulson *et al.*, *Adv Enzyme Regul* **40**, 183 (2000).
62. K. W. Caldecott, *DNA Repair (Amst)* **2**, 955 (Sep 18, 2003).
63. S. F. El-Khamisy, M. Masutani, H. Suzuki, K. W. Caldecott, *Nucleic Acids Res* **31**, 5526 (Oct 1, 2003).
64. C. Keil, T. Grobe, S. L. Oei, *J Biol Chem* **281**, 34394 (Nov 10, 2006).
65. M. Malanga, F. R. Althaus, *J Biol Chem* **279**, 5244 (Feb 13, 2004).
66. S. L. Oei, C. Keil, M. Ziegler, *Biochem Cell Biol* **83**, 263 (Jun, 2005).
67. T. B. Andrasi, A. Blazovics, G. Szabo, C. F. Vahl, S. Hagl, *Am J Physiol Heart Circ Physiol* **288**, H2972 (Jun, 2005).
68. E. R. Plummer, *Curr Opin Pharmacol* **6**, 364 (Aug, 2006).
69. K. Ratnam, J. A. Low, *Clin Cancer Res* **13**, 1383 (Mar 1, 2007).
70. S. L. Allinson, Dianova, II, G. L. Dianov, *Acta Biochim Pol* **50**, 169 (2003).
71. C. E. Strom *et al.*, *Nucleic Acids Res* **39**, 3166 (Apr, 2011).
72. H. Gu, J. D. Marth, P. C. Orban, H. Mossmann, K. Rajewsky, *Science* **265**, 103 (Jul 1, 1994).

73. R. S. Tebbs, L. H. Thompson, J. E. Cleaver, *DNA Repair (Amst)* **2**, 1405 (Dec 9, 2003).
74. W. Y. Mansour, T. Rhein, J. Dahm-Daphi, *Nucleic Acids Res* **38**, 6065 (Oct, 2011).
75. O. Mortusewicz, J. C. Ame, V. Schreiber, H. Leonhardt, *Nucleic Acids Res* **35**, 7665 (2007).
76. H. E. Bryant *et al.*, *Nature* **434**, 913 (Apr 14, 2005).
77. H. Farmer *et al.*, *Nature* **434**, 917 (Apr 14, 2005).
78. P. C. Fong *et al.*, *N Engl J Med* **361**, 123 (Jul 9, 2009).
79. M. Yaneva, T. Kowalewski, M. R. Lieber, *Embo J* **16**, 5098 (Aug 15, 1997).
80. K. O. Hartley *et al.*, *Cell* **82**, 849 (Sep 8, 1995).
81. B. L. Sibanda, D. Y. Chirgadze, T. L. Blundell, *Nature* **463**, 118 (Jan 7, 2010).
82. T. A. Dobbs, J. A. Tainer, S. P. Lees-Miller, *DNA Repair (Amst)* **9**, 1307 (Dec 10, 2011).
83. Y. Zhao *et al.*, *Cancer Res* **66**, 5354 (May 15, 2006).
84. Y. Hosoi *et al.*, *Int J Cancer* **78**, 642 (Nov 23, 1998).
85. A. Kurimasa *et al.*, *Mol Cell Biol* **19**, 3877 (May, 1999).
86. T. M. Gottlieb, S. P. Jackson, *Cell* **72**, 131 (Jan 15, 1993).
87. S. Yoo, W. S. Dynan, *Nucleic Acids Res* **27**, 4679 (Dec 15, 1999).
88. E. J. Bennett, J. W. Harper, *Nat Struct Mol Biol* **15**, 20 (Jan, 2008).
89. C. A. Koch *et al.*, *Embo J* **23**, 3874 (Oct 1, 2004).
90. R. H. Wong *et al.*, *Cell* **136**, 1056 (Mar 20, 2009).
91. B. G. Ju *et al.*, *Science* **312**, 1798 (Jun 23, 2006).

# Chapter III

## Mechanistic Rationale for Poly(ADP)-Ribose Polymerase inhibition in ETS Gene Fusion Positive Prostate Cancer

**ABSTRACT**

**While genomic rearrangements leading to ETS gene overexpression occur in about 50% of prostate cancers, transcription factors like the ETS genes have been notoriously difficult to target therapeutically. Here, I will present evidence indicating that the ETS:PARP1 interaction axis may represent a target for therapeutic intervention in cancers with ETS gene fusions. This chapter suggests that inhibition of co-factors necessary for function may represent a promising paradigm of treatment for malignancies driven by oncogenic transcription factors. This chapter also shows that overexpression of the *TMPRSS2:ERG* fusion induces DNA damage, which is potentiated by PARP1 inhibition in a manner similar to that of BRCA1/2-deficiency. Furthermore, this motivates the assessment of ETS gene fusions as a potential biomarker of response in future clinical trials incorporating**

**PARP inhibitors into the treatment of prostate cancer and other ETS fusion-positive malignancies[1].**

## INTRODUCTION

In the previous chapter, I identified an interaction between ERG and PARP1. Here I will explore the therapeutic implications of PARP inhibition for ETS positive prostate cancer. Poly(ADP) ribose Polymerases (PARPs) are a family of several enzymes with unique structures and functions, which have the ability to create ADP-ribose chains on target proteins using NAD+ as a substrate (*1*). Of PARP enzymes, PARP1 is a functionally complex molecule involved in both DNA repair (reviewed in (*2*)) and transcriptional regulation (reviewed in (*3*)). Despite the mechanistic complexity of the enzyme, however, PARP1 knockout mice are viable with almost no significant phenotypic alterations (*4*, *5*) suggesting that it is an ideal therapeutic target.

PARP1 has generated significant interest as a therapeutic target in *BRCA1/2* mutant cancer cells, which are unable to effectively execute a DNA double strand repair process called homologous recombination (*6*, *7*). Importantly, small molecules that compete for the NAD+ binding pocket of PARP1 block enzymatic activity and potentiate DNA damage in this setting (*6*, *7*). Indeed, the

[1] This chapter was published in *Cancer Cell* by the following authors: J. Chad Brenner, Bushra Ateeq, Yong Li, Anastasia K. Yocum, Qi Cao, Irfan A. Asangani, Sonam Patel, Xiaoju Wang, Hallie Liang, Jindan Yu, Nallasivam Palanisamy, Javed Siddiqui, Wei Yan, Xuhong Cao, Rohit Mehra, Aaron Sabolch, Venkatesha Basrur, Robert J. Lonigro, Jun Yang, Scott A Tomlins, Christopher A. Maher, Kojo S.J. Elenitoba-Johnson, Maha Hussain, Nora M. Navone, Kenneth J. Pienta, Sooryanarayana Varambally, Felix Y. Feng, Arul M. Chinnaiyan

initial clinical trials using a PARP inhibitor called Olaparib demonstrated early efficacy in a phase I clinical trial of *BRCA1/2* deficient cancers, with limited side effects (*8*). At the time of the research presented here, Olaparib was the primary small molecule PARP inhibitor that had been used in the clinic, thus, it was employed for our studies.

Given the observation that PARP1 interacts with ERG in prostate cancer cells and tissues, is required for ETS transcriptional activity in cell lines, we postulated that PARP inhibitors could block ETS transcriptional signaling *in vivo* and specifically disrupt ETS positive xenograft growth. Our preliminary evidence from CAM assays demonstrated that PARP inhibition disrupts the growth of ETS positive, but not ETS negative, prostate tumors to a similar extent as *BRCA1/2* deficient tumors. Thus, we extended our analysis to 9 xenograft models including 3 primary human prostate cancers that were serially propagated in mice.

## MATERIALS AND METHODS

### Cell Lines

PC3 prostate cancer cell lines were grown in RPMI 1640 (Invitrogen, Carlsbad, CA) and VCaP cells in DMEM with Glutamax (Invitrogen) both supplemented with 10% FBS (Invitrogen) in 5% $CO_2$ cell culture incubator. The immortalized prostate cell line RWPE-1 was grown in Keratinocyte media with L-glutamine (Invitrogen) supplemented with 2.5μg EGF (Invitrogen) and 25mg Bovine Pituitary Extract (Invitrogen). All cultures were also maintained with 50 units/ml of Penicillin/streptomycin (Invitrogen). The genetic identity of each cell line was confirmed by genotyping samples. Briefly, DNA samples were diluted to

0.10ng/μl and analyzed in the University of Michigan DNA sequencing Core using the Profiler Plus PCR Amplification Kit (Applied Biosystems, Foster City, CA) according to the manufacturer's protocol. The 9 loci D3S1358, D5S818, D7S820, D8S1179, D13S317, D18S51, D21S11, FGA, vWA and the Amelogenin locus were analyzed and compared to ladder control samples as previously described (*9*). Lentiviruses were generated by the University of Michigan Vector Core. PC3 cells were infected with lentiviruses expressing pLentilox-CMV-ERG or pLentilox-CMV-LACZ control and stable cell lines were selected by sorting at the University of Michigan flow cytometry core. Stable infection was monitored by confirming GFP expression every three days. Likewise, PC3-luciferase cells were created by transduction of pLentilox-CMV-Luciferase available from the University of Michigan Vector Core. Stable RWPE-ERG, RWPE-ETV1 and RWPE-LACZ cells were created and described previously (*10, 11*).

**Immunoprecipitation**

PC3-ERG cells were lysed in Triton X-100 lysis buffer (20mM MOPS, pH 7.0, 2mM EGTA, 5mM EDTA, 30mM sodium fluoride, 60mM β-glycerophosphate, 20mM sodium pyrophosphate, 1mM sodium orthovanadate, 1% Triton X-100, 1mM DTT, protease inhibitor cocktail (Roche, #14309200)). Cell lysates (0.5-1.0mg) were then pre-cleaned with protein A/G agarose beads (Santa Cruz, # sc-2003) by incubation for 1 hour with shaking at 4°C followed by centrifugation at 2000 rpm for 3 minute. Lysates were incubated with ethidium bromide (Sigma) as previously described(*12*). Antibody coupling reactions were

performed according to the Dynabeads Antibody Coupling Kit (Invitrogen, Cat# 143.11D). Briefly, 10mg Dynabeads M-270 were isolated, washed with buffer and mixed with primary antibody as indicated. Reactions were then incubated on a roller at 37°C overnight (16-24 hours), washed with buffer and resuspended to a final concentration of 10mg antibody coupled beads/mL. Lysates were then incubated overnight with the coupled antibodies as indicated. In the case of IPs performed with uncoupled antibody, 2µg antibody was used per reaction and lysates were incubated at 4°C for 4 hours with shaking prior to addition of 20µL protein A/G agarose beads. The mixture was then incubated with shaking at 4°C for another 4 hours or overnight prior to washing the lysate-bead precipitate (centrifugation at 2000 rpm for 3 minutes) 4 times in Triton X-100 lysis buffer. Beads were finally precipitated by centrifugation, resuspended in 25µL of 2x loading buffer and boiled at 80°C for 10 minutes for separation the protein and beads. Samples were then analyzed by SDS-PAGE Western blot analysis as described below.

**Immunoblot Analysis**

The cell lines were plated in two wells of a 6-well plate at 250,000 cells/mL 24hours prior to harvesting by trypsinization. Pellets were then flash frozen, briefly sonicated and homogenized in NP40 lysis buffer (50 mM Tris-HCl, 1% NP40, pH 7.4, Sigma, St. Louis, MO), and complete proteinase inhibitor mixture (Roche, Indianapolis, IN). Ten micrograms of each protein extract were boiled in sample buffer, size fractionated by SDS-PAGE, and transferred onto

Polyvinylidene Difluoride membrane (GE Healthcare, Piscataway, NJ). The membrane was then incubated overnight at 4ºC in blocking buffer [Tris-buffered saline, 0.1% Tween (TBS-T), 5% nonfat dry milk] and incubated for 4 hours at room temperature with the following: anti-DNA-PKcs mouse monoclonal (1:1000 in blocking buffer, BD Biosciences #610805, San Jose, CA), anti-Ku70 mouse monoclonal (1:1000 in blocking buffer, BD Biosciences #611892), anti-Ku80 rabbit polyclonal (1:1000 in blocking buffer, Cell Signaling Cat #2180S, Danvers, MA), anti-ERG1/2/3 rabbit polyclonal (1:1000 in blocking buffer, Santa Cruz Biotech # sc-354, Santa Cruz, CA), anti-PARP-1 polyclonal (1:1000 in blocking buffer, Santa Cruz Biotech Cat #sc-8007, Santa Cruz, CA), anti-LACZ (1:1000 in blocking buffer, GenwayBio Cat# 18-732-292258), and anti-β-Actin mouse monoclonal antibody (1:10000, Cell Signaling, Cat #: 4967). Following a wash with TBS-T, the blot was incubated with horseradish peroxidase-conjugated secondary antibody and the signals visualized by enhanced chemiluminescence system as described by the manufacturer (GE Healthcare).

**Chromatin Immunoprecipitation**

Isogenic PC3 cells were grown in complete medium and ChIP was carried out as previously described (*13*) using antibodies against ERG (Santa Cruz, #sc-354) or rabbit IgG (Santa Cruz, #sc-2027). Briefly, cultured cells were crosslinked with 1% formaledehyde for 10 min and the crosslinking was inactivated by 0.125M glycine for 5 min at room temperature (RT). Cells were then rinsed with 1x PBS twice and frozen in 1x PBS + 10µl/ml PMSF + phosphatase inhibitor

cocktail (Calbiochem) for 30 min. The following steps were performed at 4°C. Cell pellets were resuspended and incubated in cell lysis buffer + 10ul/ml PMSF and protease inhibitor (Roche) for 10 min. Nuclei pellets were spun down at 5,000 x g for 5 min, resuspended in nuclear lysis buffer, and then incubated for another 10min. Chromatin was sonicated to an average length of 600bp and then centrifuged at 14,000 x g for 10min to remove the debris. Supernatants containing chromatin fragments were incubated with agaros/protein A or G beads (Upstate) for 15 min and centrifuged at 5,000xg for 5min to remove the nonspecific binding. To immunoprecipitate protein/chromatin complexes, the supernatants were incubated with 3-5µg of antibody or IgG overnight, then added 50ul of agarose/protein A or G beads and incubated for another 1 hour. Beads were washed twice with 1X dialysis buffer and four times with IP buffer. The antibody/protein/DNA complexes were eluted with 150µl IP elution buffer twice. To reverse the crosslinks, the complexes were incubated in elution buffer + 10µg RNase A and 0.3 M NaCl at 67°C for 4 hours. DNA/proteins were precipitated with ethanol, air-dried, and dissolved in 100µl of TE. Proteins were then digested by proteinase K at 45°C for 1 hour and DNA was purified with QIAGEN PCR column and eluted with 30µl EB. The final ChIP yield was 10-30ng for each antibody. QPCR is described below and primers were either previously described (*11*) or are in **APPENDIX 1** below.

**small RNA interference**

Knockdowns of specific genes were accomplished by RNA interference using commercially available siRNA duplexes for DNA-PKcs, ATM, ATR, PARP-1, XRCC4 and DNA Ligase4 (Dharmacon, Lafayette, CO) or as previously described for ERG (*14*). At least 4 independent siRNAs were screened for knockdown efficiency against each target and the best siRNA was selected, in some cases only one siRNA was identified. Transfections were performed with OptiMEM (Invitrogen) and oligofectamine (Invitrogen) as previously described (*15*).

**Quantitative Real-Time PCR Assays**

Total RNA was isolated from VCAP, RWPE or PrEC cells that were transfected with siRNA as indicated (Qiagen). Quantitative PCR (QPCR) was performed using SYBR Green dye on an Applied Biosystems 7300 Real Time PCR system (Applied Biosystems, Foster City, CA) as described (*16*). Briefly, 1μg of total RNA was reverse transcribed into cDNA using SuperScript III (Invitrogen, Carlsbad, CA) in the presence of random primers (Invitrogen). All reactions were performed in triplicate with SYBR Green Master Mix (Applied Biosystems) plus 25ng of both the forward and reverse primer according to the manufacturer's recommended thermocycling conditions, and then subjected to melt curve analysis. Threshold levels for each experiment were set during the exponential phase of the QPCR reaction using Sequence Detection Software version 1.2.2 (Applied Biosystems). The relative quantity of the target gene was

completed for each sample using the ΔΔCt method by the comparing mean Ct of the gene to the average Ct of the housekeeping gene, β-Actin (*17*). All oligonucleotide primers were synthesized by Integrated DNA Technologies (Coralville, IA). The primer sequences for the transcript analyzed were either previously described (*11*) or are provided in **supplementary table 3**. Targeted DNA damage pathway arrays were obtained from SA biosciences and run according to the manufacturer's protocol (Qiagen).

**VCaP, PC3, DU145 and 22Rv1 Xenograft Models**

Four weeks old male Balb C nu/nu mice were purchased from Charles River, Inc. (Charles River Laboratory, Wilmington, MA). VCaP ($2 \times 10^6$ cells), DU145 ($5 \times 10^6$ cells), 22Rv1 ($0.5 \times 10^5$ cells), PC3-luciferase-control ($1 \times 10^5$ cells), PC3-luciferase-LACZ ($1 \times 10^5$ cells) or PC3-luciferase-ERG ($1 \times 10^5$ cells) stable cells were resuspended in 100µl of saline with 50% Matrigel (BD Biosciences, Becton Drive, NJ) and were implanted subcutaneously into the left flank region of the mice. Mice were anesthetized using a cocktail of xylazine (80-120 mg/kg, IP) and ketamine (10mg/kg, IP) for chemical restraint before tumor implantation. All tumors were staged for three week before starting the drug treatment. After week 3 for VCaP cells or after 48 hours for PC3 cells, mice (10 per treatment group) were treated with Olaparib (40mg/kg or 100mg/kg as indicated, IP) daily five times a week or Temozolomide (50mg/kg in two cycles of 5 days each as indicated, IP). Olaparib was obtained from Axon Medchem (Groningen, The Netherlands). We chose 100mg/kg/day because, in mice, this

dose corresponds to the maximum clinically achievable dose for which minimal side effects were observed in humans during the initial Phase I trial of Olaparib (*8, 18, 19*). Growth in tumor volume was recorded weekly by using digital calipers and tumor volumes were calculated using the formula $(\pi/6)$ $(L \times W^2)$, where L = length of tumor and W = width. Loss of body weight during the course of the study was also monitored weekly. At the end of the xenograft studies, mice were anaesthetized and blood was collected by cardiac puncture. Supernatant containing serum was separated by incubating blood on ice for 45 min, followed by centrifugation at 8000 rpm for 10 min at $4^oC$. Alanine aminotransferase (ALT) and Alanine aminotransferase (ALT) were then measured on an IDEXX Vettest 8008 biochemical analyzer according to the manufacturer's protocol (IDEXX, France). All procedures involving mice were approved by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan and conform to their relevant regulatory standards.

**Preclinical study using Human Prostate Cancer Xenografts**

Human prostate cancer xenografts MDA-PCa 2b-T (*ETV1* positive), MDA PCa 118b (*ETS*-negative) and MDA-PCa-133 (*ERG* positive) were obtained from Dr. Nora M. Navone's lab. MDA PCa 2b-T is MDA PCa 2b cell line (PMID: 9815652) maintained by serial passages in mice. MDA PCa 118b was previously reported in *Li et al* (*20*). MDA PCa 133 xenograft was recently developed in Dr Navone's Lab. Human tissue specimen used to developed MDA-PCa-133 xenograft was residual from surgery of a bone metastasis of a prostate cancer that

was undergoing castrate resistant progression. Small pieces of the human prostate cancer were implanted into subcutaneous pockets of 6- to 8-week-old male CB17 SCID mice (Charles River Laboratories). Tumor developed within 6 months of implantation and was maintained by passage through mouse as the cells did not sustain *in vitro* growth. Methods of passage and tumor processing were as reported elsewhere (*20*). Written informed consent had been obtained from the patient before sample acquisition and sample was processed according to a protocol approved by The University of Texas, MD Anderson Cancer Center institutional review board.

Each xenografted tumor tissue was aseptically harvested from the host mice. New recipient mice were anesthetized by injecting xylazine (80 mg/kg) and ketamine (10mg/kg) intraperitoneally for chemical restraint before tumor implantation. About 0.5 $cm^3$ size of the excised tumor tissue was directly implanted into subcutaneous pockets of 6- to 8-week-old male CB17 SCID mice (Charles River Laboratories) in a laminar flow sterile hood. Sub-cutaneous pocket was carefully closed by applying 1 or 2 wound clips. All mice were monitored for 24 hours after implantation procedure for any apparent physiologic disturbances or missing wound clips. Olaparib (40mg/Kg body weight) treatment was started after 72 hours. Tumor growth was recorded twice a week using digital calipers and tumor volumes were calculated using the formula (*/6) ($L \times W^2$), where L = length of tumor and W = width. Body weight was also monitored weekly during the course of the study. All procedures involving mice were approved by the

University Committee on Use and Care of Animals (UCUCA) at the University of Michigan and conform to their relevant regulatory standards.

**Immunofluorescence**

Cells were seeded at 25,000 cells/mL in 4-well chamber slide 24 hours prior to the addition of siRNA or drug as indicated. Merbarone (Sigma) was dissolved in DMSO and stored at 4°C. Cells were then washed once in PBS, fixed for 15 minutes in 100% methanol, washed in PBS, permeabilized for 5 minutes in PBS containing 0.2% Triton-X, washed twice in PBS, blocked for 30 minutes in PBS containing 0.5% donkey serum (Sigma, St. Louis, MO), held for 45 minutes in PBS containing 0.5% donkey serum and primary antibody: anti-γ-H2A.X mouse monoclonal (Millipore Cat #05-636), anti-γ-H2A.X rabbit polyclonal (Cell Signaling, Cat # 9719, Danvers, MA ), and anti-53BP1 rabbit polyclonal (Cell Signaling, Cat # 4937, Danvers, MA) washed 3 times with PBS, held for 30 minutes in PBS containing 0.5% donkey serum and secondary antibody: Alexa Fluor 488-anti-Mouse (Invitrogen) or Alexa Fluor 594-anti-Rabbit (Invitrogen). Cells were then stained with DAPI for 5 minutes and slides were mounted 12 hours prior to analysis using Vectashield (Vector laboratories, Burlingame, CA). Images were taken using 100x oil lens on an Olympus Confocal microscope at the University of Michigan microscopy imaging lab.

**Homologous Recombination Assays**

Stable PC3 cell lines were transfected with HR-reporters (Norgen Biotek Corp., Thorold, ON) using Fugene 6.0HD (Roche) according to the manufacturer's protocol. Briefly, total DNA was isolated from cells 24 hours after transfections using the DNeasy blood and tissue DNA extraction kit according to manufacturer protocol (Qiagen). qPCR was then performed to assess HR-efficiency using primer sets to specifically detect recombination (Norgen Biotek Corp.).

**COMET Assay**

VCaP, PC3, PrEC or RWPE cells were seeded at 250,000 cells/mL in a 6-well plate 24 hours prior to treatment with siRNA, drug or vehicle control. After 48 hours, cells were trypsinized, harvest by centrifugation and re-suspend in PBS. Cell counts were then normalized to 1 x 105 cells/mL. Suspended cells (25µL) were then mixed with 250µL 1.0% ultrapure low melting point agarose (Invitrogen) made in 1x Tris-Borate buffer. The agarose-cell mixture was then dropped onto slides allowed to solidify at 4°C in the dark for 20 minutes before immersion in COMET assay lysis solution (Trevigen, Gaithersburg, MD) at 4° in the dark for 45 minutes. Excess buffer was then removed and slides were submerged in freshly prepared neutral solution (Tris Base 60.57g, Sodium Acetate 204.12g, dissolve in 450 ml of dH2O. Adjust to pH = 9.0 with glacial acetic acid.) at room temperature in the dark for 40 minutes. When performed in neutral electrophoresis buffer, the COMET assay measures relative levels of DNA double

strand break fragmentation (*21*). Slides were then washed twice by immersing in 1 X TBE buffer prior to neutral electrophoresis at 20 volts for 30 minutes. Slides were then fixed in 70% ethanol for 5 minutes. Following air drying of the agarose, slides were stained with SYBR Green dye (Invitrogen) and images were collected with a 10x and 40x objective lens. COMET tail moments were then assessed using COMETscore.v1.5 (AutoCOMET.com, Sumerduck, VA) image processing software as described by the manufacturer with greater than 100 cells analyzed in triplicate experiments. Data is reported as tail moment which assesses the fluorescence intensity in the tail relative to the head while accounting for the relative area of both dipoles.Statistics were completed using previously published methods for the analysis of COMET assays (*22, 23*) with the following extension: a two-way ANOVA analysis was run using the transformation y -> log2(1+y) on the tail moments, this enabled comparison of the increase in DNA damage caused by Olaparib in control cells to the increase in DNA damage Olaparib caused in ERG overexpression cells.

## RESULTS

### ETS gene fusion prostate tumors are preferentially susceptible to PARP1 inhibition *in vivo*

Based upon our *in vivo* data from the chicken CAM assay in Chapter 1, we hypothesized that inhibition of PARP1 would inhibit ETS-positive prostate cancer growth in mouse xenograft models. Several PARP inhibitors have entered Phase I

and Phase II clinical trials (*24-26*). One of these, Olaparib, was shown to be well tolerated in cancer patients with a minimal side effect profile (*8*). Thus, to first test our hypothesis, we implanted VCaP (ERG positive) or PC3-LACZ (ETS negative) cells and studied the impact of Olaparib (100mg/kg/day, IP) on xenograft growth. Importantly, we observed a significant reduction of tumor growth in the ETS-rearranged cell line relative to that of the vehicle control, but no change was observed in the ETS-negative control cell line ($p = 0.002$) suggesting preferential sensitivity of ETS-positive tumors (**Figures 3.1A and 3.1B**).

We then extended our experiment to analyze the effects of Olaparib on a panel of ETS-positive and ETS-negative prostate cancer cell lines including an isogenic model. Because this experiment intended to test the specificity of Olaparib-induced growth inhibtion for ERG-overexpressing prostate xenografts, we chose to use a dose similar to that used in previously published xenograft experiments (*27*). Consistent with our hypothesis, this dose of Olaparib had a significant effect on VCaP cells ($p = 0.001$), but did not inhibit the growth of two additional ETS negative cell line xenografts (22RV1 or DU145) (**Figure 3.2A, 3.2B and 3.2C**). This experiment also demonstrated that the VCaP tumor growth response was dose-dependent. To then create an isogenic model, we overexpressed the primary *TMPRSS2:ERG* gene fusion product in the PC3 prostate cancer cell line (PC3-ERG). Western blotting confirmed protein overexpression and IP-Western confirmed the ERG:DNA-Pkcs and ERG:PARP1 interactions (**Figures 3.1C and 3.1D**). ChIP assays demonstrated that ERG binds

to known target genes in PC3 cells (**Figure 3.1E**) and qPCR demonstrated ERG transcriptional activity (**Figure 3.1F**). Surprisingly, ERG overexpression led to a slightly reduced growth rate of PC3 cells relative to LACZ-overexpressing PC3 cells (**Figure 3.1G**). Consistent with the model that ETS-positive, but not ETS-negative prostate tumors are susceptible to PARP inhibition, overexpression of ERG was sufficient to significantly sensitize PC3 cells to PARP inhibition ($p = 0.007$) suggesting that ERG overexpression provides a selective mechanism for Olaparib-mediated growth inhibition (**Figures 3.2D, 3.2E and 3.2F**). Western blot and qPCR analysis of flash frozen, staged PC3-ERG tumors treated with or without drug for 4 hours confirmed inhibition of PARP1 activity and loss of ERG-target gene expression after treatment with Olaparib (**Figures 3.1H-K**). In line with the clinical observation that Olaparib is well tolerated at doses capable of inhibiting PARP activity (*8*), Olaparib treatments in our xenograft models did not lead to a significant decrease in total body weight and did not lead to liver toxicity as assessed by serum levels of alanine aminotransferase (ALT) and aspartate aminotransferase (AST) (**Figure 3.1L**).

Given the specific effect of Olaparib on ERG-positive cell line xenograft growth (**Figure 3.1M**), we extended our analysis with the use of a model of primary human prostate tumors maintained in serial xenografts (*20*). We identified one ERG-positive (MDA-PCa-133), one ETV1-positive (MDA-PCa-2b-T) (FISH-confirmed (*28*)) and one ETS-negative (MDA-PCa-118b) model for this experiment by assessing relative levels of ETS gene expression by qPCR (**Figure 3.13N**). Functional ETS gene status was assessed by testing the relative

expression of several ETS target genes including *EZH2*, *DNMT3a*, *ZNF100*, *PBX1*, *ZNF618, PLA1A* and *PLAT*, between the models (**Figure 3.3A**). As shown in **Figures 3.3B and 3.3C**, low dose Olaparib altered the growth of both the ERG and ETV1 overexpressing primary human prostate cancer xenografts [MDA-PCa-133: $p = 0.05$ (day 8) and $p < 0.01$ (day 12, 16, 20, 24, 28 and 32); MDA-PCa-2b-T: $p < 0.01$ (day 8, 12 and 16)], but had no effect on the ETS-negative primary human xenograft model (**Figure 3.3D**). In all cases, Olaparib did not have an observable effect on total body weight (**Figure 3.3E-3.3M**).

Because we were able to validate the hypothesis that Olaparib specifically alters the growth of either ETS overexpressing or BRCA1/2 deficient cell lines, we sought to extend the treatment regimen to identify combination therapies that enhance the magnitude of inhibition without causing significant toxicity. Recently, an alkylating agent called temozolomide (TMZ) has been shown to potentiate the effects of other PARP inhibitors in several cancer xenograft models (*29-31*) as well as caused a complete or partial response in some patients enrolled in a Phase II trial for metastatic breast cancer (*32*). As expected, the combination treatment resulted in a significant growth reduction of VCaP tumors that was maintained over the entire 6 weeks ($p < 0.001$ for all combination treatments) (**Figure 3.2G**). Even with the combination therapy, at this dose range, no overt toxicity such as excessive weight loss was observed (**Figure 3.3N**). This suggested that the addition of PARP inhibitor therapy to existing chemotherapeutic regimens will help enhance the overall effect for ETS positive tumors.

**ETS transcription factors drive DNA double strand break formation**

To explore potential mechanisms of ETS-specific therapeutic response to these inhibitors of DNA repair, we assessed total levels of DNA double strand breaks *in vitro*. We hypothesized that constitutive overexpression of ERG may lead to an increased susceptibility to DNA damage. Thus, we first assessed the total levels of a histone mark of DNA double strand breaks called γ-H2A.X in Olaparib-treated versus untreated VCaP cells. Surprisingly, the untreated cells had a high level of γ-H2A.X foci (**Figure 3.4A**) leading us to test the hypothesis that overexpression of ETS genes induces DNA double strand breaks. Overexpression of ETS genes in primary prostate epithelial (PrEC) cells induced >5 γ-H2A.X foci in greater than 75% of the analyzed cells while the control genes *LACZ* and *EZH2* had no effect (**Figure 3.4A-B**). Quantitative PCR confirmed overexpression (**Figure 3.5A-C**). Likewise, other ETS genes were also capable of inducing γ-H2A.X foci in several different prostate cell lines (**Figure 3.5D-E**). To then confirm that ERG induces γ-H2A.X in an endogenous setting, we depleted ERG from VCaP cells by RNA interference (**Figure 3.5I**) and found significantly decreased γ-H2A.X foci, $p = 7.16 \times 10^{-3}$ and $p = 1.36 \times 10^{-3}$ for two independent siRNAs, respectively (**Figure 3.4A-B**). While γ-H2A.X foci represent an early mark of DNA damage recognition, 53BP1 is present only in the later stages of repair (*33*). As such, in the presence of a DNA damage response, we expected to observe

an increase in 53BP1 foci formation. Indeed, the ETS genes also induced 53BP1 foci formation (**Figure 3.4A-B and Figure 3.5D**).

After demonstrating that ETS gene overexpression drives the accumulation of markers of DNA double strand breaks, we sought to confirm the presence of DNA double strand breaks by directly analyzing cellular DNA for fragmentation using the COMET assay. As with the γ-H2A.X and 53BP1 foci formation assays, in PC3 cells, ERG or ETV1 overexpression was sufficient to induce significantly longer and brighter tails than those observed in controls ($p < 0.01$ for both ETS genes) and treatment with either ERG siRNA led to a reduction in relative level of DNA double strand breaks ($p < 0.01$) (**Figure 3.4C-D**).

**Olaparib potentiates ETS-induced DNA damage**

After finding that aberrantly expressed ETS transcription factors drive the accumulation of DNA double strand breaks, we hypothesized that by having a baseline level of DNA damage, ETS positive cancers may be specifically susceptible to accumulating DNA damage following inhibition of the interacting DNA repair enzyme PARP1. To test this hypothesis, we analyzed VCaP cells treated with Olaparib for 48 hours. Olaparib-treated VCaP cells had a very high level of γ-H2A.X foci (**Figure 3.5F**). Importantly, by depleting endogenous ERG using either of two independent siRNAs (confirmed in **Figure 3.5I**), we were able to reverse the gross increase in γ-

H2A.X. Similar increases in foci were observed in PC3-ERG cells or PC3 cells with BRCA2 knockdown, but not in the control cells (**Figure 3.5F**). Knockdown efficiency was confirmed by qPCR (**Figure 3.5G**). Quantification of the relative levels of DNA double strand breaks demonstrated that, while there was an increase in the tail moment of all Olaparib-treated cell lines, Olaparib caused a significantly greater increase in the tail moments of ERG-positive cells than controls ($p < 0.001$, two-way ANOVA) (**Figure 3.4C-D**). In conjunction with this observation, both ERG siRNAs led to a significant reduction in DNA damage following Olaparib and blocked the synergistic increase of DNA damage observed with Olaparib ($p = 0.001$ for ERG siRNA_1 and $p < 0.001$ for ERG siRNA_2, two-way ANOVA). Consistent with this observation, the BRCA1 mutant HCC1937 breast cells and PC3 cells with BRCA2 shRNA underwent a significant increase in the total levels of DNA damage when treated with Olaparib. Taken together, this data demonstrates that similar to BRCA1/2-deficent tumors, ETS-positive, but not ETS-negative, prostate cancer models are susceptible to PARP inhibition through the increased incidence of DNA double strand breaks (**Figure 3.4E**).

To discriminate between mechanisms of ERG-potentiated DNA damage, we performed the COMET assays after 0.5, 1, 24 and 48 hours exposure to Olaparib. Surprisingly, the potentiated DNA damage was observed in PC3-ERG cells relative to PC3-LACZ cells as early as 30 minutes after treatment (**Figure 3.5H**) ($p = 0.002$ at 30 minutes, two-way ANOVA).

This suggested that the mechanism of ERG-induced potentiation is independent of the genes regulated by PARP1-mediated transcriptional activation. Focused expression analysis of genes involved in DNA damage repair pathways demonstrated that no significant change in any of repair genes analyzed, suggesting that the DNA damage phenomena is independent of changes to ERG-regulated gene expression (**Figure 3.5I**). To analyze the role of repair pathways directly, we tested the postulate that downregulation of a protein critical for the execution of non-homologous end joining (NHEJ) pathway such as XRCC4 would lead to a synergistic induction of DNA damage in a homologous recombination (HR) deficient cell. Treatment of HR-deficient HCC1937 cells with siRNA confirmed a greater increase in DNA damage following XRCC4 knockdown (NHEJ) than by XRCC3 knockdown (HR) ($p < 0.05$, one-way ANOVA). In contrast, the synergistic induction of DNA damage following XRCC4 or XRCC3 knockdown was not observed in PC3-ERG cells as compared to PC3-LACZ cells (**Figure 3.6A and 3.6B**). This suggested that ERG overexpression does not completely block either of these double strand break repair pathways. This was further confirmed by HR-efficiency assays which demonstrated that HR is not significantly altered by ERG overexpression (**Figure 3.6C**).

## DISCUSSION

In this chapter, we discovered that therapeutic inhibition of PARP1 preferentially sensitized ETS-overexpressing prostate cancer xenografts compared to ETS negative xenografts. Thus, similar to the successful paradigm of targeting the BCR-ABL gene fusion in CML with the small molecule kinase inhibitor imatinib meyslate (*34*), one could envision targeting the ETS-PARP1 axis in prostate cancer and possibly other ETS gene fusion-dependent cancers. While directly inhibiting transcription factors, such as ERG, may be difficult, blocking the function of regulatory co-factors, such as PARP1, is more feasible, and may represent a viable treatment paradigm in cancer therapy.

In particular, PARP1 represents a very promising therapeutic target. Based on its role in base excision repair, PARP1 has been explored in both preclinical and clinical settings as a target in tumors with deficiencies in double-stranded DNA repair, such as mutations in BRCA1 and BRCA2 (*6, 7*). In these cancers, the inhibition of PARP1 in cells with an inherent defect in homologous repair results in stalled replication forks and subsequent cell death (*6, 7*). An initial Phase I trial of the PARP inhibitor Olaparib has suggested an excellent therapeutic response in patients with BRCA1/2 deficient tumors from multiple organ sites with most patients experiencing a large reduction in total tumor volume (*8*). However, most cancers do not harbor BRCA1/2 mutations; only 5-6% of all breast cancers are associated with an inherited BRCA1/2 gene mutation (*35*) and only 3% of prostate tumors from an Ashkenazi Jewish population of 832 patients were BRCA1/2-deficient (*36*).

While PARP inhibitors can exploit the DNA repair defects of BRCA deficient tumors to induce cell death, we now demonstrate that they can also inhibit ETS gene fusion protein activity by preventing ETS transcriptional activity, inhibiting ETS-associated invasion, and enhancing ETS-mediated DNA damage. Future studies will help determine if, as with AR-mediated transcription (*37*), ETS-mediated transcription is directly coupled the induction of DNA damage. Importantly, the potentiation of ETS-induced DNA damage by PARP inhibition is of particular clinical interest; analogous to the "synthetic lethality" resulting from PARP inhibition in BRCA1/2-deficient tumors. By suggesting that cancers driven by specific oncogenic transcription factors may respond to PARP inhibition independent of BRCA1/2 status, our data supports the notion that multiple tumor subtypes will be susceptible to PARP pharmacotherapy. It is important to note that the company Sanofi-BiPAR recently released a press report that their phase III trial assessing the addition of their next generation PARP inhibitor, Iniparib, to a gemcitabine-carboplatin regimen for patients with metastatic triple-negative breast cancer, was negative for an overall survival benefit (*38*). This is in direct contrast to the recently reported phase II trial showing that the addition of Iniparib approximately doubled overall survival in this setting (*26*), and some questions about specificity have been raised (*39*). Nonetheless, these results highlight the importance of target selection; it is expected that ongoing phase III trials assessing chemotherapy with or without PARP inhibitor in BRCA-mutant cancers (instead of a non-specific triple-negative breast cancer population) will be positive, because the patient population

is selected based on the presence of the PARP inhibitor target – BRCA mutation (*40*). Here, we have shown that Olaparib very specifically, and in a dose dependent manner, delays tumor growth of ETS positive, but not ETS negative prostate cancer xenografts.

By exploiting the ETS:PARP1 interaction to selectively target ETS-overexpressing xenografts, our studies significantly expand the total population of patients who could benefit from PARP inhibition. Consequently, the data presented here also have implications on the design of subsequent clinical trials which will follow the recently reported Phase I trial of Olaparib (*8*). While most trials will undoubtedly be designed to target and subtype BRCA-deficient tumors, trials designed in organ sites that are also known to harbor aberrantly expressed ETS genes, such as breast, melanoma, Ewing's sarcoma and especially prostate, should also subtype the disease by ETS status. Based on the data presented here, ETS-positive tumors are expected to respond with a higher probability to PARP inhibition than ETS-negative tumors, potentially making ETS status an important predictive biomarker. In line with the observation that PARP inhibitors can significantly increase the mean overall survival of patients with triple negative breast cancer when added onto an existing regimen, our data suggests that the best design for a clinical trial in hormone-refractory metastatic prostate cancer will be to add PARP inhibitors in combination with chemotherapeutics known to potentiate the effects of PARP inhibition such as temozolomide.

Finally, the observation that gene fusions which drive the gross overexpression of ETS genes also induce DNA double strand break formation

provides additional mechanistic insight into how ETS gene fusions drive cancer progression. Specifically, by causing DNA double-strand breaks, ETS gene overexpression may also play a role in the gradual evolution of genomic rearrangements. This finding may explain why recurrent ETS gene fusions may have been difficult to discover, because they simply lead to the accumulation of additional gene fusions only some of which will drive disease progression. In fact, this model may partially explain the clinical behavior of prostate cancers which lie dormant for years only to spontaneously become extremely aggressive.

**FIGURES**

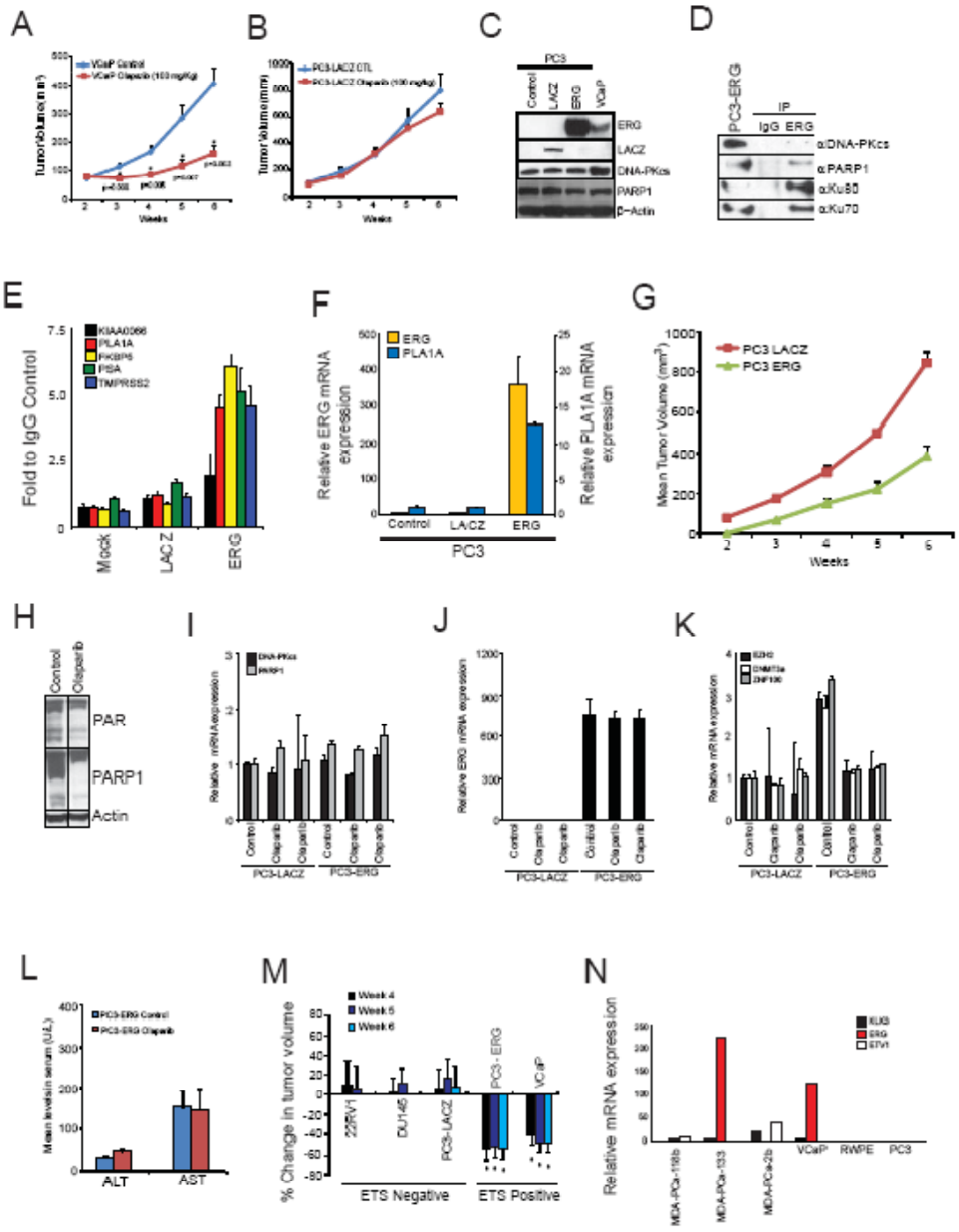



**Figure 3.1: Establishment of PC3-ERG model cell line and *in vivo* effects of Olaparib and characterization of primary prostate cancer serial xenograft model.**

**(A)** VCaP cells and **(B)** PC3-LACZ (ETS negative) cells were xenografted and mice were treated with Olaparib (100mg/kg, 5 days/week) as indicated.

**(C)** Western blot analysis of LACZ and ERG overexpression in PC3 model.

**(D)** IP-Western shows ERG interacts with DNA-PKcs, PARP1, Ku70 and Ku80 in PC3 cells.

**(E)** Chromatin immunoprecipitation assay using anti-ERG antibody. qPCR was performed for known ERG-regulated targets. Samples were input normalized before relative quantification.
**(F)** qPCR of ERG overexpression and increased expression of the ERG target gene *PLA1A*.
**(G)** PC3-LACZ and PC3-ERG xenograft growth curves ploted on same axis for comparison.
**(H)** Western blot analysis of ERG overexpressing PC3 cell xenografts treated with or without 40mg/kg Olaparib 4 hrs prior to harvesting tumor from mice.
**(I)** RNA harvested from 4 hour staged xenograft tumors as in (H) was analyzed by qPCR.
**(J)** and **(K)**, as in (I), analyzing for total ERG mRNA expression or several ERG target genes as indicated, respectively. All qPCR experiments were run in triplicate.
**(L)**, Serum from PC3-ERG mice treated with 40mg/kg Olaparib were assessed for alanine aminotransferase (ALT) and aspartate aminotransferase (AST) to determine liver toxicity. Normal range for ALT and AST levels in mouse serum: 28-132 and 59-247, respectively.
**(M)**, Percent reduction in tumor volume of xenografts from specificity screen.
**(N)**, qPCR assessment of ETS gene expression in the primary prostate cancer serial xenograft models and control cell lines. Normalizations were to RWPE expression for relative quantification.
For all experiments mean +/- SEM shown, * $p < 0.05$, ** $p < 0.01$ unless indicated.

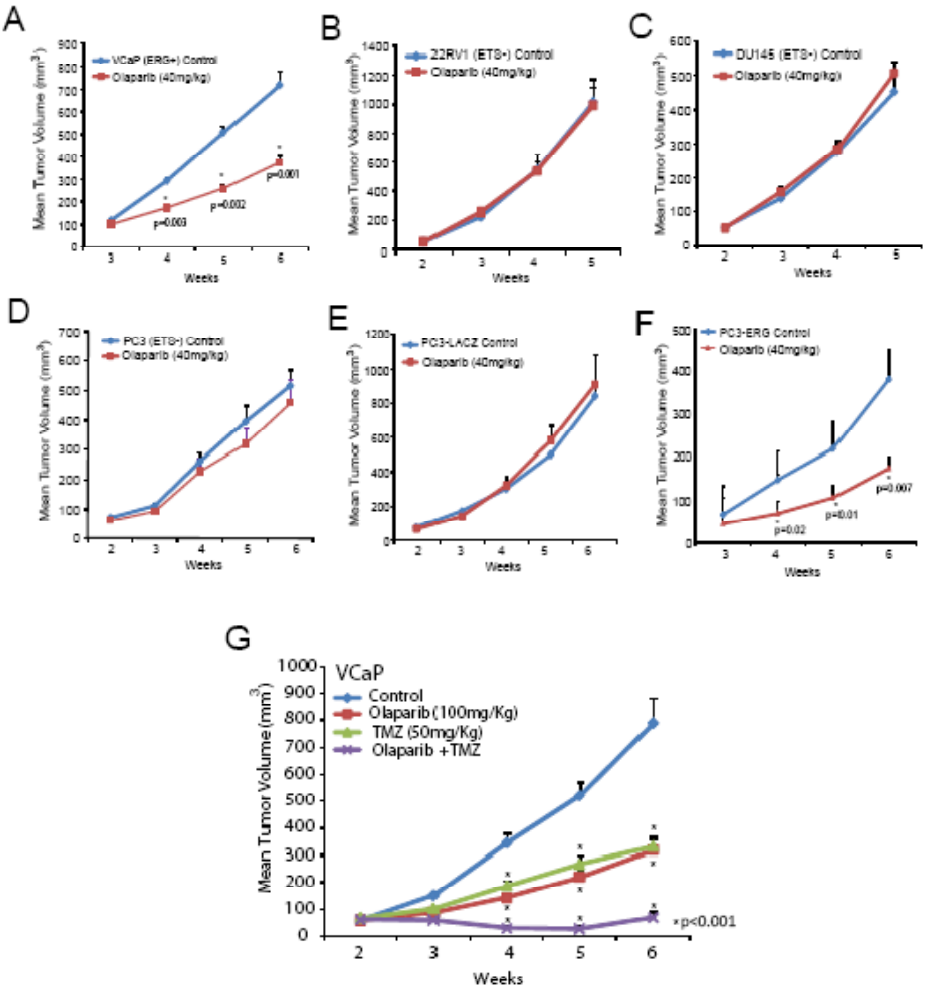



**Figure 3.2: Inhibition of PARP1 alters ETS positive, but not ETS negative cell line xenograft growth.**

**(A-F)** Specificity screen for ETS positive and ETS negative tumor cell line xenografts. Cell lines were injected subcutaneously and grown until tumors were palpable. Xenografted mice then received I.P. injections of Olaparib 40mg/kg as indicated 5 days/week. Caliper measurements were taken weekly. ETS positive cell line xenografts were **(A)** VCaP (ERG rearrangement) and **(F)** PC3-ERG cells and the ETS negative xenografts are **(B)** 22RV1, **(C)** DU145 and **(D, E)** PC3-Control/-LACZ, respectively.

**(G)** Mice xenografted with VCaP cells were treated as in as in (A) except with 100mg/kg Olaparib and/or 50 mg/kg Temozolomide (TMZ) as indicated. Olaparib was administered I.P. five days per week. TMZ was administered in two 5 day cycles with the first occurring during week 3 and the second occurring during week 5.

For all experiments mean +/- SEM shown, * $p < 0.05$, ** $p < 0.01$ unless indicated.

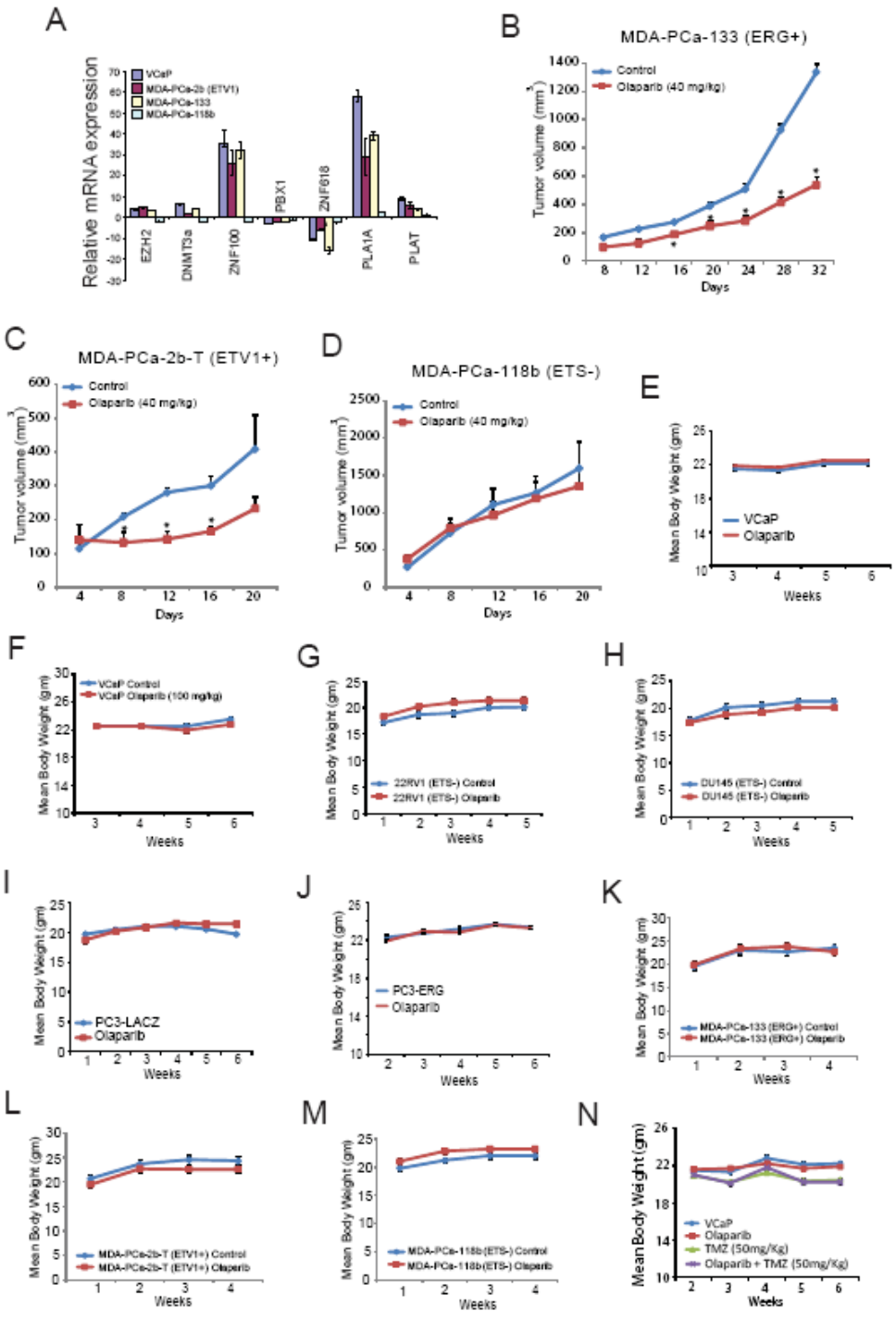



**Figure 3.3: Response of primary prostate cancer serial xenograft model to Olaparib and effects of Olaparib on mouse body weight in all xenograft models.**

**(A)**, qPCR was performed for several ETS-regulated target genes in primary prostate cancer xenograft models following treatment with Olaparib as indicated.

qPCR samples in were normalized to RWPE expression for relative quantification.
**(B), (C)**, and **(D)** MDA-PCa-133, MDA-PCa-2b-T and MDA-PCa-118b were treated with or without Olaparib as indicated.
**(E)** and **(F)** Mice from VCaP cell xenograft experiments treated with 40mg/kg Olaparib unless otherwise indicated were weighed at the time of caliper measurement.
**(G)**, **(H)**, **(I)** and **(J)**, As with (E), except 22RV1, DU145, PC3-LACZ and PC3-ERG models, respectively.
**(K)**, **(L)**, **(M)**, as with (E), except MDA-PCa-133, MDA-PCa-2b-T as well as MDA-PCa-118b primary human tumor xenografts (gray axis), respectively. Standard error of the mean is shown on the plot.
**(N)** VCaP cells treated with Olaparib and/or TMZ.
Bar graphs are shown with +/- SEM unless otherwise indicated, * $p < 0.05$, ** $p < 0.01$ unless indicated.

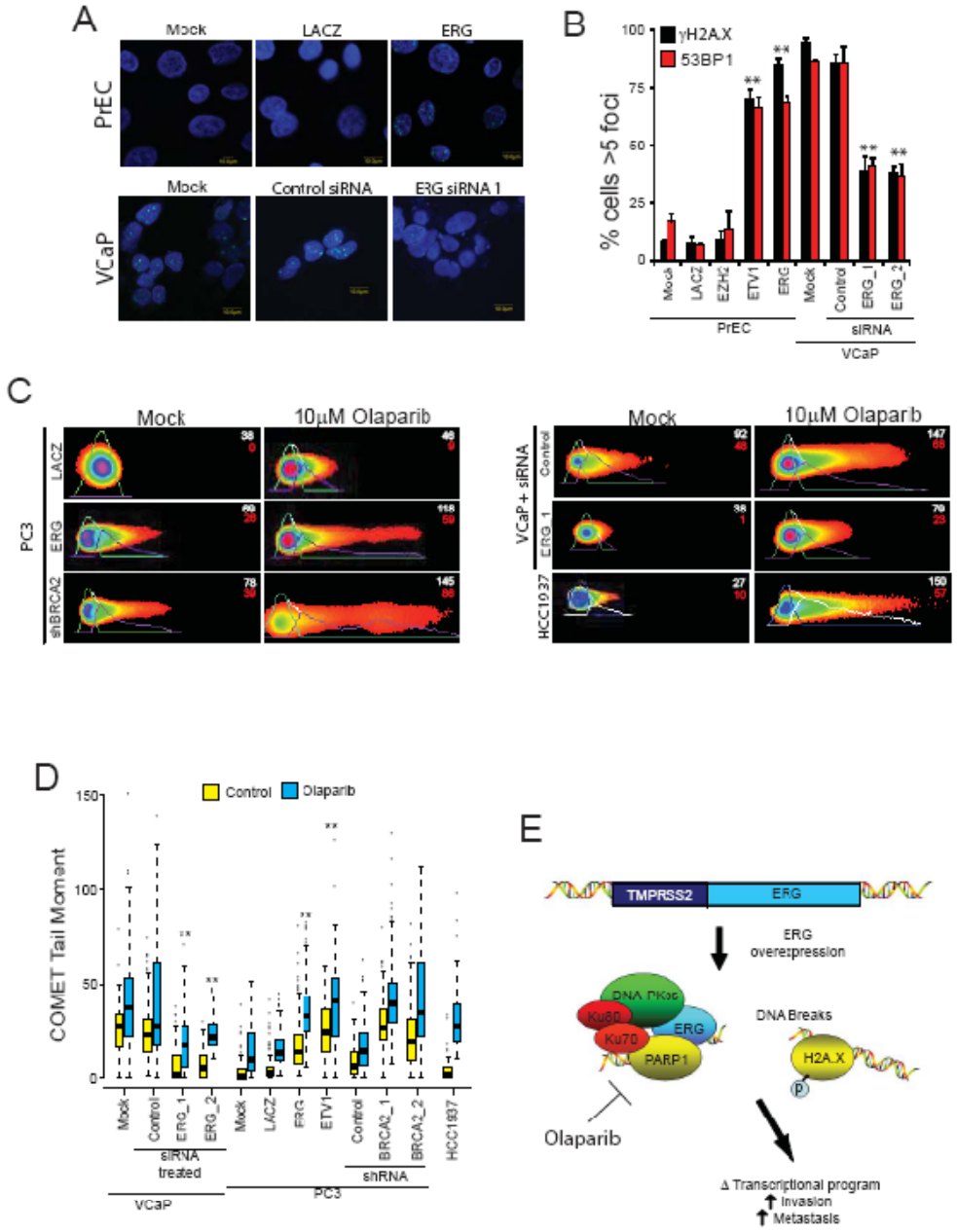



**Figure 3.4: ETS transcription factors induce DNA damage which is potentiated by PARP inhibition.**
**(A)** γ-H2A.X immunofluorescence staining shows that ERG induces the formation of γ-H2A.X foci. Top row, benign prostate epithelial cells(PrEC) were infected with lentiviruses expressing LACZ or ERG. Bottom row, VCaP cells treated with control or ERG siRNA.
**(B)** Quantification of γ-H2A.X and 53BP1 IF staining in PrEC or VCaP cells.
**(C)** ETS overexpression or BRCA2 knockdown (with shRNA) induces DNA damage as assessed by neutral COMET assay in VCaP cells. Cells were

treated with or without 10μM Olaparib for 48 hours. Cells with DNA damage have an extended "tail moment" of fragmented DNA shown in red. Relative tail length is shown in white. Representative images showing quantification of head and tail height, length and fluorescence intensity are shown (as indicated).
**(D)** Quantification of average COMET tail moments following treatment as noted in the box plot. Statistical tests were performed using the two-way ANOVA test (described in S.O.M.) to determine if the increase in DNA damage in Olaparib-treated ETS overexpressing cells (PC3-ERG, PC3-ETV1 and VCaP) was statistically greater than the increase observed in Olaparib-treated control cells with low ETS expression (PrEC-LACZ, PC3-LACZ or VCaP treated with ERG siRNA) as indicated in the text. Similar statistical tests were used to compare the increase in BRCA2 shRNA expressing cells to PC3 cells transduced with control shRNA ** $p < 0.01$
**(E)** Proposed model to therapeutically target ETS gene fusions via their interacting enzyme, PARP1.
All bar graphs are shown with +/- SEM unless otherwise indicated, * $p < 0.05$, ** $p < 0.01$.

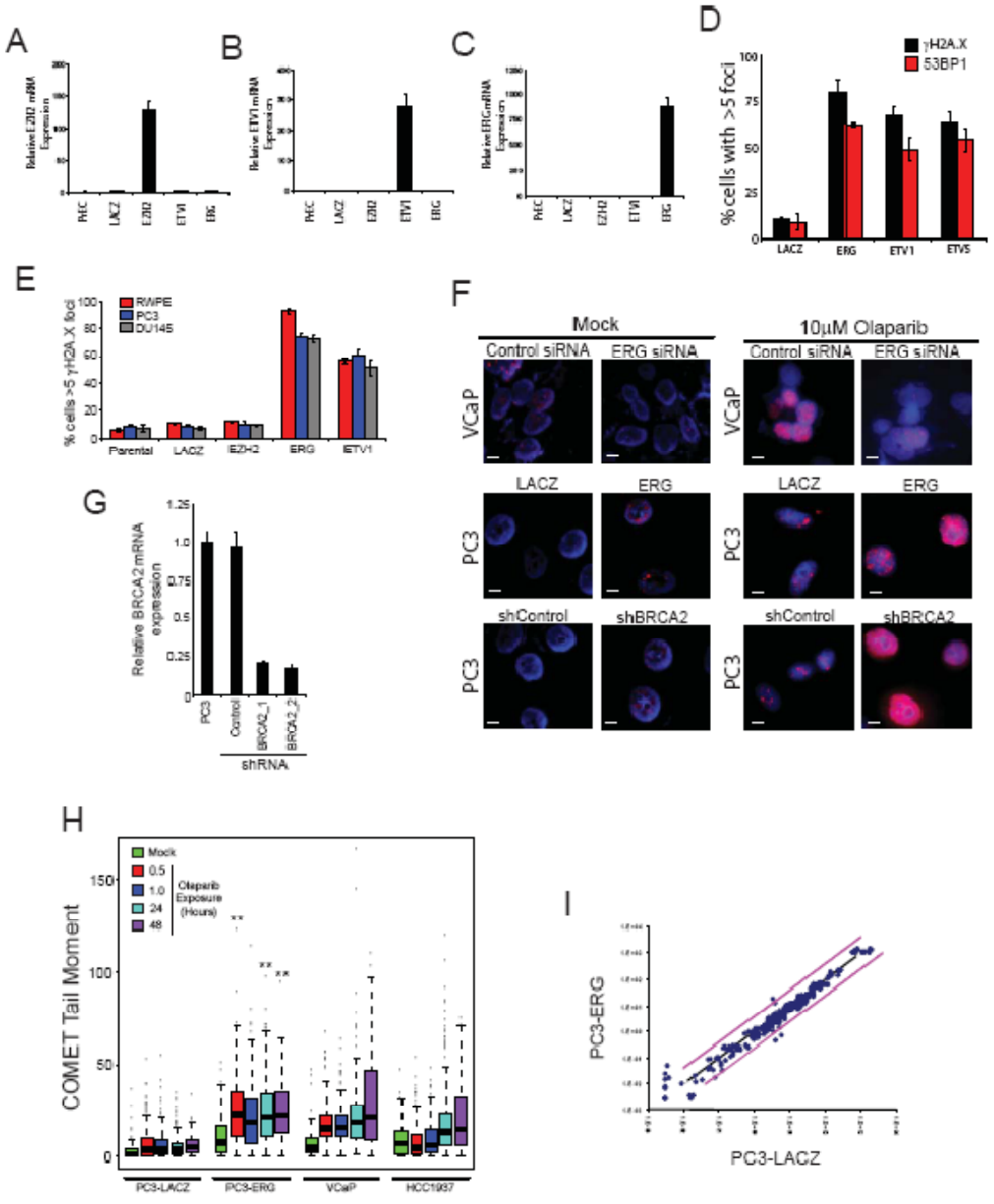



**Figure 3.5: ETS gene fusion proteins induce γH2A.X and 53BP1 foci formation.**

**(A)-(C)** PrEC cells treated with LACZ, EZH2, ETV1 or ERG lentivirus for 48 hours and used for γH2A.X imaging were assessed for changes in the total RNA expression level of **(A)** EZH2, **(B)**, ETV1 or, **(C)** ERG as compared to GAPDH control.

**(D)** RWPE cells transduced with lentivirus overexpressing ERG, ETV1 or ETV5 were analyzed by immunofluorescence.

**(E)** Various prostate cell lines were assessed for percentage of cells displaying γH2A.X foci following lentiviral infection as indicated. Expression changes were confirmed by qPCR (Data not shown).

**(F)** VCaP cells treated with siRNA or PC3 cell transduced with lentivirus as indicated were treated with or without Olaparib for 48 hours. Cells were stained with rabbit α-γH2A.X and goat a-rabbit IgG 555nm (red) was used as a secondary. Images were taken using 100x objective oil lens. Experiment was run three times. **(G)** PC3 cells transduced with lentivirus expressing BRCA2 shRNAs were analyzed for changes in BRCA2 mRNA expression as compared to GAPDH. Experiment was run in triplicate. ETS gene fusion proteins.
**(H)** Cells were treated with Olaparib for indicated times and harvested for COMET assays. A two-way ANOVA was used to compare Olaparib-induced increases between the isogenic PC3-LACZ and PC3-ERG cell lines at each time point.
**(I)** Targeted DNA damage qRT-PCR arrays were run with PC3-LACZ versus PC3-ERG overexpressing cells.
All assays were performed in triplicate and S.E.M. is shown unless otherwise indicated. * $p < 0.05$ ** $p < 0.01$

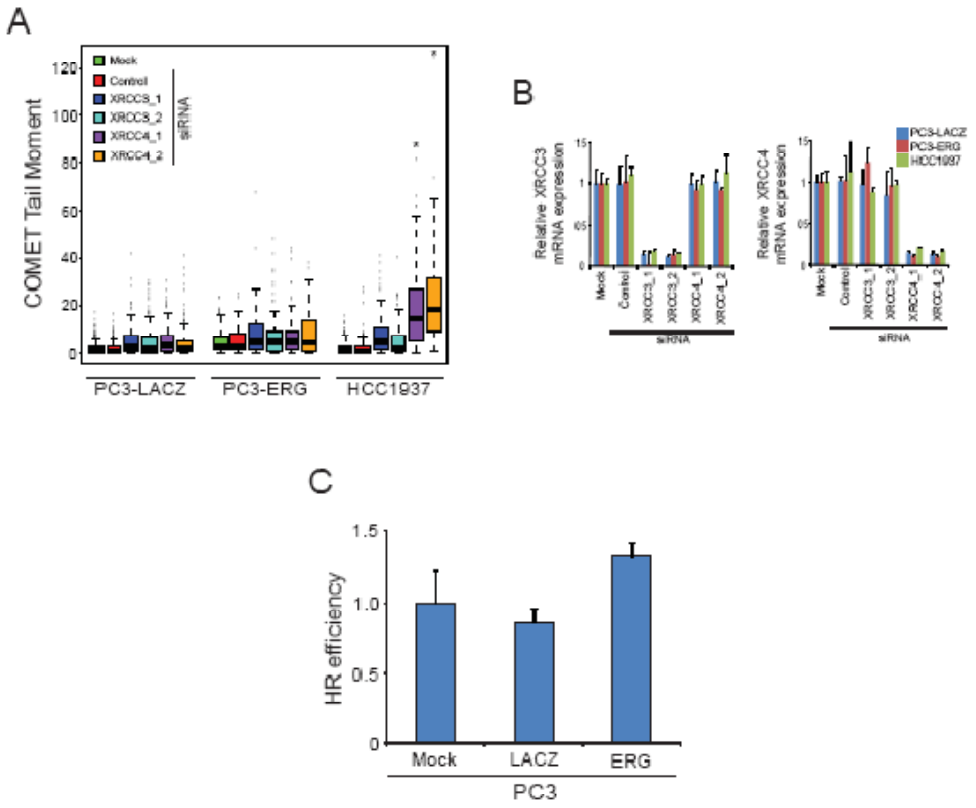



**Figure 3.6: ETS proteins for not alter homologous recombination efficiency in prostate cells.**
**(A)** COMET assays run on PC3-LACZ, PC3-ERG or HCC1937 (BRCA1 mutant) cells treated with siRNA for 48 hours as indicated. One way ANOVA was performed between mock and siRNA treated groups.
**(B)** qPCR quantification of knockdown efficiency of cells treated as in (J). All samples were run in quadruplicate three times.
**(C)** Homologous recombination efficiency assays were run by transfecting reporter constructs into stable cell lines. Twenty four hours later, DNA was extracted and repair efficiency was determined by qPCR. Transfection efficiency was controlled using primers in the construct backbones. All assays were performed in triplicate and S.E.M. is shown unless otherwise indicated. * $p < 0.05$ ** $p < 0.01$

## REFERENCES


1. J. A. Brown, R. B. Marala, *J Pharmacol Toxicol Methods* **47**, 137 (May-Jun, 2002).
2. T. Helleday, *Mol Oncol* **5**, 387 (Aug, 2011).
3. W. L. Kraus, *Curr Opin Cell Biol* **20**, 294 (Jun, 2008).
4. J. M. de Murcia *et al.*, *Proc Natl Acad Sci U S A* **94**, 7303 (Jul 8, 1997).
5. Z. Q. Wang *et al.*, *Genes Dev* **11**, 2347 (Sep 15, 1997).
6. H. E. Bryant *et al.*, *Nature* **434**, 913 (Apr 14, 2005).
7. H. Farmer *et al.*, *Nature* **434**, 917 (Apr 14, 2005).
8. P. C. Fong *et al.*, *N Engl J Med* **361**, 123 (Jul 9, 2009).
9. J. C. Brenner *et al.*, *Head Neck* **32**, 417 (Sep 16, 2009).
10. N. Palanisamy *et al.*, *Nat Med* **16**, 793 (Jul, 2010).
11. S. A. Tomlins *et al.*, *Neoplasia* **10**, 177 (Feb, 2008).
12. K. Yano *et al.*, *EMBO Rep* **9**, 91 (Jan, 2008).
13. J. Yu *et al.*, *Cancer Cell* **12**, 419 (Nov, 2007).
14. B. E. Helgeson *et al.*, *Cancer Res* **68**, 73 (Jan 1, 2008).
15. S. Varambally *et al.*, *Science* **322**, 1695 (Dec 12, 2008).
16. S. A. Tomlins *et al.*, *Science* **310**, 644 (Oct 28, 2005).
17. K. J. Livak, T. D. Schmittgen, *Methods* **25**, 402 (Dec, 2001).
18. C. f. B. E. a. R. Center for Drug Evaluation and Research, *Estimating the safe starting dose in clinical trials for therapeutics in adult healthy volunteers* (U.S. Food and Drug Administration Rockville, Maryland, USA, 2002), pp.
19. S. Reagan-Shaw, M. Nihal, N. Ahmad, *Faseb J* **22**, 659 (Mar, 2008).
20. Z. G. Li *et al.*, *J Clin Invest* **118**, 2697 (Aug, 2008).
21. D. W. Fairbairn, P. L. Olive, K. L. O'Neill, *Mutat Res* **339**, 37 (Feb, 1995).
22. P. Duez, G. Dehon, A. Kumps, J. Dubois, *Mutagenesis* **18**, 159 (Mar, 2003).
23. P. E. Verde *et al.*, *Mutat Res* **604**, 71 (Apr 30, 2006).
24. M. W. Audeh *et al.*, *Lancet* **376**, 245 (Jul 24, 2010).
25. A. Tutt *et al.*, *Lancet* **376**, 235 (Jul 24, 2010).
26. J. O'Shaughnessy *et al.*, *N Engl J Med* **364**, 205 (Jan 20, 2011).
27. S. Rottenberg *et al.*, *Proc Natl Acad Sci U S A* **105**, 17079 (Nov 4, 2008).
28. S. A. Tomlins *et al.*, *Nature* **448**, 595 (Aug 2, 2007).
29. C. K. Donawho *et al.*, *Clin Cancer Res* **13**, 2728 (May 1, 2007).
30. X. Liu *et al.*, *Mol Cancer Res* **6**, 1621 (Oct, 2008).
31. J. P. Palma *et al.*, *Clin Cancer Res* **15**, 7277 (Dec 1, 2009).
32. O. B. Isakoff S J, Tung N M, Gelman R S, Giranda V L, Bernhard K M, Habin K R, Ellisen L W, Winer E P and Goss P E, paper presented at the American Society of Clinical Oncology, Chicago, IL, June 4, 2010 2010.
33. E. J. Bennett, J. W. Harper, *Nat Struct Mol Biol* **15**, 20 (Jan, 2008).
34. B. J. Druker *et al.*, *N Engl J Med* **344**, 1038 (Apr 5, 2001).
35. K. E. Malone *et al.*, *Jama* **279**, 922 (Mar 25, 1998).
36. D. J. Gallagher *et al.*, *Clin Cancer Res* **16**, 2115 (Mar 9, 2010).
37. M. C. Haffner *et al.*, *Nat Genet* **42**, 668 (Jul 4, 2010).
38. Sanofi-aventis, "Results from Phase III Study with Iniparib (BSI-201) in Metastatic Triple-Negative Breast Cancer," *Press release*, January 27, 2011 2011.
39. L. A. Carey, N. E. Sharpless, *N Engl J Med* **364**, 277 (Jan 20, 2011).
40. L. W. Ellisen, *Cancer Cell* **19**, 165 (February 15, 2011, 2011).

# Chapter IV

# PARP-1 inhibition as a targeted strategy to treat Ewing's sarcoma


**ABSTRACT**

**Ewing's sarcoma family tumors (ESFTs) are aggressive malignancies which frequently harbor characteristic EWS-FLI1 or EWS-ERG genomic fusions. Here we report that these fusion products interact with the DNA damage response protein and transcriptional co-regulator PARP-1. ESFT cells, primary tumor xenografts and tumor metastases were all highly sensitive to PARP1 inhibition. Addition of a PARP1 inhibitor to the second-line chemotherapeutic agent temozolomide resulted in complete responses of all treated tumors in an EWS-FLI1-driven mouse xenograft model of ESFT. Mechanistic investigations revealed that DNA damage induced by expression of EWS-FLI1 or EWS-ERG fusion genes was potentiated by PARP1 inhibition in ESFT cell lines. Notably, EWS-FLI1 fusion genes acted in a positive feedback loop to maintain the expression of PARP1, which was required for EWS-FLI-mediated transcription, thereby enforcing oncogene-dependent sensitivity to PARP-1 inhibition. Our findings offer a strong preclinical rationale to target the EWS-FLI1: PARP1 intersection as a**

**therapeutic strategy to improve the treatment of Ewing's sarcoma family tumors. Finally, this study prompts immediate attention to reposition small molecule inhibitors of the DNA damage response protein PARP-1, a new class of experimental agents currently in trials in the adult oncology clinic, for treatment of a particularly aggressive and deadly class of pediatric sarcomas[1].**

## INTRODUCTION

The Ewing's sarcoma family of tumors (ESFTs) are malignant neoplasms of bone and soft tissue which frequently harbor reciprocal chromosomal translocations that result in the production of chimeric proteins containing the transcriptional activation domain from the Ewing's sarcoma breakpoint region 1 gene (*EWS*) fused to the DNA binding domain of an *ETS* transcription factor (*FLI1* or *ERG*) (*1, 2*). Dysregulation of these chimeric proteins have previously been implicated in abnormal proliferation, invasion, and tumorigenesis (*3, 4*). While the EWS-FLI or EWS-ERG fusion proteins represent potential disease-specific molecular targets, transcription factors such as these have been notoriously difficult to inhibit therapeutically. However, given the poor outcomes of patients with metastatic EFST (*5*), there is a clear need to improve therapy for these patients.

[1] The work presented in this chapter was published in *Cancer Research* and completed in collaboration with the following co-authors: Felix Y. Feng, Sumin Han, Sonam Patel, Siddharth V. Goyal, Laura M. Bou-Maroun, Meilan Liu, Robert Lonigro, John R. Prensner, Scott A. Tomlins, Arul M. Chinnaiyan

We have recently shown that an alternative strategy to inhibit transcription factor activity is to target critical interacting enzymes (*6*). Similar to ESFTs, 50% of prostate cancers harbor genomic rearrangements of *ETS* transcription factors (*7*). However, unlike the EWS-FLI or EWS-ERG fusions which produce a multifunctional chimeric protein, prostate cancer rearrangements usually place an androgen-regulated promoter upstream of an *ETS* gene (*ERG* or *ETV1*), resulting in increased *ETS* transcription factor activity by overexpression (*3*). We have previously demonstrated that PARP1 and DNA-dependent protein kinase (DNA-PKcs) are key ETS cofactors in prostate cancer, and therapeutic inhibition of PARP1 disrupts the growth of *ETS* positive, but not *ETS* negative, prostate cancer xenografts (*6*). Importantly, we demonstrated that this ETS protein interaction axis is mediated by the ETS DNA binding domain (*6*). While this interaction site is present in the EWS-FLI1 and EWS-ERG chimeras, these fusion proteins activate oncogenic pathways independent of their DNA binding domain (*8*). As such, despite the presence of the ETS DNA binding domain, the effectiveness of PARP inhibitors against these multifunctional gene fusion proteins is unclear.

Mechanistically, PARP1 is a functionally complex enzyme that has been shown to both drive transcription and to accelerate base excision repair (*9-11*). PARP1 inhibitors have demonstrated promising activity in early clinical trials (*12*), particularly in BRCA-mutant cancers defective in homologous repair (HR), in which they may cause replication fork stalling and subsequent synthetic lethal cell death (*13, 14*). We now hypothesize that the *EWS-FLI* and *EWS-ERG* gene

fusions in ESFTs depend on the activity of PARP1 and may be targeted with PARP1 inhibition.

## MATERIALS AND METHODS

### Cell Lines and inhibitors

PC3 (ATCC), CADO-ES1 (DSMZ), RD-ES (ATCC), SAOS-2 (ATCC), A-204 (ATCC), COG10 and COG258 (www.COG.org) Ewing's sarcoma cell lines were grown in RPMI 1640 (Invitrogen, Carlsbad, CA) supplemented with 10% FBS (Invitrogen) in 5% $CO_2$ cell culture incubator. All cultures were also maintained with 50 units/ml of Penicillin/streptomycin (Invitrogen). The genetic identity of each cell line was confirmed by genotyping samples. Briefly, DNA samples from each cell line were diluted to 0.10ng/µl and analyzed in the University of Michigan DNA sequencing Core using the Profiler Plus PCR Amplification Kit (Applied Biosystems, Foster City, CA) as previously described (*15*). Olaparib was synthesized to >99% purity by Axon Biochem. EWS-FLI1_3xFlag and EWS-ERG_3xFlag constructs were created by gene synthesis (Invitrogen) and cloned into the pLL_IRES_GFP vector backbone available from the University of Michigan vector core. Virus and stable isogenic cell lines were made as previously described (*6*). Cells were monitored for GFP expression and construct overexpression was confirmed by Western blot analysis.

**Immunoprecipitation**

Cells were lysed in Triton X-100 lysis buffer (20mM MOPS, pH 7.0, 2mM EGTA, 5mM EDTA, 30mM sodium fluoride, 60mM β-glycerophosphate, 20mM sodium pyrophosphate, 1mM sodium orthovanadate, 1% Triton X-100, protease inhibitor cocktail (Roche, #14309200)). Cell lysates (0.5-1.0mg) were then pre-cleaned with protein A/G agarose beads (Santa Cruz, # sc-2003) by incubation for 1 hour with shaking at room temperature followed by centrifugation at 2000xg for 1 minute. Lysates were incubated with 100μg/mL ethidium bromide (Sigma) as previously described (*16*). Then, after adding 2μg anti-ERG antibody (Epitomics, Burlingame, CA), anti-PARP-1 polyclonal (1:1000 in blocking buffer, Santa Cruz Biotech Cat #sc-8007, Santa Cruz, CA or anti-FLI1 (Santa Cruz C19, Cat# sc-356), lysates were incubated at 4°C overnight with shaking prior to addition of 20μL protein A/G agarose beads. The mixture was then incubated with shaking at 4°C for another 4 hours prior to washing the lysate-bead precipitate (centrifugation at 2000xg for 1 minute) 3 times in Triton X-100 lysis buffer. Beads were finally precipitated by centrifugation, resuspended in 20μL of 2x loading buffer and boiled at 80°C for 10 minutes for separation the protein and beads. Samples were then analyzed by SDS-PAGE Western blot analysis as described below.

**Immunoblot Analysis**

The cell lines were plated in two wells of a 6-well plate at 250,000 cells/mL 24hours prior to harvesting by trypsinization. Pellets were then flash frozen,

briefly sonicated and homogenized in NP40 lysis buffer (50 mM Tris-HCl, 1% NP40, pH 7.4, Sigma, St. Louis, MO), and complete proteinase inhibitor mixture (Roche, Indianapolis, IN). Ten micrograms of each protein extract were boiled in sample buffer, size fractionated by SDS-PAGE, and transferred onto Polyvinylidene Difluoride membrane (GE Healthcare, Piscataway, NJ). The membrane was then incubated overnight at 4ºC in blocking buffer [Tris-buffered saline, 0.1% Tween (TBS-T), 5% nonfat dry milk] and incubated for 4 hours at room temperature with the following: anti-DNA-PKcs mouse monoclonal (1:1000 in blocking buffer, BD Biosciences #610805, San Jose, CA), anti-Ku70 mouse monoclonal (1:1000 in blocking buffer, BD Biosciences #611892), anti-Ku80 rabbit polyclonal (1:1000 in blocking buffer, Cell Signaling Cat #2180S, Danvers, MA), anti-PARP-1 polyclonal (1:1000 in blocking buffer, Santa Cruz Biotech Cat #sc-8007, Santa Cruz, CA, anti-ERG (Epitomics, Burlingame, CA), anti-FLI1 (Santa Cruz C19, Cat# sc-356) and anti-GAPDH rabbit polyclonal antibody (1:1000, Millipore, Cat # ABS16). Following a wash with TBS-T, the blot was incubated with horseradish peroxidase-conjugated secondary antibody and the signals visualized by enhanced chemiluminescence system as described by the manufacturer (GE Healthcare).

**Soft Agar Colony Formation**

A solid layer of 1% agarose was used to coat each well of a 96-well plate. After allowing two hours to cool, five hundred prostate or Ewing's cells were suspended in a 33% agarose (10% solution), 33% media with 2x FBS and 33%

cell suspension in quadruplicate wells of a 96-well plate. Cells were treated with or without Olaparib in an upper layer of media. Every three days, the media was exchanged for fresh Olaparib containing media. After 3-weeks in culture, images were taken on a 10x objective lens and colonies were counted for each condition. This assay was run in three times.

**Chemosensitivity Assays**

Five thousand cells were plated in each well of a 96-well plate in sets of ten. Cells were then treated with a single dose of Olaparib as indicated for 72 hours. WST assays (Roche) were performed according to company protocol. Briefly, 10% WST was added to each well, the plates were incubated at 37C with 5% $CO_2$ for 4 hours and the OD was measured.

**small RNA interference**

Knockdown of specific genes was accomplished by RNA interference using commercially available siRNA duplexes for EWS, ERG, FLI1, DNA-PKcs or PARP-1 (Dharmacon, Lafayette, CO). At least 4 independent siRNAs were screened for knockdown efficiency against each target and the best siRNA was selected. In each case, two independent siRNAs were used. We were unable to identify a 3'-FLI1 siRNA with >80% knockdown efficiency, therefore, these siRNA transfections were run 3 times. Transfections were performed with OptiMEM (Invitrogen) and oligofectamine (Invitrogen) as previously described (*17*).

**Quantitative Real-Time PCR Assays**

Total RNA was isolated from cell lines transfected with siRNA or as indicated (Qiagen). Quantitative PCR (QPCR) was performed using SYBR Green dye on an Applied Biosystems 7300 Real Time PCR system (Applied Biosystems, Foster City, CA) as described (*7*). The relative quantity of the target gene was completed for each sample using the ΔΔCt method by the comparing mean Ct of the gene to the average Ct of the geometric mean of two housekeeping genes, GAPDH and RPL19 (*18*). All oligonucleotide primers were synthesized by Integrated DNA Technologies (Coralville, IA). The primer sequences for the transcript analyzed are provided in **Appendix 1** unless used previously (*6*).

**Basement Membrane Matrix Invasion Assays**

For invasion assays, the cell lines were transfected with siRNA, drug or negative controls as indicated. Forty-eight hours post-treatment, cells were seeded onto the basement membrane matrix (EC matrix, Chemicon, Temecula, CA) in the chamber insert with 8.0μM pores of a 24-well culture plate in serum free media. Cells were attracted to the lower chamber by the addition of complete media as a chemoattractant. After 48 hours incubation at 37˚C with 5% $CO_2$, the non-invading cells and EC matrix were gently removed with a cotton swab. Invasive cells, which were located on the lower side of the membrane, were stained with crystal violet, air dried and photographed. To quantify the relative number of invaded cells, colorimetric assays were performed by treating the inserts with

150µl of 10% acetic acid (v/v) and measuring absorbance of each condition at 560nm using a spectrophotometer (GE Healthcare).

**Xenografts**

Five week-old male SCID mice (CB.17. SCID), were purchased from Charles River, Inc. (Charles River Laboratory, Wilmington, MA). RD-ES cells ($2 \times 10^6$ cells/injection), PC3-luciferase-LACZ ($1 \times 10^6$ cells) or PC3-luciferase-EWS-FLI ($1 \times 10^6$ cells) stable cells were resuspended in 100µl of saline with 50% Matrigel (BD Biosciences, Becton Drive, NJ) and were implanted subcutaneously into the left and right flank regions of the mice. Mice were anesthetized using a cocktail of xylazine (80-120 mg/kg, IP) and ketamine (10mg/kg, IP) for chemical restraint before tumor implantation. All tumors were staged for two weeks before starting the drug treatment. At the beginning of the third week, mice with RD-ES tumors (10 tumors per treatment group, average size 150-200 $mm^3$) were treated with Olaparib (100mg/kg, IP twice daily five times per week) and/or Temozolamide (50mg/kg daily five times per week, IP). Similarly, at the beginning of the third week, mice with PC3-luciferase-LACZ or PC3-luciferase-EWS-FLI were treated with Olaparib (100mg/kg, IP twice daily five times per week). Olaparib was obtained from Axon Medchem (Groningen, The Netherlands). We chose 100mg/kg/twice daily because, in mice, this dose corresponds to the maximum clinically achievable dose for which minimal side effects were observed in humans during the initial Phase I trial of Olaparib (Center for Drug Evaluation and Research, 2002; Fong et al., 2009; Reagan-Shaw et al., 2008). Temozolamide

was obtained from Selleck Chemicals (Houston, Texas). Growth in tumor volume was recorded weekly by using digital calipers and tumor volumes were calculated using the formula $(\pi/6)$ $(L \times W^2)$, where L = length of tumor and W = width. Loss of body weight during the course of the study was also monitored weekly; all changes in mouse weights occurred within tolerable limits as set by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan. At the end of the xenograft studies, mice were sacrificed, and mouse lungs were subsequently fixed in 10% formalin via gavage through a tracheal incision. Lungs were subsequently harvested, maintained in formalin, and subsequently embedded in paraffin. The number of metastases per lung was quantitated via light microscopy by a pathologist (S.A.T.). All procedures involving mice were approved by the UCUCA at the University of Michigan and conform to their relevant regulatory standards.

**Immunofluorescence**

Cells were seeded at 25,000 cells/mL in 4-well chamber slide 24 hours prior to the addition of siRNA or drug as indicated. Merbarone (Sigma) was dissolved in DMSO and stored at 4°C. Cells were then washed once in PBS, fixed for 15 minutes in 100% methanol, washed in PBS, permeabilized for 5 minutes in PBS containing 0.2% Triton-X, washed twice in PBS, blocked for 30 minutes in PBS containing 0.5% donkey serum (Sigma, St. Louis, MO), held for 45 minutes in PBS containing 0.5% donkey serum and primary antibody: anti-γ-H2A.X mouse monoclonal (Millipore Cat #05-636) washed 3 times with PBS, held for 30

minutes in PBS containing 0.5% donkey serum and secondary antibody: Alexa Fluor 488-anti-Mouse (Invitrogen). Cells were then stained with DAPI for 5 minutes and slides were mounted 12 hours prior to analysis using Vectashield (Vector laboratories, Burlingame, CA). Images were taken using 100x oil lens on an Olympus Confocal microscope at the University of Michigan microscopy imaging lab.

**COMET Assay**

Ewing's sarcoma cell lines were seeded at 250,000 cells/mL in a 6-well plate 24 hours prior to treatment with siRNA, drug or vehicle control. After 48 hours, cells were trypsinized, harvest by centrifugation and re-suspend in PBS. Cell counts were then normalized to 1 x $10^5$ cells/mL. Suspended cells (25µL) were then mixed with 250µL 1.0% ultrapure low melting point agarose (Invitrogen) made in 1x Tris-Borate buffer. The agarose-cell mixture was then dropped onto slides allowed to solidify at 4°C in the dark for 20 minutes before immersion in COMET assay lysis solution (Trevigen, Gaithersburg, MD) at 4°C in the dark for 45 minutes. Excess buffer was then removed and slides were submerged in freshly prepared neutral solution (Tris Base 60.57g, Sodium Acetate 204.12g, dissolve in 450 ml of dH2O. Adjust to pH = 9.0 with glacial acetic acid.) at room temperature in the dark for 40 minutes. Slides were then washed twice by immersing in 1 X TBE buffer prior to neutral electrophoresis at 20 volts for 60 minutes unless otherwise indicated. Slides were then fixed in 70% ethanol for 5 minutes. Following air drying of the agarose, slides were stained with SYBR

Green dye (Invitrogen) and images were collected with a 10x and 40x objective lens. COMET tail moments were then assessed using COMETscore.v1.5 (AutoCOMET.com, Sumerduck, VA) image processing software as described by the manufacturer with greater than 100 cells analyzed in triplicate experiments.

**Promoter Reporter Luciferase Assays**

Cells were transfected with PARP1, emptry vector (R01) or GAPDH promoter reporter constructs (Switchgear Genomics, Menlo Park, CA) using Fugene 6.0 (Roche) according to standard protocol. Luciferase activity was then assessed using the Lightswitch Assay kit available (Switchgear Genomics) according to the manufacturer's instructions.

**Oncomine expression analysis**

The expression signatures from Ewing's sarcoma patients (*19*) and prostate cancer patients (*20, 21*) were downloaded from the Oncomine and used to generate scatter plots. Correlation coefficients were determined using Oncomine (*22-25*).

**Statistical Analysis**

All statistical assays were two-tailed t-test unless otherwise indicated. On all plots and graphs the standard error of the mean is shown. * $P < 0.05$. For xenograft experiments, comparison of tumor volume growth rates between treatment groups was achieved via likelihood ratio tests. Volume measurements were modeled

using linear mixed-effects models incorporating a random intercept for each mouse.

**RESULTS**

To test the hypothesis that EWS fusions interact with PARP1, we performed IP-immunoblot analysis on cell lines naturally harboring the rearrangements. Using an ERG or FLI1 antibody that recognizes the EWS-ERG and EWS-FLI1 fusion proteins, respectively, we pulled down both PARP1 and DNA-PKcs in a DNA-independent manner from two cell lines, CADO-ES1 (EWS-ERG) or RD-ES (EWS-FLI1) (**Fig. 4.1A**). Reverse IP from RD-ES cells performed using a PARP1 antibody enriched for EWS-FLI1 and DNA-PKcs in a DNA-independent manner (**Fig. 4.1B**), confirming the interaction between EWS fusions and PARP1.

We then tested the sensitivity of ESFT cells to the PARP1 inhibitor Olaparib. To analyze the sensitivity of ESFT cells overexpressing the *EWS-FLI1* and *EWS-ERG* fusions, compared to the prostate cancer harboring the *TMPRSS2-ERG* fusion, we performed soft agar colony formation assays on PC3 cells in which we had stably overexpressed these fusion products. Overexpression of any of these three *ETS* fusion genes caused sensitivity to Olaparib ($p < 0.05$ at 3 and 10μM), and the *EWS-FLI1* and *EWS-ERG* fusions were more sensitive than the *TMPRSS2-ERG* fusion ($p < 0.05$ at 1μM) (**Fig. 4.1C**). Soft agar assays performed using CADO-ES1, RD-ES or VCaP (*TMPRSS2-ERG* rearranged prostate cancer) cells confirmed that while all *ETS* positive cell lines were more susceptible to

Olaparib than the *ETS* negative cell line PC3, ESFT cell lines were even more sensitive than VCaP cells ($p < 0.05$) (**Fig. 4.1D**). Importantly, Olaparib did not have an effect on an osteosarcoma cell line, SAOS-2, or a rhabdomyosarcoma cell line, A-204, which were used as sarcoma model controls without *ETS* rearrangement (**Fig. 4.1D**).

Given that the vast majority of patients with Ewing sarcomas present with overt or subclinical metastases, front-line therapy for this disease includes an intense regimen of five cytotoxic chemotherapy agents. While this regimen can be effective in the short term, patients can also quickly relapse with even more aggressive disease. Additionally, because none of these front-line therapeutics are tailored to target the *EWS-FLI1* or *EWS-ERG* fusion products, it is possible that the recurrences from these patients have activated multiple general drug resistance pathways. For these reasons, we tested the Olaparib sensitivity of two Ewing's cell lines called COG10 and COG258 that were derived from patients that had relapsed from front-line chemotherapy. Surprisingly, these cell lines were just as sensitive to PARP inhibition as the other Ewing's cell lines tested and were still more sensitive than VCaP cells ($p < 0.5$ as indicated) (**Fig. 4.1D**).

Given the sensitivity of ESFTs to PARP inhibition, we sought to explore the mechanism underlyng this sensitivity. In prostate cancer, *TMPRSS2-ERG* overexpression causes DNA damage, which is then potentiated by PARP inhibition (*6*). Surprisingly, Ewing's cells demonstrated significantly more γH2A.X foci (a marker of DNA double strand breaks) per cell than *TMPRSS2-ERG* positive VCaP cells ($p < 0.05$ for all EFST cells) (**Fig. 4.2A**) and 1μM

Olaparib caused massive induction of γH2A.X foci (**Fig. 4.3A**). Neutral COMET assays confirmed that natural EFST cells have more DNA double strand breaks per cell than VCaP cells ($p < 0.05$) (**Fig. 4.2A**). Importantly, overexpression of EWS-FLI1 in PC3 cells led to increased DNA breaks ($p < 0.05$) (**Fig. 4.3B**) and knockdown of either *EWS-FLI1* or *EWS-ERG* fusions caused a reduction in COMET tail moment, and significantly decreased the induction of DNA damage after Olaparib treatment ($p < 0.05$) (**Fig. 4.2B and Figs. 4.4A, 4.4B, 4.5A and 4.5B**) suggesting that Olaparib-potentiated DNA damage is due to expression of the *EWS-FLI1* and *EWS-ERG* fusion genes. Inhibition of non-homologous end-joining (NHEJ) can synergistically potentiate DNA damage in the context of deficient homologous recombination (HR). However, siRNA knockdown of XRCC4 (needed for NHEJ) had little effect (**Fig. 4.6**), suggesting that EFST cells have functional HR.Although PARP1 inhibition dramatically accentuated DNA damage, it did not have an effect on short-term cell viability (**Fig. 4.7A and 4.7B**).

We next assessed the effect of PARP1 inhibition on short-term cell invasion. Intriguingly, siRNA against *EWS*, *PARP1* or *DNA-PKcs* reduced CADO-ES1 and RD-ES cell invasion ($p < 0.05$), while knockdown of key components of DNA repair pathways (*XRCC1* for base excision repair, *XRCC3* for HR and *XRCC4* for NHEJ) had no effect (**Fig. 4.2C and Figs. 4.4A and 4.4B**), supporting a DNA repair-independent role of the enzymes. Importantly, Olaparib disrupted ESFT cell invasion ($p < 0.05$) at approximately one log-fold

lower concentration than VCaP cell invasion (*6*) and did not affect ETS negative, control sarcoma cell invasion (**Fig. 4.2D and Figs. 4.4A and 4.4B**).

Given the *in vitro* data demonstrating preferential sensitivity of ESFT cell lines to PARP1 inhibition, we next assessed the ability of the *EWS-FLI1* gene fusion to sensitize xenografts to PARP inhibition, since this fusion accounts for the largest percentage of *ETS* fusions in Ewing's Sarcoma. Specifically, we established xenografts from our isogenic PC3-EWS-FLI and PC3-LACZ cells, and treated these xenografts with Olaparib. We chose the PC3 system as our proof-of-principle isogenic model because 1) PC3-LACZ xenografts are resistant to Olaparib (*6*), 2) the PC3 cell line was derived from a prostate cancer bone metastasis, a site commonly infiltrated by Ewing's Sarcomas, and, 3) the Ewing's Sarcoma "cell-of-origin" is currently unclear (*26*).While Olaparib had almost no effect on controlling PC3-LACZ xenograft growth, Olaparib significantly impeded PC3-EWS-FLI1 xenograft growth ($P < 0.001$) (**Fig. 4.8A and 4.8B**) demonstrating that *EWS-FLI1* overexpression alone is sufficient to sensitize PC3 xenografts to Olaparib.

Next, because PARP1 inhibitors are unlikely to be used clinically as a monotherapy, we also combined PARP1 inhibition with the DNA alkylating agent temozolomide (TMZ) to treat a highly aggressive ESFT xenograft model (RD-ES) harboring an endogenous *EWS-FLI1* rearrangement. TMZ potentiates the effects of PARP1 inhibitors (*27*) and is currently being assessed in Ewing's sarcoma patients as second-line therapy (*28*). Surprisingly, TMZ alone only delayed RD-ES xenograft growth ($p < 0.001$) (**Fig. 4.8C**). However, Olaparib also caused

significant delay in RD-ES xenograft growth ($p < 0.001$), and, when administered in combination with TMZ, led to an immediate tumor regression and a highly significant sustained complete response. Mice were maintained without therapy for an additional 30 days without any observable recurrence ($p < 0.001$). Furthermore, we harvested lungs from RD-ES xenografted mice, assessed metastasis and found only 1/5 mice treated with Olaparib developed lung metastasis as compared to 6/8 untreated mice ($p = 0.058$) (**Fig. 4.8D**). This is the first mouse model demonstrating that PARP1 activity is required for metastasis.

Given that PARP1 has roles in both DNA repair and transcriptional regulation (*9, 10*), we hypothesized that the sensitivity of ESFT xenograft growth and metastasis depends on both PARP1-mediated roles. Consistent with this, knockdown of PARP1 or DNA-PKcs, but not other DNA repair proteins, disrupted transcriptional activity to a similar extent as fusion targeting *EWS* siRNA as evidenced by disruption of consensus EWS-FLI1 target genes (*29*) (**Fig. 4.9A and Figs. 4.4A, 4.4B**).

In modulating EWS-FLI1 and EWS-ERG expression, we observed that both EWS fusions maintain *PARP1* mRNA expression levels ($p < 0.05$) (**Figs. 4.4A, 4.5A, 4.10A and 4.10B**). This was unexpected because the *ETS* gene fusions found in prostate cancer do not regulate *PARP1* mRNA expression (*6*), but consistent with reports demonstrating that ETS transcription factors drive PARP1 expression in ESFTs (*30, 31*). Likewise, in EFST cells, knockdown of the EWS-FLI1 fusion led to a decrease in PARP1 protein expression and promoter activity ($p < 0.05$), while ERG siRNA did not have these effects in VCaP cells

(**Fig. 4.9B**). Conversely, EWS-FLI1 expression was not altered by Olaparib in RD-ES cells (**Figs. 4.10C and 4.10D**). Finally, Oncomine tumor microarray compendium analysis demonstrated that *PARP1* mRNA expression correlates with *FLI1* expression in EFST patient samples ($p < 0.01$, Spearman correlation shown in figure), but did not correlate with *ERG* expression in prostate cancer (**Fig. 4.9C**). Taken together, our data suggest that PARP1 is critical to a positive-feedback loop in EFST, in which the EWS-FLI fusion product drives expression of PARP1, which then further promotes transcriptional activation by EWS-FLI.

## DISCUSSION

In conclusion, our data demonstrate that ESFT cells and xenografts are sensitive to PARP1 inhibition, to a significantly greater degree than their *ETS*-fusion-positive prostate cancer counterparts. This increased sensitivity may stem from both greater potentiation of DNA damage by PARP1 inhibition, as well as a EWS-FLI-PARP1 positive feedback loop in transcriptional activation (**Fig. 4.9D**). Our findings suggest that PARP1 inhibition should be explored as a strategy for targeted therapy in Ewing's sarcoma, which may represent a tumor type that may be as profoundly affected by PARP inhibitors as *BRCA1/2* mutated carcinomas (*12-14*).

**FIGURES**

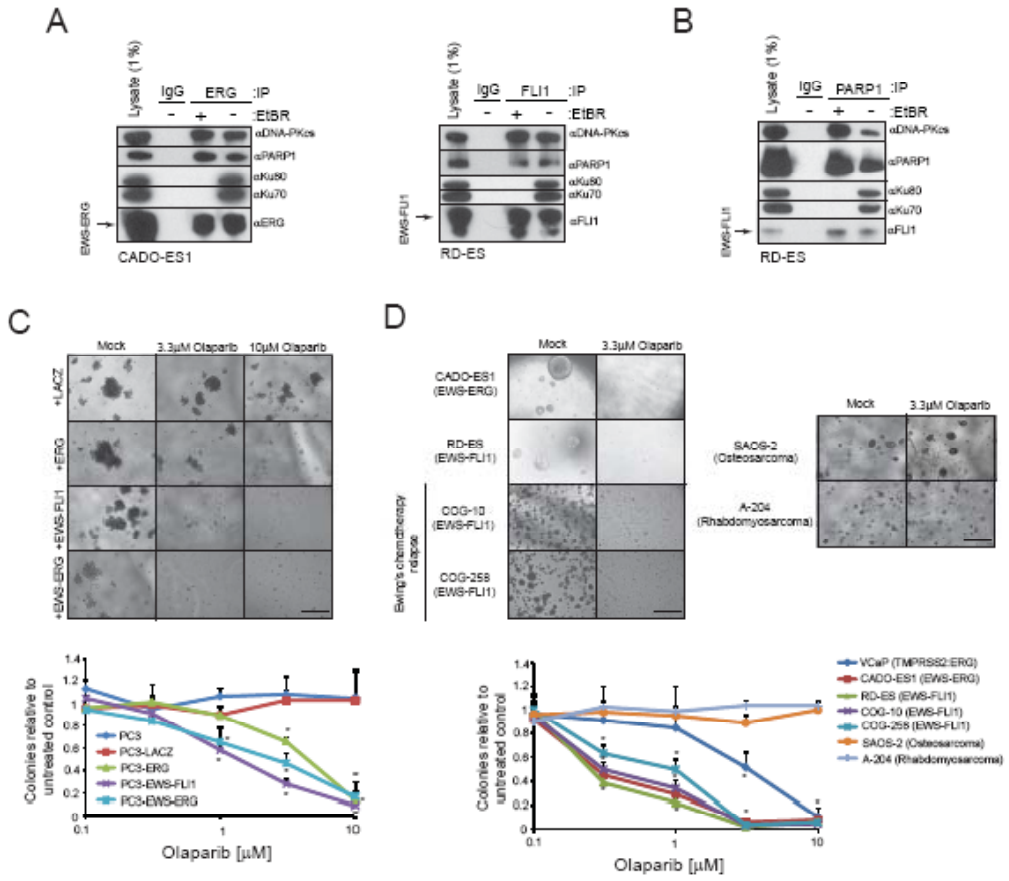



**Figure 4.1: Ewing's sarcoma gene fusion products (EWS-FLI1 and EWS-ERG) interact with PARP1 and DNA-PKcs, and confer sensitivity to PARP1 inhibition.**
**(A)** Immunoprecipitation (IP)-immunoblot analysis performed on CADO-ES1 (*EWS-ERG*) cells using an ERG antibody and RD-ES (*EWS-FLI1*) cells using a FLI1 antibody. IPs were treated +/-100μg/mL ethidium bromide.
**(B)** As in (A), except using a PARP1 antibody on RD-ES cells.
**(C)** Soft agar colony formation assays were performed using PC3 cells, overexpressing ERG, EWS-FLI1, EWS-ERG, or LACZ. Colonies counted after 3 weeks in culture. Representative photomicrographs shown. Scale bar = 50μM.
**(D)** As in (C), except in ETS-rearranged cells (VCaP, RD-ES, CADO-ES1, COG10 and COG258) and ETS-negative sarcoma controls (SAOS-2 and A-204). Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

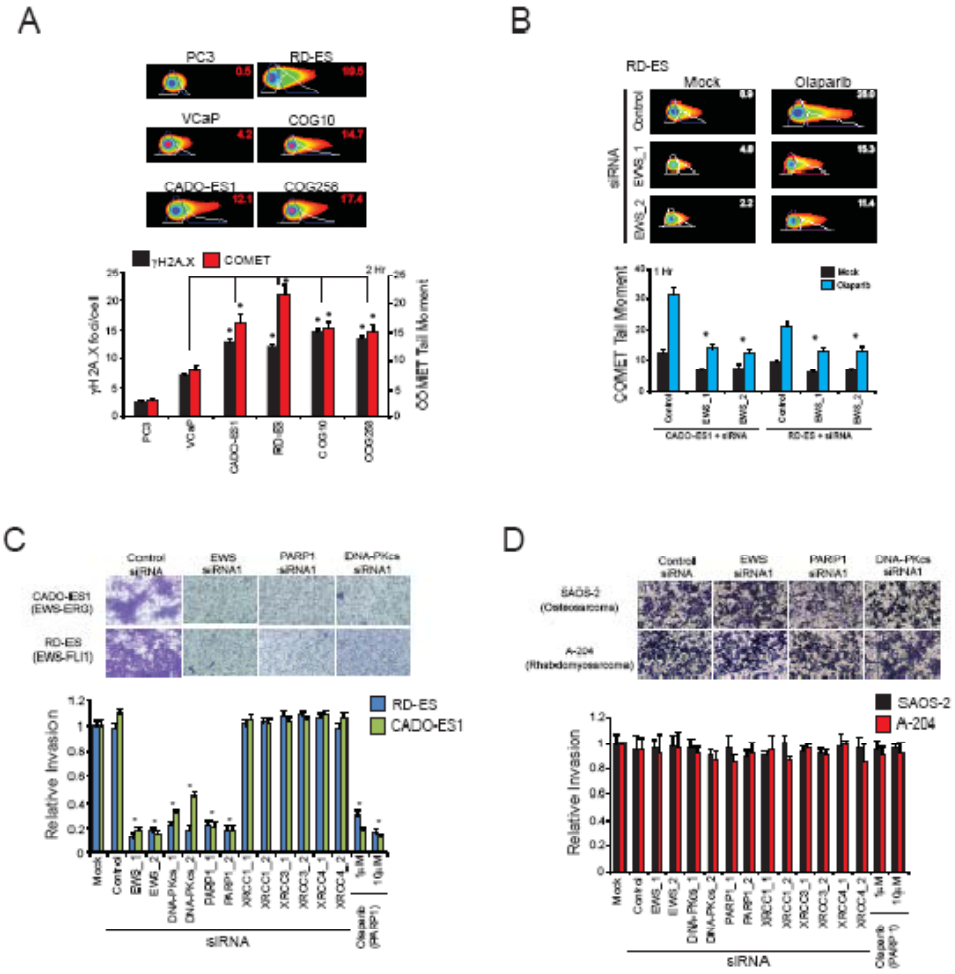



**Figure 4.2: Ewing sarcoma gene fusion products confer increased DNA damage and cell invasion, which can be modulated by genetic or pharmacologic approaches.**

**(A)** Staining and quantification of γH2A.X foci in prostate and EFST cells shown on left axis. On right axis, average tail moments from neutral COMET assays are shown. Representative COMET tails are shown, with calculated tail moments depicted numerically.

**(B)** Neutral COMET assays performed on cells treated with siRNA and/or 1μM Olaparib as indicated.

**(C)** Boyden chamber transwell migration assays were performed 48 hours after siRNA transfection or Olaparib treatment. For siRNA experiments, cells were plated in serum-free media and incubated for another 48 hours before fixing and staining. Representative photomicrographs are shown.

**(D)** As in (C), except sarcoma control cell lines that do not harbor EFST gene fusions were used.

Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

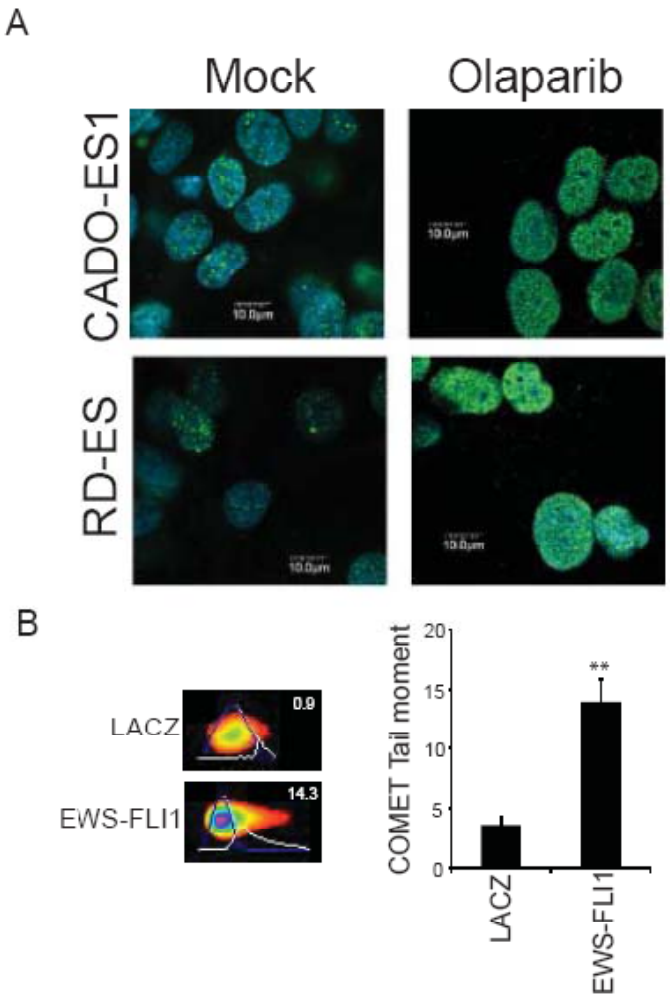



**Figure 4.3: DNA breaks are potentiated by Olaparib in Ewing's sarcoma cell lines and by overexpression of EWS-FLI1 in PC3 cells.**
**(A)** CADO-ES1 and RD-ES cells were treated with or without 1μM Olaparib as indicated. Cells were stained for γH2A.X foci. Confocal images were collected using a 60x Oil objective and representative images are shown.
**(B)** PC3 cells transduced with either LACZ or EWS-FLI1 lentivirus were subjected to neutral COMET assays. Representative images are shown on the left with COMET tail moments indicated in white. Relative DNA damage is quantified on the right.
Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

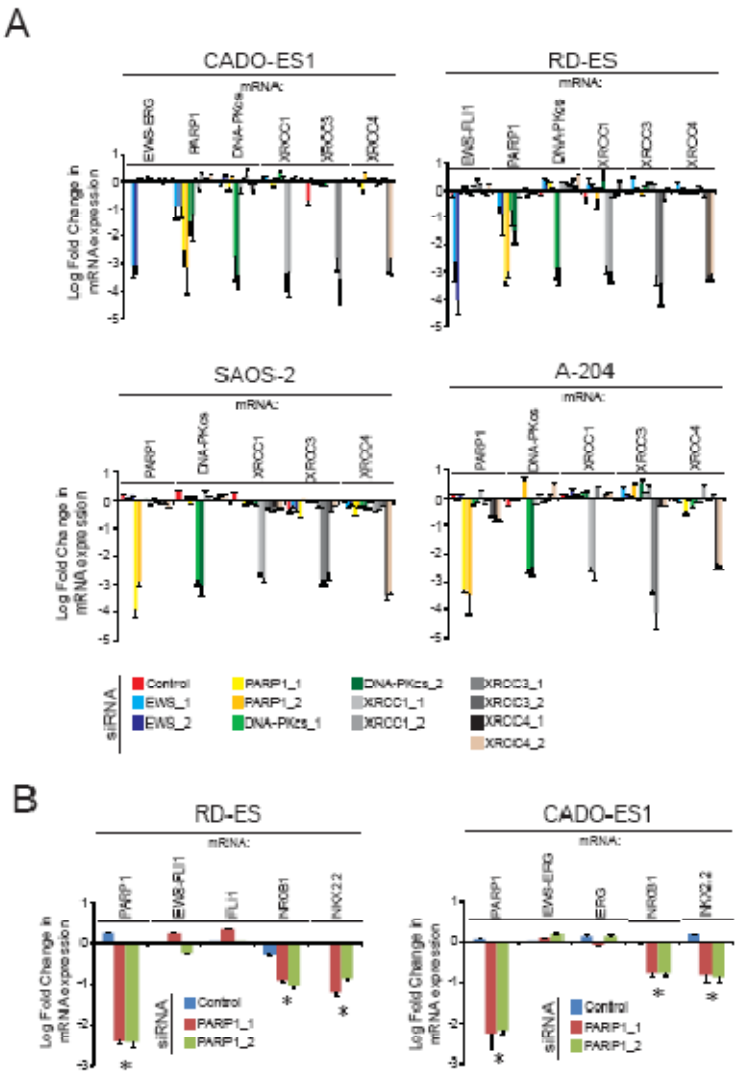



**Figure 4.4: Confirmation of knockdown.**
**(A)** qPCR analysis of CADO-ES1, RD-ES, SAOS-2 or A-204 cells treated with or without siRNA as indicated. Experiments were run in parallel to COMET and invasion assays.
**(B)** Additional mRNA targets were tested in CADO-ES1 and RD-ES cell lines.
Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

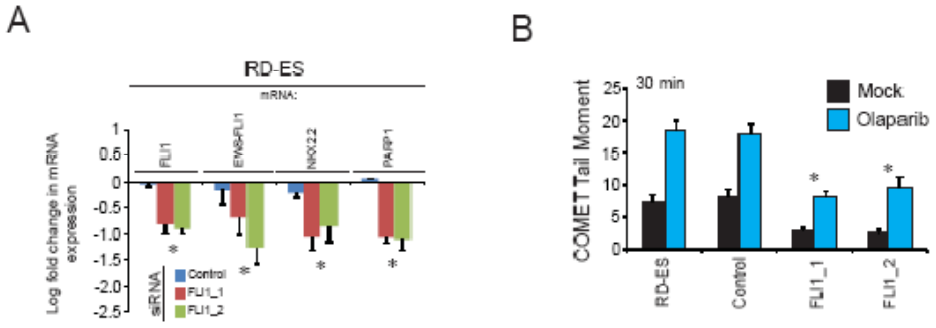



**Figure 4.5: FLI1 siRNA replicates EWS siRNA phenotype in EWS-FLI1 rearranged cells.**
**(A)** qPCR analysis of knockdown efficiency in RD-ES cells 48hours after treatment with siRNA as indicated.
**(B)** Neutral COMET assays were performed in parallel to the qPCR assays using RD-ES cells treated with or without 10μM Olaparib as indicated.
Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

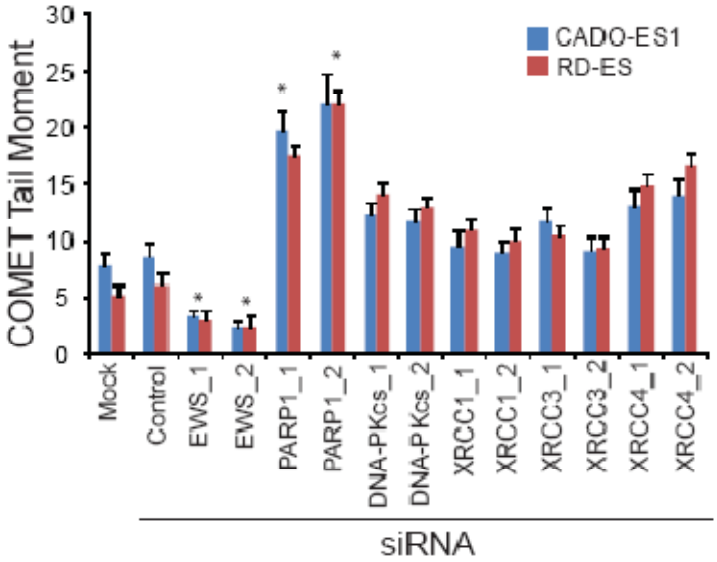



**Figure 4.6: Depletion of PARP1 potentiates the relative number of DNA double strand breaks in Ewing's Sarcoma.**
Cells were treated with siRNA for 48 hours as indicated. Neutral COMET assays were then performed to assess the relative level of DNA damage.
Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

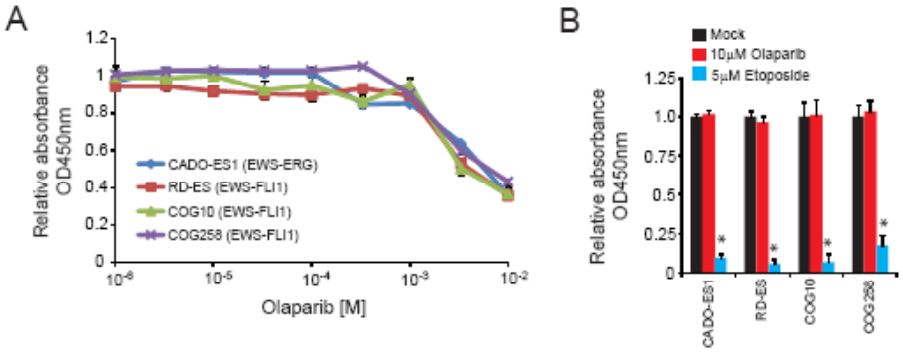



**Figure 4.7: Olaparib chemosensitivity of Ewing's cell lines.**
Various Ewing's cell lines were plated in 96-well plates, allowed 24 hours to attach and treated with different doses of Olaparib as indicated. WST assays were then performed after an additional 72 hours **(A)** or after 5 days **(B)**. IC50s were not significantly different.
Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

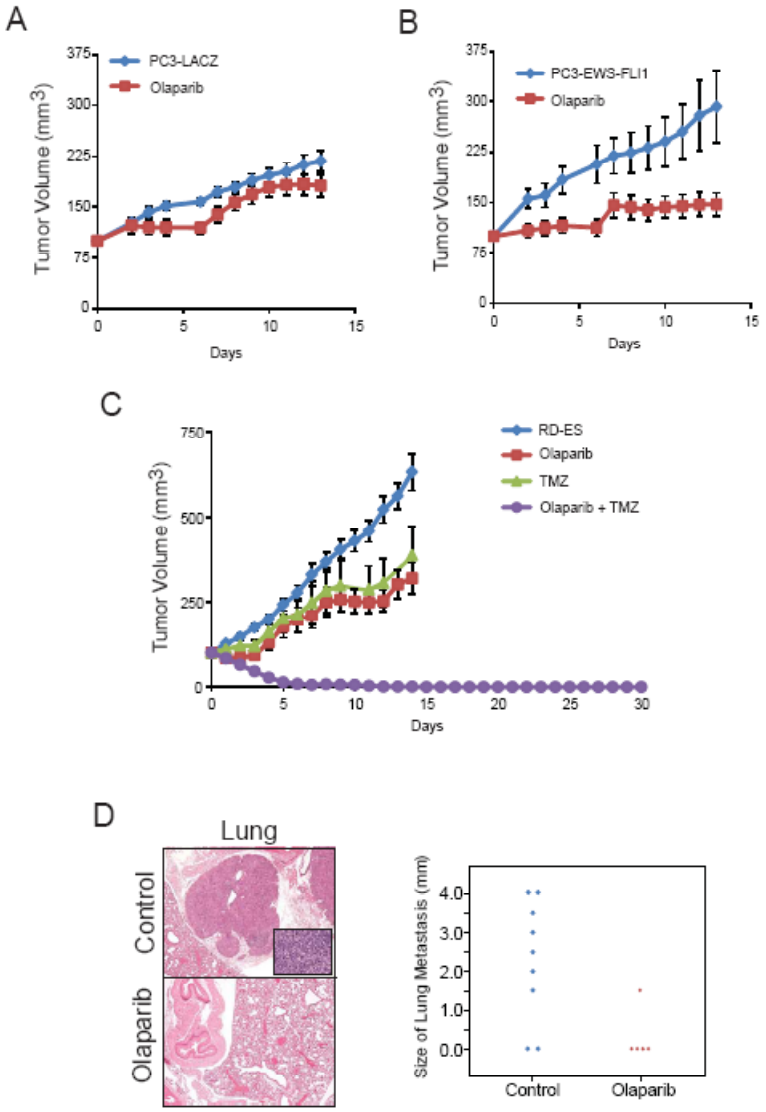



**Figure 4.8: PARP inhibition prevents EWS-FLI1 positive xenograft growth and metastasis.**

**(A)** Two weeks after engraftment, PC3-LACZ xenografts were treated with 100mg/kg Olaparib BID. Tumor volume measurements were recorded daily.

**(B)** As in (A), except with PC3-EWS-FLI1 xenografts.

**(C)** As in (A), except RD-ES were treated with Olaparib (100mg/kg BID) alone, TMZ (50mg/kg QD) alone or in combination.

**(D)** At the RD-ES xenograft assay endpoint, lungs were analyzed for metastases. Representative images were taken with a 4x objective (inset at 40x). In the plot, each dot represents an individual lung metastasis for different treatments as indicated along the x-axis. Metastasis diameter was measured and depicted along the y-axis.

Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

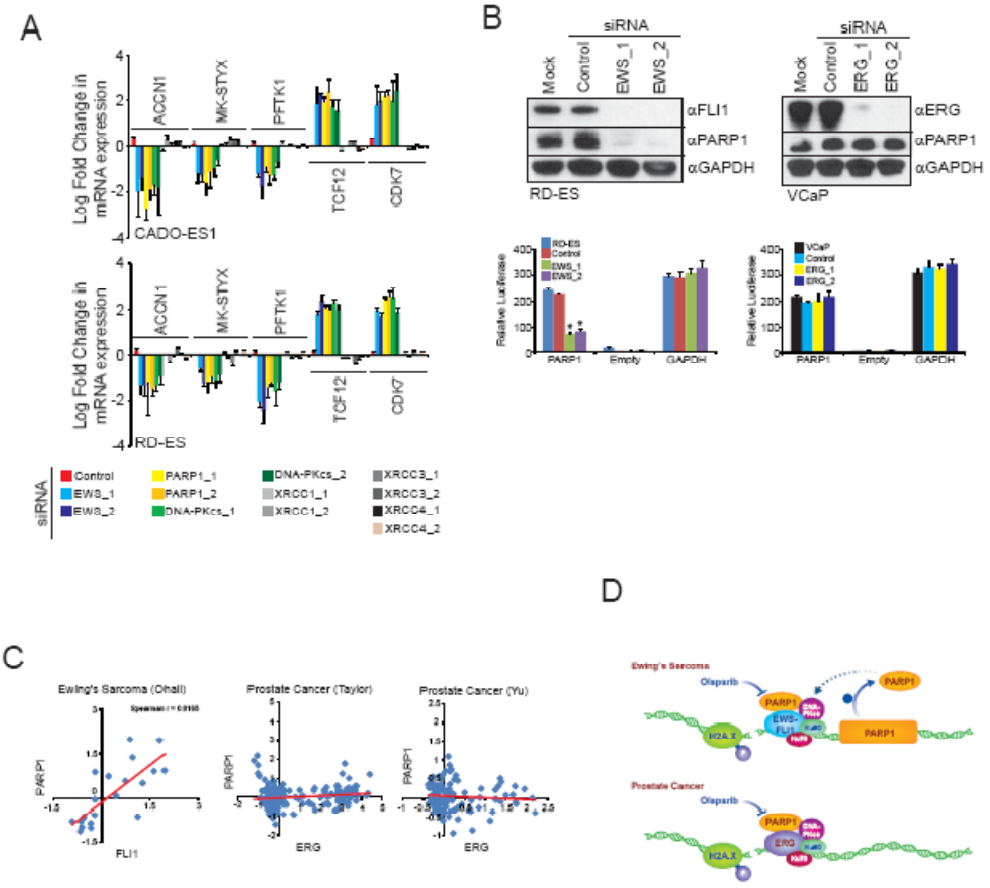



**Figure 4.9: EWS-FLI1 maintains a feed forward loop that drives PARP1 expression.**

**(A)** CADO-ES1 or RD-ES cells were treated with siRNA for 48 hours. qPCR was then performed for several *EWS-ETS* target genes.

**(B)** Immunoblot analysis of cells treated with control or *EWS* siRNA. Bar graphs are *PARP1* promoter reporter luciferase assays were performed on cells treated with siRNA for 24 hours and then transfected with the promoter reporter for an additional 24 hours.

**(C)** Oncomine scatter plots of publically available gene expression data sets. Each point represents gene expression values from an individual patient. First author of each published study is indicated and referenced in the S.O.M..

**(D)** Model depicting proposed mechanism of increased PARP sensitivity of ESFTs as compared to ETS positive prostate cancers.

Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

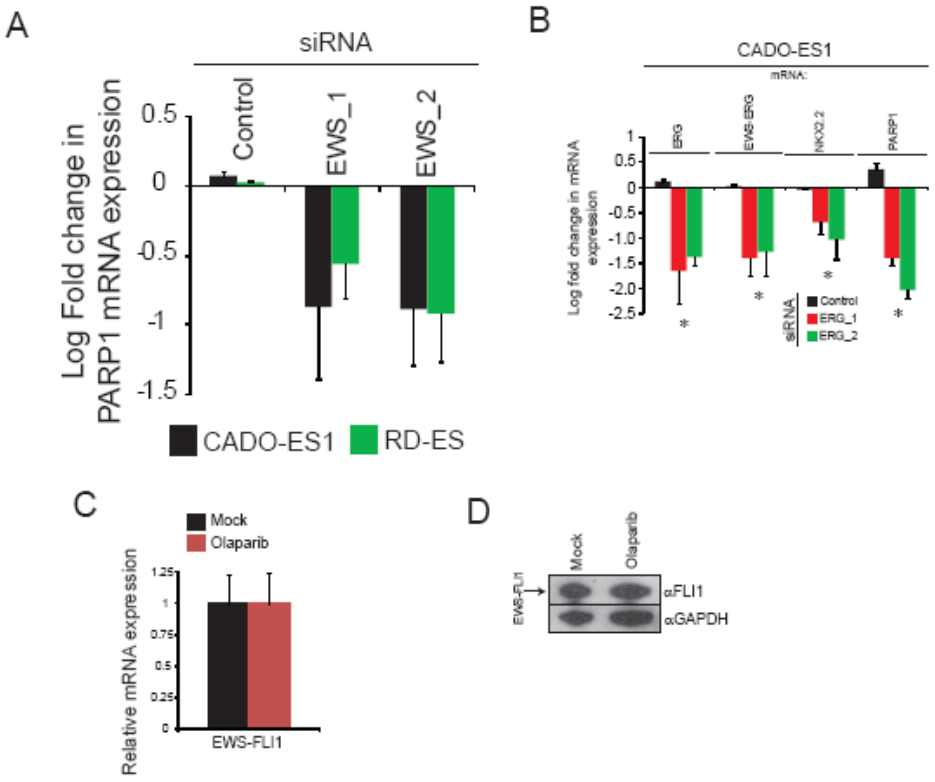



**Figure 4.10: EWS-FLI1 knockdown causes a decrease in PARP1 mRNA expression.**
**(A)** CADO-ES1 and RD-ES Ewing's cell lines were plated in 6-well plates, allowed 24 hours to attach and treated with different siRNAs as indicated. RNA was isolated and qPCR was performed for PARP1 mRNA expression.
**(B)** qPCR analysis of knockdown efficiency in CADO-ES1 cells 48 hours after treatment with siRNA as indicated.
**(C)** qPCR analysis for changes in EWS-FLI1 mRNA expression following treatment with Olaparib.
**(D)** As in (C), except Western blot analysis.
Error bars indicate S.E.M. of three replicates, * $p < 0.05$, ** $p < 0.01$.

## REFERENCES


1. S. A. Burchill, *J Clin Pathol* **56**, 96 (Feb, 2003).
2. E. de Alava *et al.*, *J Clin Oncol* **16**, 1248 (Apr, 1998).
3. J. C. Brenner, A. M. Chinnaiyan, *Biochim Biophys Acta* **1796**, 201 (Dec, 2009).
4. S. A. Tomlins *et al.*, *Nature* **448**, 595 (Aug 2, 2007).
5. C. R. Pinkerton *et al.*, *Eur J Cancer* **37**, 1338 (Jul, 2001).
6. J. C. Brenner *et al.*, *Cancer Cell* **19**, 664 (May 17, 2011).
7. S. A. Tomlins *et al.*, *Science* **310**, 644 (Oct 28, 2005).
8. S. M. Welford, S. P. Hebert, B. Deneen, A. Arvand, C. T. Denny, *J Biol Chem* **276**, 41977 (Nov 9, 2001).
9. M. Y. Kim, S. Mauro, N. Gevry, J. T. Lis, W. L. Kraus, *Cell* **119**, 803 (Dec 17, 2004).
10. R. Krishnakumar *et al.*, *Science* **319**, 819 (Feb 8, 2008).
11. M. Rouleau, A. Patel, M. J. Hendzel, S. H. Kaufmann, G. G. Poirier, *Nat Rev Cancer* **10**, 293 (Apr, 2010).
12. P. C. Fong *et al.*, *N Engl J Med* **361**, 123 (Jul 9, 2009).
13. H. E. Bryant *et al.*, *Nature* **434**, 913 (Apr 14, 2005).
14. H. Farmer *et al.*, *Nature* **434**, 917 (Apr 14, 2005).
15. J. C. Brenner *et al.*, *Head Neck* **32**, 417 (Apr, 2010).
16. K. Yano *et al.*, *EMBO Rep* **9**, 91 (Jan, 2008).
17. S. Varambally *et al.*, *Science* **322**, 1695 (Dec 12, 2008).
18. K. J. Livak, T. D. Schmittgen, *Methods* **25**, 402 (Dec, 2001).
19. A. Ohali *et al.*, *Oncogene* **23**, 8997 (Nov 25, 2004).
20. B. S. Taylor *et al.*, *Cancer Cell* **18**, 11 (Jul 13, 2010).
21. Y. P. Yu *et al.*, *J Clin Oncol* **22**, 2790 (Jul 15, 2004).
22. D. R. Rhodes *et al.*, *Nat Genet* **37**, 579 (Jun, 2005).
23. D. R. Rhodes *et al.*, *Neoplasia* **9**, 166 (Feb, 2007).
24. D. R. Rhodes *et al.*, *Neoplasia* **9**, 443 (May, 2007).
25. D. R. Rhodes *et al.*, *Neoplasia* **6**, 1 (Jan-Feb, 2004).
26. P. P. Lin, Y. Wang, G. Lozano, *Sarcoma* **2011** (2011).
27. C. K. Donawho *et al.*, *Clin Cancer Res* **13**, 2728 (May 1, 2007).
28. L. M. Wagner *et al.*, *Pediatr Blood Cancer* **48**, 132 (Feb, 2007).
29. C. Siligan *et al.*, *Oncogene* **24**, 2512 (Apr 7, 2005).
30. V. A. Soldatenkov *et al.*, *Oncogene* **18**, 3954 (Jul 8, 1999).
31. R. E. Newman, V. A. Soldatenkov, A. Dritschilo, V. Notario, *Oncol Rep* **9**, 529 (May-Jun, 2002).

# Chapter V

# Development of peptidomimetic inhibitors of the ERG transcription factor in prostate cancer


**ABSTRACT**

**Transcription factors play a key role in the development of a number of cancers, and therapeutically targeting them has remained a challenge. In prostate cancer, the ETS transcription factor ERG is recurrently rearranged and likely plays a critical role in prostate oncogenesis. Here we identified a consensus peptide inhibitor from a phage display library which interacts specifically with the ETS DNA binding domain of ERG. The interactive interface was mapped to 9-amino acid residues in the 3$^{rd}$ $\alpha$-helix of the ETS domain, which is critical for ERG activity as a transcription factor. X-ray crystallography confirmed the interaction residues are accessible for interaction. ERG inhibitory peptides (EIPs) were found to efficiently disrupt ERG-mediated protein-protein interactions, transcription, DNA damage, and cell invasion. Furthermore, a retroinverso peptidomimetic EIP suppressed tumor growth, intravasation, and metastasis *in vivo*. Taken together, our results suggest that transcription factors have specific residues**

**that are important for protein-protein interactions, that may be amenable to therapeutic targeting[1].**

## INTRODUCTION

A majority of prostate cancers harbor recurrent gene rearrangements of the ETS family of transcription factors (*1-5*). The most common of these rearrangements (approximately 50% of prostate cancers), fuse the androgen-regulated *TMPRSS2* promoter to the coding region of the ETS transcription factor *ERG (4, 6)*. This rearrangement results in the pathogenic overexpression of an androgen-responsive *TMPRSS2-ERG* chimeric transcript that encodes a nearly full length ERG transcription factor. The *TMPRSS2-ERG* gene fusion product has been shown to drive a unique transcriptional program as well as induce DNA damage and drive invasion and metastasis (*6-11*). While genetically engineered mice expressing an androgen-regulated *ERG* gene exhibit precursor-like lesions (*9, 12-14*), mice that have both androgen-regulated *ERG* and disrupted *PTEN* rapidly develop invasive prostate cancer (*12, 13*). In some models, combined *AR* and *ERG* overexpression have been shown to drive invasive prostate cancer (*14*). Importantly, siRNA-mediated knockdown of the *TMPRSS2-ERG* fusion gene inhibits the growth of prostate cancer xenografts that endogenously harbor the gene fusion, suggesting its significance as a therapeutic target (*6, 8*). Recently, we

[1] The data for this chapter is in review at *Science* and was completed in collaboration with the following authors: Xiaoju Wang, Bushra Ateeq, Irfan A. Asangani, Qi Cao, Cynthia X. Wang, Rui Wang, Jeanne A. Stuckey, Zaneta Nikolovska-Coleska, Kenneth J. Pienta, and Arul M. Chinnaiyan

have shown that an indirect way to target ERG in prostate cancer is by inhibiting interacting enzymatic proteins such as DNA-PKcs or PARP1 (*11*)

Frontline therapy for advanced prostate cancer often includes androgen ablation which leads to a reduction in *TMPRSS2-ERG* expression in patients harboring the rearrangement, however, androgen independence develops by a variety of mechanisms, many of which restore androgen signaling and *TMPRSS2-ERG* expression (*15*). Although, a variety of newer agents targeting androgen biosynthesis (*16*) or the androgen receptor (*17*) have shown efficacy in castrate resistant prostate cancer, the tumors invariably develop resistance mechanisms. As such, the *TMPRSS2-ERG* gene fusion product is an attractive therapeutic target as it is an oncogenic "driver" transcription factor in both hormone-naïve as well as castrate resistant prostate cancer (*1, 5-11*). The ERG gene fusion in prostate cancer is also an ideal model for the study of therapeutic approaches to target oncogenic transcription factors. Unlike oncogenic kinases (e.g., BCR-ABL in chronic myeloid leukemia), transcription factors have thus far been considered "undruggable" (*18*), and thus, screening modalities to discover inhibitory small molecules or peptidomimetics are much needed.

## MATERIALS AND METHODS

### Protein Expression by Baculovirus and Purification

Baculovirus that could express either ERG or GUS protein was constructed using the Bac-to-Bac Baculovirus Expression System (Invitrogen). Briefly, cDNAs encoding full-length ERG and GUS were cloned into pDEST10 vector

(Invitrogen) and transferred to FastBac vector (Invitrogen), which was then used to make recombinant virus. Virus was then amplified following the manufacturer's protocol and used to infect SF9 cells. After 3 days of infection, infected SF9 cells were harvested and lysed (50 mM Tris [pH 7.9], 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT, 0.1% NP40) in the presence of protease inhibitors (Roche). Cell lysates were then homogenized by three rounds of sonication (30s each) on ice and supernatant was recovered by centrifuging at 11,000 rpm for 10 min. Recombinant his tagged ERG and GUS proteins were purified using $Ni^{2+}$-NTA agarose (Invitrogen) following the manufacturer's protocol and the elutes were dialyzed against lysis buffer overnight at 4°C, and stored at -80°C in 10% glycerol.

**Enrichment of phage display random peptide library**

The Ph.D.-7 Phage Display Peptide Library (New England Biolabs) was used for biopanning experiments. This library contains $10^9$ unique 7-mer peptides that are linked to the N-terminus of a phage coat protein by a 4 amino acid spacer (GGGS). Using this library, ERG-binding phage clones were selected by panning over immunotubes (Maxisorb, Nunc) coated with recombinant protein at 10μg/ml. First, nonspecific clones were removed by pre-absorbing phage clones from the phage library using a GUS protein coated immunotube for 1 h at 4 °C. ERG-binding phage clones were then enriched by panning the pre-cleared phage library over an ERG protein coated immunotube for 2 h at 4 °C, eluted with glycine buffer (pH 2.2) and neutralized with Tris buffer (pH 9.0). DNA from eluted and

neutralized phage clones was then amplified and precipitated for subsequent rounds of selection. A total of 4 rounds of selection were performed.

**Reverse Phage ELISA**

After screening the phage display library, phage clones were serially diluted, spread on to Luria-Bertani (LB) agar plates and grown overnight at 37 °C. A total of 59 single colonies were randomly picked and propagated into 96-well plates. As controls, 4 random clones from the original, un-enriched phage library and one insertless phage clone were also selected. We then performed ELISA assays to confirm the interaction between each phage clone and ERG. Briefly, ELISA plates (Max-Sorb™, NUNC) were incubated overnight at 4°C with either purified his tagged ERG or GUS protein (10 μg/ml in carbonate buffer pH 9.6). Protein coated ELISA plates were then blocked (PBS with 5% BSA) for 1h at RT prior to incubation with expressed phage supernatant (diluted 1:10 in PBS with 1% BSA) for 2h at RT. The plates were then washed three times with PBS–0.1% Tween-20 and bound phage was then detected with horseradish peroxidase-conjugated anti-M13 antibody (GE Healthcare) using the TMB substrate system (Sigma). Detection reactions were quenched using 250mM HCl and absorbance was recorded on a SpectraMAX plate reader (OD 450nm). DNA from phage clones was then sequenced using the 96-gIII sequencing primer according to the manufacture's protocol.

**Peptide synthesis**

Synthetic peptides (including biotin or FITC conjugated peptides) were custom-synthesized and HPLC-purified by AnaSpec (San Jose, CA). Peptides were stored lyophilized at -20°C until reconstituted with sterile water immediately before use for both *in vitro* and *in vivo* experiments. HPLC-MS was used to confirm 95% or higher purity for each peptide.

### *In vitro* transcription/translation system

*In vitro* protein expression was completed by cloning the desired expression cassettes downstream of a Halo protein (approximately 35kDa) to create Halo fusion proteins. Briefly, ERG and ERG sub-domains were cloned into pFN19A vector (Promega) following the manufacturer's instructions. Point mutations were then created in the Halo-ETS pFN19A expression vector using the Quikchange XL site directed mutagenesis kit (Stratagene). After cloning, Halo fusion proteins were expressed using the TNT® SP6 High-Yield Wheat Germ kit (Promega) following the manufacturer's protocol. For each reaction, protein expression was confirmed by incubating the cell-free reactions with Biotin-linked Halo ligand and performing Western blot analysis using an HRP-streptavidin.

### HaloLink Protein Array

Identification of the peptide binding site on ERG was completed by creating a custom protein array of full length ERG and ERG subdomains that could be used to capture ERG-binding phage clones. First, constructs expressing either full length or individual domains of ERG were used as templates to create Halo-fusion

proteins. Protein expression was confirmed by Western blot as described above and expressed proteins were covalently immobilized on HaloLink Slides (Promega) in a humidity chamber for 1 hr at RT. Protein coated slides were then washed using PBS-I (1X PBS, 0.05% IGEPAL CA-630) to reduce any non-specific binding interactions and incubated with ERG-interacting phage clones (diluted in PBS-B: 1X PBS, 10mg/ml BSA) in a humidity chamber for 1hr at RT. After washing, slides were incubated with Cy3 labeled, rabbit anti-M13 (GE Healthcare). In all cases, equal protein loading was confirmed using an anti-HaloTag antibody. For peptide inhibition assays, the protein array was pre-incubated with the synthetic peptides for 1hr prior to incubating with phage clones for an additional 1hr at RT. The signal was detected using an Axon scanner and images were quantified using GenePix Pro 6 software (Molecular Devices).

**ERG:Peptide Binding Assays**

The interaction coefficients between ERG and synthetic peptides or peptidomimetics were determined by biolayer interferometry technology using the Octet Red system (ForteBio). Purified ERG and GUS proteins were biotinylated by EZ-Link NHS-PEG4 Biotinylation Kit (Thermo Scientific) following the manufacturer's protocol and any unincorporated biotin was removed from the reactions with Zeba 2ml desalt columns (Thermo). Biotinylated proteins (500µg/ml) were then incubated with super streptavidin biosensors in binding buffer (20mM HEPES pH 7.4, 150mM NaCl), blocked with 10µg/ml biocytin, and washed three times in binding buffer. Peptides, which were then serially

diluted in binding buffer, mixed with the protein-coupled biosensors and peptidomimetic:protein association/dissociation was monitored for 10 min at 25 °C. Nonspecific binding was controlled by subtracting the signal obtained from peptidomimetic:GUS interactions from that of peptidomimetic:ERG interactions and baseline signal drift was controlled by monitoring immobilized ERG without peptide. Octet Red analysis software was used to analyze the data.

**Protein Expression and Purification**

The ETS DNA binding domain fragment encoding ERG residues 307-395 was cloned into the pMCSG7 expression vector (courtesy of William Clay Brown) by Ligation-Independent Cloning technique (*1*). This vector expresses a fusion protein with a NH2-terminal, and tobacco etch virus (TEV)-cleavable, 6-histidine tag. The ETS DNA binding domain fusion protein was expressed in *Escherichia coli* strain C41. Bacteria cells were subsequently lysed by sonication in buffer containing 50mM Tris, pH 7.5, 500mM NaCl, and 10% glycerol. Supernatant was then purified using Ni-nitrilotriacetic acid (NTA) affinity agarose (Invitrogen) from which protein was eluted with buffer containing 300mM imidazole. To remove the His6 tag, protein elutes were dialyzed overnight with TEV protease at 4°C against 50mM Tris, pH 7.5, 150mM NaCl, 1mM dithiothreitol (DTT), and 10% glycerol. The ETS DNA binding domain was further purified by gel filtration on a Superdex 75 column (GE Healthcare). Finally, proteins (primarily monomers) were eluted in a buffer of 25mM HEPES, pH 7.5, 400mM NaCl, and

0.1% β-Mercaptoethanol. The protein was then concentrated to 4.7 to 7.5 mg/ml for crystallization.

**Crystallization and Structure Solution**

ERG ETS crystals were grown by the hanging-drop vapor diffusion method at 20°C (*2*). Very small (10 x 30 mm), rod shaped crystals of ETS grew after 10 days in 45-60% Tacsimate, pH 7.0 (Hampton Research) and 0.1 M Tris, pH 8.5. Prior to data collection, crystals were cryoprotected in well solution containing 20% glycerol and flash frozen in liquid nitrogen.

Diffraction data for the ERG ETS crystals were collected using the microbeam at the Life Sciences Collaborative Access Team (LS-CAT) beamline (21-ID-F) at the Advanced Photon Source, Argonne IL. The crystals diffracted to 2.6 Å with $C222_1$ symmetry and contained 3 molecules in the asymmetric unit. The structure was solved using Phaser (*3*) in the CCP4 Program suite (*4*) with a search model derived from PDB ID:1K7A. Prior to molecular replacement, the sequence of 1K7A structure was modified to reflect that of ERG ETS. The ETS structure was refined using Buster (*5*) with alternating rounds of model building in Coot (*6*). The final Rwork/Rfree was 18.4/22.8, respectively. Residues 333-414 from ETS-1 which contain the DNA interacting sequences, align with an RSMD of 0.562 Å2. Structure figures were prepared with PyMol.

**ELISA-based transcription factor binding assay**

Free and 5`-biotinylated oligonucleotides were synthesized from Integrated DNA Technologies as follows: ETS binding sequence (EBS), (5`→3`) GATCTTCGAAACGGAAGTTCGAG, p53 binding sequence, (5`→3`) TACAGAACATGTCTAAGCATGCTGGGG, and, PARP binding sequence, (5`→3`) TGCCCCTCCCATTGCCCCTCCCAT. To construct the double stranded (ds) biotin-DNA probes, sense and antisense oligonucleotides were diluted in $H_2O$, incubated in a 95°C heating block for 3min, and slowly cooled to RT. The ds biotin-DNA probe was stored at -20°C.

Before performing protein:DNA binding reactions, 100µL of ds biotin-DNA (2pmol/µl in TBS-T) was incubated with streptavidin-coated plates (Thermo) at 37°C for 1hr. After washing 3X with TBS-T, plates were blocked with 5% non-fat dry milk in TBS-T for 30min at RT and incubated with purified ERG protein (diluted in 20mM HEPES, 150mM NaCl) for 1hr at RT. Plates were then washed three times using PBS-T, incubated with rabbit anti-ERG antibody (Santa Cruz, C-17), washed 3 times and signal was detected by incubating with HRP-anti-rabbit antibody. The absorbance $OD_{450}$ was measured with TMB (Sigma) as above. All experimental data was analyzed using Prism 5.0 software (Graphpad) and the inhibition constants were determined by nonlinear curve fitting to identify the concentration of ERG protein at which 50% of the DNA is bound.

For competitive assays, ERG protein was pre-incubated with free EBS, control oligonucleotide sequences or synthetic peptides/peptidomimetics for 1hr.

The signal was measured as above and the absorbance data was analyzed using a fixed slope dose-response inhibition equation in Prism 5.0 software for determination of $IC_{50}$ values.

**Cell Lines**

PC3 (ATCC) and DU145 (ATCC) prostate cancer cell lines were grown in RPMI 1640 (Invitrogen) and VCaP (ATCC) cells in DMEM with Glutamax (Invitrogen) both supplemented with 10% FBS (Invitrogen) in 5% CO2 cell culture incubator. The immortalized prostate cell line RWPE-1 (ATCC) was grown in Keratinocyte media with L-glutamine (Invitrogen) supplemented with 2.5µg EGF (Invitrogen) and 25mg Bovine Pituitary Extract (Invitrogen). All cultures were also maintained with 50 units/ml of penicillin/streptomycin (Invitrogen). The genetic identity of cell lines was determined as previously described (*7*). Virus and stable isogenic cell lines were made as previously described (*8*).

**Biotinylated peptide pull-down assays**

Approximately 1mg of soluble VCaP cell lysate prepared in RIPA buffer (Invitrogen) was incubated with different amount of biotinylated peptides for 2hrs at 4°C. HaloLink agarose beads (20µl) were then added to the peptide-lysate mixture and incubated for an additional 4hrs. Beads were then pelleted by centrifugation, washed three times in PBS-T, and loaded into an SDS-PAGE gel. Western blot analysis was subsequently used to identify ERG binding. Total VCaP cell lysate was loaded for positive control.

### *In vitro* peptide competitive pull-down assay

Halo-ERG fusion protein was expressed using the *in vitro* transcription/translation system as above. After confirming protein expression, the Halo-ERG fusion protein was incubated with either 100ng of GST-AR protein (Millipore) or 224U (2μl) of purified DNA-PKcs (Promega). Protein mixtures were incubated with or without peptidomimetics as indicated in PBS-T (0.1% Tween) at 4°C overnight. Simultaneously, HaloLink beads (Promega) were blocked in 5% BSA at 4°C overnight. The next day 10μl of HaloLink beads were mixed with the protein-peptide mixture, incubated at RT for 1 hr, washed with 4X in PBS-T and eluted into SDS sample buffer. Eluted proteins were then separated by SDS gel electrophoresis and blotted with either anti-GST mAb (Sigma) or anti-DNA-PKcs (Santa Cruz, H-163) as described below. HaloLink beads without fusion proteins were used as negative controls.

### Western blot analysis

VCaP cells were plated in two wells of a 6-well plate at $0.5x10^6$ cells/mL. After allowing cells to attach, cells were harvested and homogenized in NP40 lysis buffer (Sigma) containing a complete protease inhibitor cocktail (Roche). Ten micrograms of each protein extract were boiled in SDS sample buffer, size fractionated by SDS-PAGE, and transferred onto a PVDF membrane (GE Healthcare). After blocking in 5% nonfat dry milk, membranes were incubated with the following primary antibodies: anti-DNA-PKcs rabbit polyclonal (1:500,

Santa Cruz, H-163), anti-ERG1/2/3 rabbit polyclonal (1:1000, Santa Cruz, C-17) or anti-AR monoclonal (1:1000, Millipore). Following three washes in TBS-T, the blot was incubated with horseradish peroxidase-conjugated secondary antibody and the signals visualized by enhanced chemiluminescence system as described by the manufacturer (GE Healthcare).

**Immunofluorescence staining**

VCaP cells were seeded in a 4-well chamber slide for 24 hours prior to the addition FITC labeled peptide to a final concentration of 10μM. After 2hr incubation, cells were washed two times in PBS, and fixed for 15 minutes in 100% methanol. After washing an additional three times in PBS, cells were mounted using Vectashield with DAPI (Vector laboratories). For foci formation experiments, cells were fixed for 15 minutes in 100% methanol and permeabilized for 5 minutes in PBS containing 0.2% Triton-X100. Slides were then washed twice in PBS, blocked for 30 minutes in PBS containing 0.5% donkey serum (Sigma, St. Louis, MO), held for 45 minutes in PBS containing 0.5% donkey serum and primary antibody: anti-γ-H2A.X mouse monoclonal (Millipore), washed 3 times with PBS, held for 30 minutes in PBS containing 0.5% donkey serum and secondary antibody: Alexa Fluor 488 anti-Mouse (Invitrogen). Images were then taken using 100x oil objective lens on an Olympus Confocal microscope at the University of Michigan microscopy imaging lab.

**Chromatin Immunoprecipitation**

VCaP cells were grown in complete medium and treated with peptidomimetics 48 hours prior to harvesting cells. ChIP assay was carried out as previously described (*8, 9*) using antibodies against ERG (Santa Cruz, #sc-354), DNA-PKcs (BD Biosciences, #610805), rabbit IgG (Santa Cruz, #sc-2027) or mouse IgG (Santa Cruz, #sc-2025). Briefly, chromatin collected from formaldehyde crosslinked and lysed cells was sonicated to an average length of 600bp. Supernatants containing chromatin fragments were pre-cleared using protein A/G beads (Upstate), and incubated with 3-5µg of antibody or IgG overnight prior to the addition of 50µl of protein A/G beads for 1 hr. Immunoprecipitation-bead complexes were then washed six times and antibody/protein/DNA complexes were eluted with 150µl IP elution buffer (50mM NaHCO3, 1% SDS). To reverse the crosslinks, the complexes were incubated in elution buffer containing 10µg RNase A and 0.3 M NaCl at 67°C for 4-5 hours. DNA/protein complexes were then precipitated with ethanol, air-dried, and dissolved in 100µl of TE. Protein was then digested by incubation with proteinase K at 45°C for 1 hr and DNA was purified using a QIAGEN PCR column according to the manufacturer's protocol. QPCR was performed using appropriate primers as previously described (*8*).

**Quantitative Real-Time PCR Assays**

Total RNA was isolated from VCaP cells and tumors using Qiazol and following the manufacturer's protocol (Qiagen). Quantitative PCR (QPCR) was performed

using SYBR Green dye on an Applied Biosystems 7300 Real Time PCR system (Applied Biosystems) as described (*10*). Briefly, 2μg of total RNA was reverse transcribed into cDNA using SuperScript III (Invitrogen) in the presence of random primers (Invitrogen). All reactions were performed in triplicate with SYBR Green Master Mix (Applied Biosystems) plus 25ng of both the forward and reverse primer according to the manufacturer's recommended thermocycling conditions, and then subjected to melt curve analysis. Threshold levels for each experiment were set during the exponential phase of the QPCR reaction using Sequence Detection Software version 1.2.2 (Applied Biosystems). The relative quantity of target gene was calculated for each sample using the ΔΔCt method by the comparing mean Ct of the gene to the average Ct of the housekeeping gene, β-Actin (*11*). All oligonucleotide primers were synthesized by Integrated DNA Technologies (Coralville, IA). The primer sequences for the transcript analyzed were previously described (*8*).

**Immunoprecipitation**

Cell pellets or tumors treated with or without peptidomimetics were lysed in cell lysis buffer (20mM MOPS, pH 7.0, 2mM EGTA, 5mM EDTA, 30mM sodium fluoride, 60mM β-glycerophosphate, 20mM sodium pyrophosphate, 1mM sodium orthovanadate, 1% Triton X-100, 1mM DTT, protease inhibitor cocktail (Roche)). Cell lysates (0.5-1.0mg) were then pre-cleaned with protein A/G agarose beads (Santa Cruz) by incubation for 1 hour with shaking at 4°C followed by centrifugation at 2000 rpm for 3 minute. Lysates were then incubated with 2μg

ERG antibody (Santa Cruz, C-17) at 4°C for 4 hours with shaking prior to addition of 20µL protein A/G agarose beads. The mixture was then incubated at 4°C overnight. After washing 4 times with lysis buffer, beads were finally precipitated and resuspended in 25µL of 2x loading buffer and boiled at 80°C for 10 minutes. Samples were then analyzed by SDS-PAGE Western blot analysis as described above.

**COMET Assays**

COMET assays were performed in a neutral electrophoresis buffer in order to specifically measure the relative level of DNA double strand break fragmentation (*12*) as previously described (*8*). Briefly, peptidomimetic treated cells were trypsinized, harvested by centrifugation and re-suspend in PBS to a concentration of 1 x $10^5$ cells/ml. Suspended cells were then mixed with 250µL 1.0% ultrapure low melting point agarose (Invitrogen) made in 1x Tris-Borate buffer and dropped onto slides. After solidifying in the dark for 20 minutes at 4°C, slides were immersed in COMET assay lysis solution (Trevigen, Gaithersburg, MD) at 4° in the dark for 45 minutes. Excess buffer was then removed and slides were submerged at RT in the dark for 40 minutes in freshly prepared neutral solution (Tris Base 60.57g, Sodium Acetate 204.12g dissolve in 450 ml of $dH_2O$. Adjust to pH = 9.0 with glacial acetic acid.). Slides were then washed twice by immersing in 1 X TBE buffer prior to neutral electrophoresis at 20 volts for the indicated amount of time prior to fixation in 70% ethanol for 5 minutes. Following air drying, slides were stained with SYBR Green dye (Invitrogen) and images were

collected with a 40x objective lens. COMET tail moments were then assessed using COMETscore.v1.5 (AutoCOMET.com, Sumerduck, VA) image processing software as described by the manufacturer with greater than 100 cells analyzed in triplicate experiments. Statistical analysis was completed using the Wilcoxon Rank test.

**Basement Membrane Matrix Invasion Assays**

For invasion assays, the cell lines were treated with peptidomimetics or negative controls for 24hrs prior to harvesting. Twenty-four hours post-treatment, cells were seeded onto the basement membrane matrix (Chemicon, Temecula, CA) in the chamber insert with 8.0μM pores of a 24-well culture plate in serum free media. Cells were attracted to the lower chamber by the addition of complete media as a chemoattractant. After 48 hours incubation at 37˚C with 5% CO2, the non-invading cells and EC matrix were gently removed with a cotton swab. Invasive cells, which were located on the lower side of the membrane, were stained with crystal violet, air dried and photographed. To quantify the relative number of invaded cells, colorimetric assays were performed by treating the inserts with 150μl of 10% acetic acid (v/v) and measuring absorbance of each condition at 560nm using a spectrophotometer (GE Healthcare).

**Chicken CAM assays**

The CAM assay was performed as described previously (*8*). Briefly, fertilized eggs were incubated in a rotary humidified incubator at 38°C for 10 days. After

releasing the CAM by applying mild amount of low pressure to the hole over the air sac and cutting a square 1cm$^2$ window encompassing a second hole near the allantoic vein, cultured VCaP cells that stably express Cherry Red and had been pre-treated with peptidomimetics as indicated were detached by trypsinization and re-suspended in complete medium prior to implantation of 2 x $10^6$ cells adjacent to the mesenchyme in each egg. The windows were subsequently sealed and the eggs were returned to a stationary incubator. For invasion and intravasation experiments, implanted eggs were treated with 25mg/kg body weight of peptidomimetic 6 hours after cell inoculation and both the upper and lower CAM was isolated after 72 hours, respectively. Invasion assays were processed and stained for chicken collagen IV as previously described (*8*). For metastasis and tumor growth assays, eggs were then treated as indicated with 25mg/kg body weight peptide which was administered every other day. At day 8 embryos were sacrificed. The extra-embryonic tumor were excised and weighed. The embryonic lungs were harvested and analyzed for the presence of tumor cells by quantitative human alu-specific PCR. Genomic DNA from lower CAM and lungs were prepared using Puregene DNA purification system (Qiagen). Quantification of human cells in the extracted DNA was done as described (*13*). Fluorogenic TaqMan qPCR probes were applied as above, and DNA copy numbers were quantified.

**VCaP and PC3 xenograft mouse models**

Hsd: Athymic Nude-Foxn1nu mice were procured from Harlan Laboratories, Inc. (Indianapolis, IN). VCaP ($2 \times 10^6$ cells) and PC3 ($5 \times 10^5$ cells) were collected and resuspended in 100µl of saline with 50% Matrigel (BD Biosciences, Becton Drive, NJ) prior to being implanted subcutaneously into the both flank regions of the mice. Mice were anesthetized using a cocktail of xylazine (80-120 mg/kg, IP) and ketamine (10mg/kg, IP) for chemical restraint before tumor implantation. For short term assays, all tumors were staged for three weeks before starting the peptidomimetic treatment. After week 3 for VCaP cells or after 4 days for PC3 cells, mice (10 per treatment group) were treated with peptidomimetics (25mg/kg or 10mg/kg as indicated, IP) daily five times a week. Growth in tumor volume was recorded every 3 days by using digital calipers and tumor volumes were calculated using the formula $(\pi/6)$ $(L \times W2)$, where L = length of tumor and W = width. Loss of body weight during the course of the study was also monitored weekly. At the end of the xenograft studies, mice were anaesthetized and blood was collected by cardiac puncture. Supernatant containing serum was separated by incubating blood on ice for 45 min, followed by centrifugation at 8000 rpm for 10 min at 4°C. All procedures involving mice were approved by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan and conform to their relevant regulatory standards.

**Serum kinetics of peptide concentration**

Following IP injection of peptidomimetics, the amount of biotinylated peptides in the serum of mice was determined by using the Quant-Tag Biotin kit (Vector) and following the manufacturer instructions. Standards with known concentration of biotin were used to generate a curve to calculate the concentration of biotin in the serum samples.

**Peptide distribution in xenograft tumors**

Mice with VCaP xenografts were IP injected with biotinylated cell permeable peptidomimetics. At the indicated time points, tumors were collected and immediately frozen in embedding medium in liquid nitrogen. Tissue sections were cut at -12°C using a cryostat, air-dried overnight, and fixed in 100% acetone at −20 °C for 5 min at RT. Tumor sections were then rehydrated in PBS and incubated with Alexa 594-streptavidin (Invitrogen) for 1hr at room temperature. Slides were mounted with permanent mounting medium Vectashield Hard with DAPI (Vector) to prevent photo bleaching. Pictures were taken using 100x oil lens on an Olympus Confocal microscope at the University of Michigan microscopy imaging lab.

**Statistical Analysis**

All statistical assays were two-tailed t-test unless otherwise indicated. On all plots and graphs the standard error of the mean is shown. * $P < 0.05$.

**RESULTS**

We started by iteratively screening a phage display random peptide library (complexity of $1.28 \times 10^9$) to identify peptides that interacted with the wild-type ERG protein, but not a negative control protein, GUS (**Fig. 5.1A; Fig. 5.2A**). After 4 rounds of enrichment, the bound phage library was eluted and 59 clones were randomly selected to pursue as potential ERG-interacting peptides. DNA sequence analysis revealed that the 59 phage clones encoded 12 unique peptides generated from 18 unique nucleotide sequences (**Table 5.1**). To confirm the specificity of these ERG binding phage clones, we coated ELISA plates with purified ERG or control GUS proteins and monitored binding of phage with anti-M13 antibody (M13 bacteriophage). All 12 unique ERG selected phage clones bound to ERG, while the random phage clones did not (**Fig. 5.1B**). Analysis of the 12 ERG binding phage clones revealed a consensus peptide motif of LSFGSLP (**Fig. 5.1C**). To confirm the enrichment of this consensus peptide motif by phage display, we analyzed randomly selected phage clone DNA sequences from earlier enrichment cycles and identified 5/20 in the second round and 16/20 in the third round of enrichment.

After identifying a consensus ERG-interacting peptide motif, we sought to specifically map the interaction along the ERG protein, which contains 479 amino acids with two conserved domains called pointed (PNT) and the ETS DNA binding domain (*19*) (**Fig. 5.1D**). The pointed domain has been shown to function in hetero- or homo-dimerization (*20*) and mediate transcriptional repression (*21*). Other named domains include the central alternative exons (CAE), central domain

(CD), and C-terminal transactivator domain (CTD); however, their functions are unknown (*20, 22*). To characterize the peptide binding domain, a total of 6 ERG segments, namely N-terminus (N-term), PNT, CAE, CD, ETS and CTD, together with full-length ERG and GUS as controls, were cloned and expressed by *in vitro* transcription/translation as HaloTag fusion proteins (**Fig. 5.2B**). The proteins were then immobilized on HaloLink Protein Arrays and subsequently incubated with individual phage clones. The interaction signal was detected by Cy3-labeled anti-M13 antibody (**Fig. 5.1D**). Interestingly, the phage clone encoding the consensus ERG binding peptide LSFGSLP strongly bound to both full length ERG and the ETS DNA binding domain, but not other domains or control proteins, suggesting that the ETS DNA binding domain is the critical region for interaction.

In order to further pinpoint the critical interactive residues in the ETS DNA binding domain of ERG, a panel of 12 tiling fragments spanning the ETS domain (including two fragments with point mutations) was expressed **(Fig. 5.2C)** and immobilized on HaloLink Arrays. This finer analysis allowed us to specifically localize the interaction site of the ERG binding phage to a 9 amino acid stretch (RALRYYYDK) (**Fig. 5.1E**), corresponding to the residues 367 to 375 within the ETS domain of ERG. Importantly, a single amino acid substitution of R367K completely abolished ERG phage peptide binding, indicating that $R^{367}$ is a critical residue for ERG: phage peptide interactions (**Fig. 5.1E**). This was intriguing, as it has been showed that $R^{367}$ is a critical residue for the ability ERG to bind DNA (*22*). Previously, we had also demonstrated that a Y373A mutation

in ERG disrupts ERG interaction with AR and DNA-PKcs (*11*). Taken together, we would postulate that mutations of amino acids 367 to 375 have a profound effect on the functional activities ERG.

To confirm the functional significance of amino acids 367 to 375, we created two expression constructs encoding ERG mutants within this critical region including R367A/R370A and Y373A (**Fig. 5.1F**). Of note, in *in vitro* cell line studies, both mutants significantly abolished ERG-mediated transcription of known target genes (**Fig. 5.1G**), cell migration (**Fig. 5.1H, Fig. 5.3A**) and DNA damage as assessed by a histone mark of DNA double-strand breaks, γH2A.X foci (**Fig. 5.1I, Fig. 5.3B**) supporting the importance of this region for ERG function.

To study the specificity of the ERG binding phage clones, we employed synthetic peptides for competition experiments. The synthetic 7-mer peptide, LSFGSLP (corresponding to the ERG binding consensus peptide), but not a random peptide, HSKINPT, competitively disrupted the ability of ERG to interact with ERG binding phage clones in a dose-dependent fashion (**Fig. 5.1J**). This suggested that synthetic peptides corresponding to the ERG consensus peptide, rather than phage clones, could be used for *in vitro* studies. To determine the kinetic binding of the synthetic peptides to ERG we utilized biolayer interferometry by Octet RED (**Fig. 5.1K)**. The dissociation constant (Kd) was calculated as 21μM for the LSFGSLP peptide, while a single alanine substitution significantly decreased its binding affinity by 10-fold. There was no specific binding when 3 amino acids were substituted with alanine, emphasizing the

specific and high binding affinity between the synthetic peptide and the ERG transcription factor.

To better understand the physical location of the ERG binding peptide interaction site, we determined the X-ray crystal structure of the ETS domain (aa 307-395) of ERG. The crystals contained 3 molecules in the asymmetric unit (**Fig. 5.4A**). The structure was solved with a search model derived from PDB ID:1K7A (crystal data for the ERG paralogue, ETS1), and found no DNA bound to the protein in the crystal, thus representing the first DNA-independent structure of the DNA binding domain of an ETS transcription factor (**Fig. 5.5A**). Interestingly, the ETS DNA binding domain of ERG aligned extremely closely to the ETS1:DNA structure (using residues 333-414 from 1K7A) with a root mean square deviation (RMSD) of 0.562 $Å^2$ (**Fig. 5.5B**). The structural differences lie at the ends of the termini as our ERG ETS domain structure has 3 additional residues at its N-terminus, but is truncated by 13 residues at its C-terminus when compared to the ETS domain fragment in the ETS-1:DNA complex (**Fig.5.5B, Fig. 5.4B**). The structure reveals a winged helix-turn-helix motif for the ERG ETS domain with a high degree of structural conservation to the ETS1:DNA crystal, which also contains three α-helices and four β-sheets. The main protein-DNA contacts in the ETS1 crystal are from residues that are located in the third α-helix. Interestingly, the ERG:peptide interaction domain, RALRYYYDK, involves part of the third α-helix, where the $R^{367}$ residue perfectly fits into the large groove of DNA double-strand helix structure (**Fig. 5.5B)**. Of note, $Y^{373}$ is adjacent to the α-helix and is accessible to potential interactive proteins

suggesting that this residue may be specifically required for ERG:protein interactions, and in particular ERG interaction with AR and DNA-PKcs (*11*). This led to the hypothesis that the ERG binding peptides may disrupt ERG function by blocking ERG:DNA and/or critical ERG:protein interactions. To test the hypothesis, we first determined the binding affinity between the ERG transcription factor and its consensus DNA-binding sequence (ETS binding sequence, **EBS**) (*22*) using an ELISA-based transcription factor binding assay (*23*). We observed a strong binding for wt-ERG (dissociation constant Kd = 46nM), but not a control ERG deletion protein lacking the ETS domain (ERG-ΔETS) (**Fig. 5.6A**). The binding specificity was further validated by competitive binding assays using unlabeled EBS as competitor oligonucleotides and unrelated oligonucleotides (p53 binding sequence and PARP binding sequence) as negative controls. Interestingly, increasing amounts of free EBS significantly decreased the signal of the ERG transcription factor; in contrast, unrelated competitors, which have no consensus binding sequence for ERG, do not affect the signal strength (**Fig. 5.6B**), confirming the assay specifically measures the binding of ERG to its target DNA binding site EBS. Next, the competitive binding assay was performed with a fixed ERG concentration at 0.2μM in the presence of either synthetic ERG binding peptide LPPYFLT or a derivative control peptide LAAYLFT. For the ERG binding peptide, the absorbance signal decreased dramatically and in a dose dependent fashion resulting in an $IC_{50}$ = 45 +/- 1.1μM, while no significant difference was detected for the derivative control peptide (**Fig. 5.5C**), indicating

that the ERG binding peptide could specifically interrupt the ERG:DNA interaction *in vitro*.

To demonstrate that the ERG binding peptide binds to ERG in cells, we synthesized a biotinylated version. The biotinylated ERG binding peptide LPPYLFT, or a control peptide, was incubated with cell lysate from the *TMPRSS2-ERG* positive prostate cancer cell line VCaP. Following pull down with Streptavidin-linked agarose beads, Western blot analysis showed biotin-LPPYLFT readily enriches for the *TMPRSS2-ERG* gene fusion product in a concentration dependent manner, while mutant control peptide does not (**Fig. 5.5D**). Next, we explored the ability of the ERG binding peptides to disrupt ERG protein-protein interactions in a cell free system. Halo-tagged ERG was incubated with recombinant GST-AR or purified DNA-PKcs in the presence of increasing concentrations of ERG binding peptide LPPYLFT or control peptide. HaloLink magnetic beads were then used to pull down ERG and immunoblot analysis to detect AR or DNA-PKcs. We observed that the direct ERG:AR and ERG:DNA-PKcs protein-protein interactions were disrupted by the ERG binding peptide, but not control peptide, in a dose-dependent fashion (**Fig. 5.5E**).

We next sought to determine whether cell-penetrating synthetic peptides could inhibit ERG transcription factor activities in cells. To address this, we designed cell permeable peptides by conjugating the peptides with the cationic HIV-TAT motif RKKRRQRRR, as a protein transduction domain to facilitate uptake into cells (*24*) (**Fig. 5.5F**). Two cell penetrating ERG binding peptides were created for pre-clinical studies and designated EIP1 (ERG Inhibitory Peptide

1) and EIP2 (**Fig. 5.5F**). EIP2 represented the consensus ERG binding peptide motif (**Fig. 5.1C**) while EIP1 was the most divergent peptide found in the phage display screen that still maintained ERG binding. Because TAT also functions as a nuclear localization signal (*25*), it is well suited for the delivery of transcription factor inhibitors. Immunofluorescence confocal microscopy showed that FITC-labeled EIP1 clearly enters VCaP cells, while EIP1 without the TAT motif did not penetrate the cell line (**Fig. 5.5G**).

Previous studies showed ERG and its transcriptional co-regulators, such as AR and DNA-PKcs, bind to ERG targets, including the *EZH2, PLAT, PLAU* and *ZNF100* promoters and *FKBP5*, *PSA*, and *TMPRSS2* enhancers (*11*). Therefore, we performed chromatin immunoprecipitation (ChIP) assays in VCaP cells to assess whether peptide treatment could block endogenous ERG transcriptional activity. These experiments clearly demonstrated that EIP1/2, but not control peptide, inhibits ERG and DNA-PKcs enrichment at ERG-regulated genomic loci (**Fig. 5.5H**). Importantly, while AR is an ERG binding protein it is also a transcription factor that binds AR-responsive elements, some of which are directly adjacent to ETS binding sites (*26*). As EIP1/2 did not affect AR recruitment, the data supports the hypothesis that the presence of the peptide specifically blocks the recruitment of ERG and its non-transcription factor co-factors such as DNA-PKcs to ERG-regulated genomic loci. As EIP1/2 blockage disrupted the ERG-mediated recruitment of DNA-PKcs, we then determined the functional consequences of the EIP1/2 by testing the relative expression of several ERG target genes, including *EZH2*, *PLAT*, *PLAU* and *ZNF100* (*6, 11, 26*), after

the treatment with EIP1/2. EIP1/2, but not a control peptide, caused the repression of ERG-mediated transcription in VCaP cells (**Fig. 5.5I**). Finally, we determined whether EIP disrupts endogenous ERG protein-protein interactions in VCaP cells. In concordance with the cell free system (**Fig. 5.5E**), IP-Western blot analysis showed EIP1/2 significantly blocks the ERG:AR and ERG:DNAPKcs interactions in cells, whereas the control peptide had no effect on these interactions (**Fig. 5.5J**).

Recently, we have showed that constitutive overexpression of ERG leads to increased DNA damage at baseline in prostate cancer cells (*11*). Thus, we sought to test the hypothesis that our EIPs might alter the level of DNA damage in cells. Incubation of 10μM EIP1/2, but not control peptide, in VCaP, PC3-ERG, or RWPE-ERG cells caused a significant decrease in the average number of γ-H2A.X foci (**Fig. 5.7A, Fig. 5.8A**). In contrast, the EIPs have no effect on DNA damage caused by etoposide, a topoisomerase II inhibitor, in either RWPE-LacZ and 22RV1 cells suggesting that the peptides are not generally enhancing DNA repair. cMYC is another transcription factor commonly dysregulated in cancer that has been shown to induce DNA damage at baseline (*27*). Thus, we created constitutive RWPE-cMYC cells which had a higher baseline level of γH2A.X foci than RWPE-control cells. Importantly, the EIPs were unable to block cMYC-induced DNA damage (**Fig. 5.7A**). The presence of DNA double-strand breaks was further confirmed by directly analyzing cellular DNA for fragmentation using the neutral COMET assay (**Fig. 5.7B, Fig. 5.8B**). As with the γ-H2A.X foci formation assays, in PC3 cells, ERG or c-Myc overexpression was sufficient to induce significantly longer and brighter tails indicative of DNA damage than

those observed in control cell lines, and treatment with 10μM EIP1/2 led to a reduction in relative level of DNA double-strand breaks caused by ERG overexpression. Likewise, 10μM EIP1/2 blocked DNA damage in the endogenous *TMPRSS2-ERG* cell line model, VCaP (**Fig. 5.7B**).

Next, we sought to determine if the EIPs could prevent tumor cell transwell invasion through matrigel in boyden chamber assays. We previously developed two RWPE stable cell lines overexpressing either ERG or EZH2 that are highly invasive as compared to parental RWPE or RWPE-LacZ cells (*6, 28*). Importantly, addition of 10μM EIP1/2, but not control peptide significantly attenuated invasion of RWPE-ERG, but not RWPE-EZH2 cells (**Fig. 5.7C**). Likewise, EIP1/2 also inhibits VCaP cell invasion to a comparable extent as ERG-siRNA (**Fig. 5.7C**) while having no effect on ETS gene fusion negative prostate cancer cells, DU145 and PC3 (**Fig. 5.7D**).

To explore the effect of the EIPs using *in vivo* models, we employed multiple xenograft systems; however, we observed rapid degradation of the EIPs with no overall efficacy in these models (data not shown). In contrast to L-isomer peptides, cell-penetrating peptides in the retroinverso configuration (D-isomers) have been shown to enhance the biologic activity of peptides in xenograft models as these peptides retain a similar structural configuration to their L-isomers, but are resistant to cleavage by serum or cellular proteases (*29, 30*). Thus, we sought to explore whether retroinverso peptidomimetics (**Fig. 5.9A**) still retain the specific anti-ERG properties of the parental L-amino acid peptide. Competitive transcription factor binding assays revealed stronger inhibitory activities for

retroinverso-EIP1 (RI-EIP1) with $IC_{50}$=13 +/- 1.2μM as compared to EIP1 which had an $IC_{50}$=45 +/- 1.1μM (**Fig. 5.9B**). To determine whether RI-EIP1 still disrupts the ERG interactome, we performed immunoprecipitation in the presence of retroinverso peptides. At a concentration of 50μM, RI-EIP1, but not a RI-control peptide, attenuated ERG:AR interaction both in an *in vitro* transcription/translation system and endogenously in VCaP cells (**Fig. 5.10A** and **5.10B**), indicating that this peptidomimetic retained the ability to bind to ERG and disrupt its interactome. Likewise, treatment of VCaP cells with RI-EIP1 repressed the ERG-mediated gene expression to a similar extent as the L-isoform peptides (**Fig. 5.10C**). We previously showed that L-isoform peptides specifically block the biologic activities of ERG, thus inhibiting DNA damage as well as cell invasion and migration. Likewise, RI-EIP1 also specifically blocks DNA damage in VCaP cells in either γH2AX foci or neutral COMET tail moment assays (**Fig. 5.10D** and **5.10E**). Interestingly, we performed this experiment over a time course to determine the length of effect in cell culture and even at 72 hours, RI-EIP1 was still able to inhibit DNA damage, indicating the longer stability of retroinverso peptidomimetics. Further, like its L-isoform counterpart, RI-EIP1 specifically blocks ERG-, but not EZH2-driven, cell migration in transfected RWPE cells at 5μM (**Fig. 5.10F**) as well as inhibits VCaP, but not PC3 or DU145 cell migration (**Fig. 5.10G**). Thus, at less than half the concentration of the L-isomer, the retroinverso peptides achieved comparable inhibitory effects (**Fig. 5.10H**). Collectively, these data suggest that the retroinverso peptidomimetics of EIP may have superior inhibitory performance than the L-isoform peptides.

We next sought to define the role of RI-EIP1 in ERG-mediated invasion and metastasis *in vivo* using established cancer models. To ascertain this, we implanted *TMPRSS2:ERG* positive VCaP cells onto the upper chorioallantoic membrane (CAM) of a fertilized chicken embryo and analyzed the relative number of cells that invade and intravasate into the vasculature of the lower CAM after treatment with the peptidomimetics (*31*). Interestingly, RI-EIP1 dramatically blocked both ERG-mediated invasion and intravasation ($P = 0.05$) (**Fig. 5.9C** and **5.9D**). To extend our observation that the peptidomimetic blocks VCaP cell line metastasis, we harvested lungs and analyzed the metastatic potential of VCaP cells in this model. As shown in **Figure 5.9E**, RI-EIP1 treatment blocked the formation of lung metastases from VCaP tumors ($P = 0.03$), suggesting therapeutic inhibition of ERG biologic activities using the RI-EIPs could effectively block tumor invasion as well as metastatic spread.

To assess the basic pharmacokinetic properties of RI-EIP1, we first sought to test whether the peptidomimetic penetrates xenografted tumor cells and is stable in serum after intraperitoneal (IP) administration. For this purpose, VCaP xenograft mice were injected with a single dose of 25 mg/kg biotinylated RI-EIP1 and serum and tumors were collected at different times as indicated. ELISA assays showed that serum peptide levels reached a peak at 30 minutes after IP injection (**Fig. 5.11A**), and were still detectable in serum 24 hours after administration (as noted earlier, this was not the case with the L-isomer peptides). Immunofluorescence analysis demonstrated that VCaP tumors had a high level of

peptide in the tumor 3 hours after injection that decreased over time, but was still detectable 24 hours after administration (**Fig. 5.11B**).

In order to evaluate the anti-tumor activity of RI-EIP1, we implanted VCaP (ERG fusion positive) or PC3 (ERG fusion negative) cells in nude mice and studied the impact of RI-EIP1 on xenograft growth. Once palpable tumors were detected, pairs of mice were randomized to receive either 50mg/kg per day (n=10) or 10mg/kg (n =10) of the retroinverso peptidomimetics. At day 6, however, the group with high dose of RI-EIP1 had a 10% body weight loss, indicating that 50mg/kg peptide was beyond the maximum tolerated dosage. We therefore decreased the maximum dosage to 25mg/kg for rest of the study. Importantly, we observed a significant reduction of VCaP tumor growth for RI-EIP1 treatment relative to that of control peptide at day 18 ($P$<0.01) (**Fig. 5.11C**), suggesting specific targeting of ERG-positive tumors. Interestingly, after categorizing the tumor sizes into four groups to analyze the tumor size variance (<100mm$^3$; 100-500 mm$^3$; 500-1000 mm$^3$ and >1000 mm$^3$), we found a remarkable difference in the tumor size distribution (**Fig. 5.11D**) as the majority of tumors from RI-EIP1 treatment were <500 mm$^3$, while most control groups exhibited tumor size > 500 mm$^3$. Particularly, among the high dosage group with RI-EIP1, 10% (2/20) of the tumors were completely regressed with no palpable tumor at day 18. Because we noticed that over the extended treatment period, the VCaP tumors were significantly reduced by RI-EIP1 monotherapy, we continued to treat the high dosage group with RI-EIP1. At day 30 of treatment, we found that 35% (7/20) of the tumors did not recur. These mice were then maintained for

an additional 30 days without therapy, and no recurrences were observed. Importantly, the effect of RI-EIP1 was specific to VCaP cell line xenografts as there was no difference in tumor size for PC3 cell line xenografts treated with the RI-EIP1 as compared to the control peptidomimetic (**Fig. 5.11E**). Doses of the peptidomimetic at or below 25mg/kg did not have a significant effect on total body weight (**Fig. 5.12A** and **5.12B**). This data demonstrates that RI-EIP1 can markedly attenuate the growth of ERG-positive prostate cancer cells that have become addicted to ERG overexpression *in vivo*.

To assess the basic pharmacodynamics properties of RI-EIP1, we determined whether the peptidomimetics can block the ERG interactome *in vivo*. We performed IP-Western blot analysis using tumor lysate harvested from VCaP cell xenografts treated with RI-EIP1 or a control peptidomimetic. Remarkably, RI-EIP1, but not the control, attenuated the ERG:AR and ERG:DNA-PKcs interactions in the VCaP xenograft model (**Fig. 5.11F**). Furthermore, qPCR analysis of ERG-target genes in 26 tumors demonstrated that RI-EIP1 inhibited gene expression of ERG target genes *EZH2*, *PLAT* and *ZNF100*, but not *TMPRSS2*:*ERG* gene itself in the VCaP xenograft model (**Fig. 5.11G**) indicating that *in vivo*, peptidomimetic doses can be achieved that disrupt ERG protein-protein interactions thereby blocking ERG transcription factor activity.

## DISCUSSION

Taken together, our studies support the notion that ETS gene fusions may represent an attractive therapeutic target in prostate cancer. ETS genes also

function as “driver” oncogenic transcription factors in almost all Ewing’s Sarcoma (*32, 33*) as well as certain lymphoid leukemias (*34, 35*). Many other transcription factors, most notably MYC, have been implicated as an oncogenic driver in a variety of cancers (*27, 36*), however, direct targeting of MYC and transcription factors in general, has represented a formidable challenge in the field. The development of ERG inhibitory peptidomimetics demonstrates that tumors addicted to transcription factor signaling are susceptible to disruption of transcription factor-protein interactions. Specifically, approaches to disrupt the ERG interactome may have therapeutic benefit for ERG gene fusion positive prostate cancer. While the peptidomimetic strategy to disrupt ERG and transcription factors in general, may have clinical potential in the future, much work remains to actually develop these into drugs. As a probe compound however, the characterization of RI-EIP allows for the in depth study of ERG function in *in vitro* and *in vivo* model systems, and may serve as a valuable tool to screen for more traditional small molecules capable of disrupting the ERG-EIP interaction.

**FIGURES**

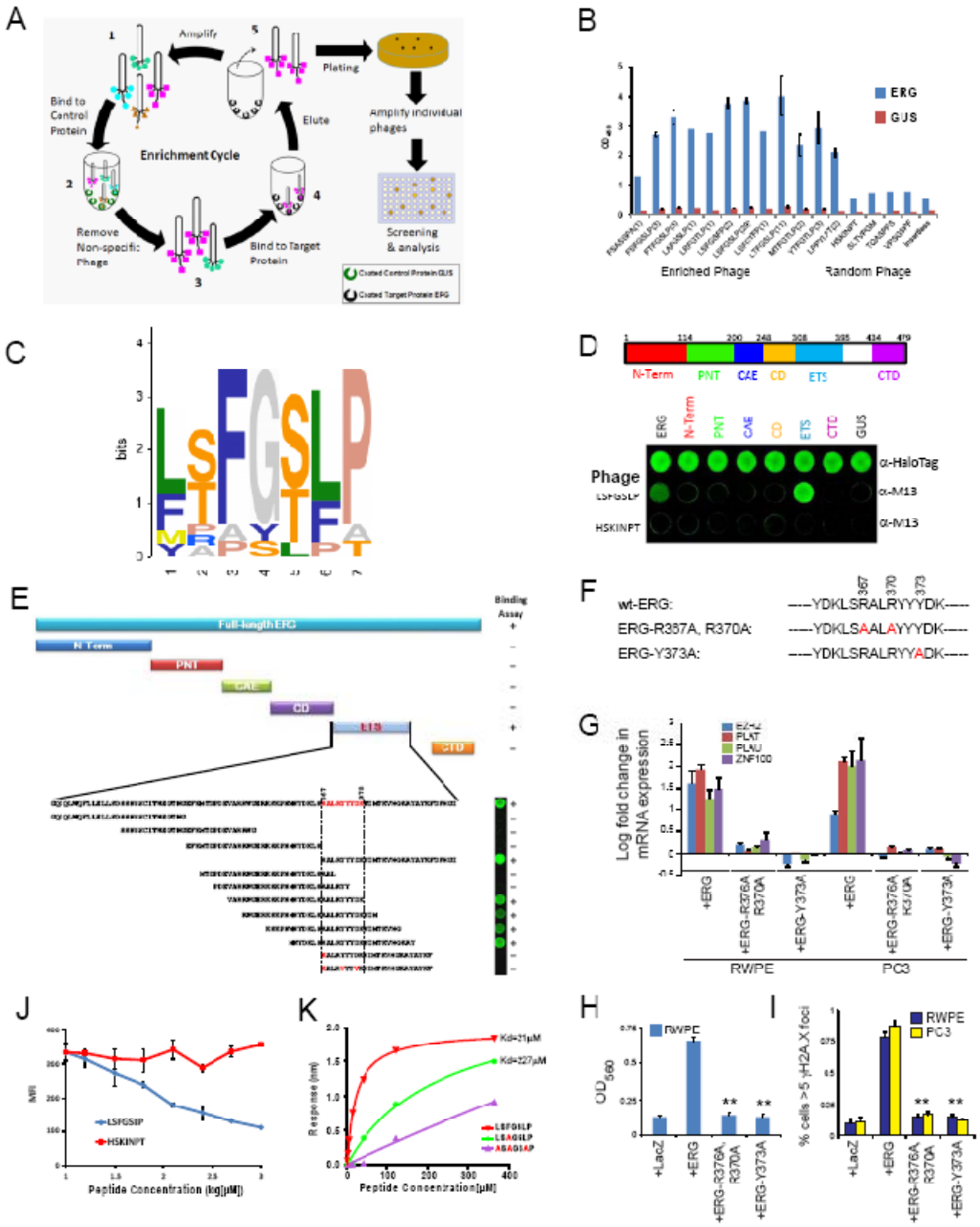



**Figure 5.1: Identification and characterization of ERG inhibitory peptides.** (**A**) Schematic representation of the phage display approach to identify ERG binding peptides. A random 7 amino acid phage display library (1), was pre-adsorbed onto purified GUS control protein to remove non-specific peptides (2). The pre-cleared phage peptides (3) were then enriched for ERG-binding peptides

by employing purified recombinant ERG protein as bait (4). The bound phage clones were then eluted and propagated. After 4 rounds of selection, enriched ERG-binding peptide phage clones were cultured on LB agar plates, and randomly picked for further screening and analysis (5). (**B**) Validation of selected ERG binding phage peptides by ELISA. ELISA plates were coated with either recombinant ERG or GUS protein. Phage clones on x-axis are represented by encoded peptide sequence. The numbers in parentheses indicate clone repeats for each peptide sequence. (**C**) Consensus sequence for aligned ERG binding phage peptides. The height of each residue is proportional to its frequency. (**D**) Protein array analysis to localize the interaction domain of ERG binding phage peptides. Respective Halo-tagged recombinant proteins (either full length or specific ERG domains as indicated) are immobilized on a protein array surface as a matrix with each column representing a different ERG related protein. The first row indicates the relatively equal amounts of recombinant proteins spotted as detected with anti-halo-tag antibody. The second and third rows represent ERG binding peptide (LSFGSLP) phage or control peptide (HSKINPT) phage incubated with each recombinant protein and detected with anti-M13 antibody. N-term, N-terminal domain (1-113); PNT, pointed domain (114-200); CAE, central alternative exons (201-248); CD, central domain (249-308); ETS, ETS DNA binding domain (309-395); CTD, C-terminal trans-activator domain (434-479). (**E**) Refinement of precise residues in ERG that the ERG binding phage peptide interacts. Protein array analysis using recombinant versions of ERG domains as in **D**. Summary of data from **D** and further mutational analysis of the ETS domain in lower expansion panel. Positive interactions are labeled as "+" and negative as "-". Numbers indicate amino acid residue. Inset is a representative array image. (**F**) Amino acid sequences of wild-type ERG protein and site-directed mutants (mutated residues in ETS domain are indicated). (**G**) qPCR analysis of ERG target genes isolated from RWPE or PC3 stable cell lines transduced with lentivirus. (**H**) Cell invasion assay of RWPE cells expressing *TMPRSS2:ERG* or *TMPRSS2:ERG* mutants. (**I**) DNA damage induction as assessed with γH2A.X immunofluorescence staining of RWPE or PC3 stable cells expressing *TMPRSS2:ERG* or *TMPRSS2:ERG* mutants. (**J**) Competitive displacement of phage LSFGSLP binding of ERG with synthetic peptide LSFGSLP. Protein array analysis as in **D** and median fluorescence intensity (MFI) is shown on the y-axis. (**K**) Steady-state analysis of real-time binding data between ERG and synthetic peptide LSFGSLP by an Octet RED biolayer interferometry. The plots show the response versus peptide concentration curves derived from the raw binding data. Kd represents the peptide concentration yielding half-maximal binding to ERG. For all experiments mean ± SEM is shown (**$P < 0.01$).

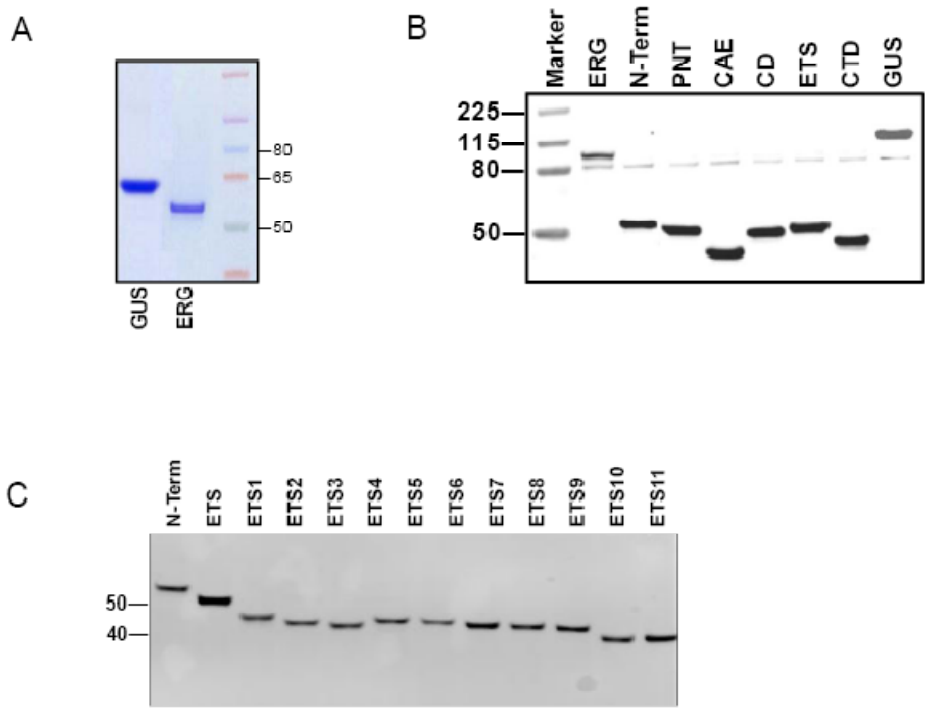



**Figure 5.2: Expression of recombinant proteins.**
(**A**) Coomassie staining of recombinant his-ERG and his-GUS proteins expressed by Baculovirus in insect cells. (**B**) HaloTag-fusion ERG proteins and sub-domains were expressed using SP6 High-Yield Wheat Germ lysate analyzed by western blot. (**C**) Western blot demonstrated the expression level of ETS domain and its segments as Halo-fusion proteins.

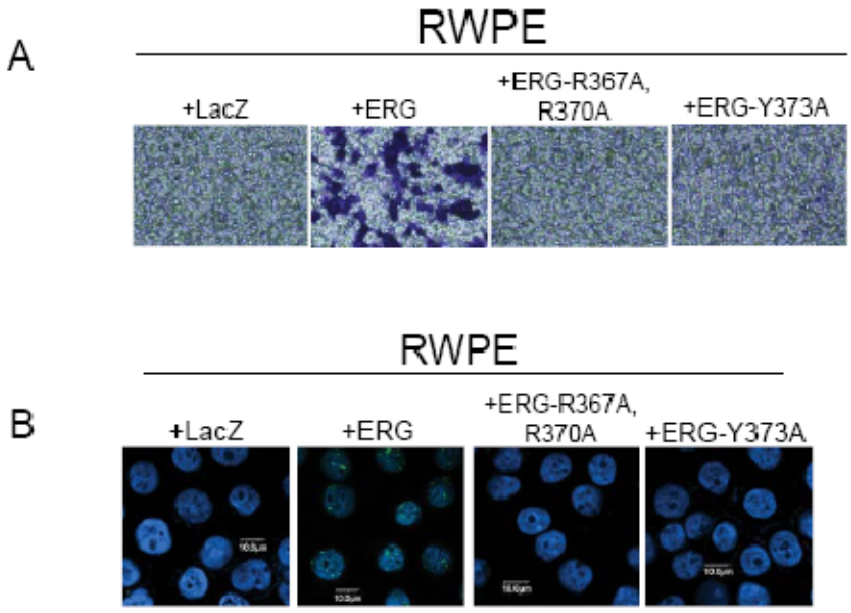



**Figure 5.3: Point mutations in ETS domain abolished *TMPRSS2:ERG*-induced invasion and DNA damage.** (**A**) Infected RWPE cells as indicated were assayed for invasion through a modified basement membrane. Photomicrographs of invaded cells are shown. (**B**) γ-H2A.X immunofluorescence staining shows that mutant ERG attenuates the formation of γ-H2A.X foci driven by *TMPRSS2:ERG*.

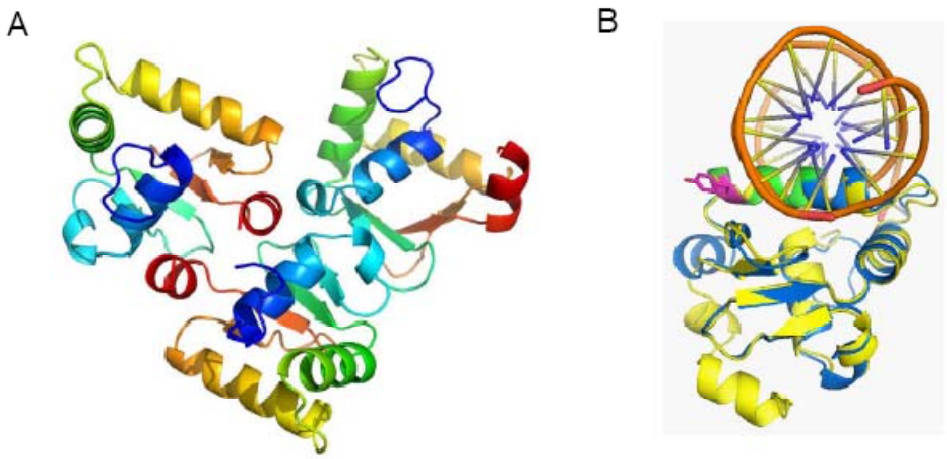



**Figure 5.4: Analysis of ERG ETS DNA binding domain structure.** (**A**) X-ray crystal structure of ERG ETS DNA binding domain shows three molecules in the asymmetric unit. (**B**) Structural comparison of ERG ETS domain and Ets-1:DNA complex. $Y^{373}$ (highlight as pink color) is adjacent to α-helix formed by ERG:peptide interactive domain.

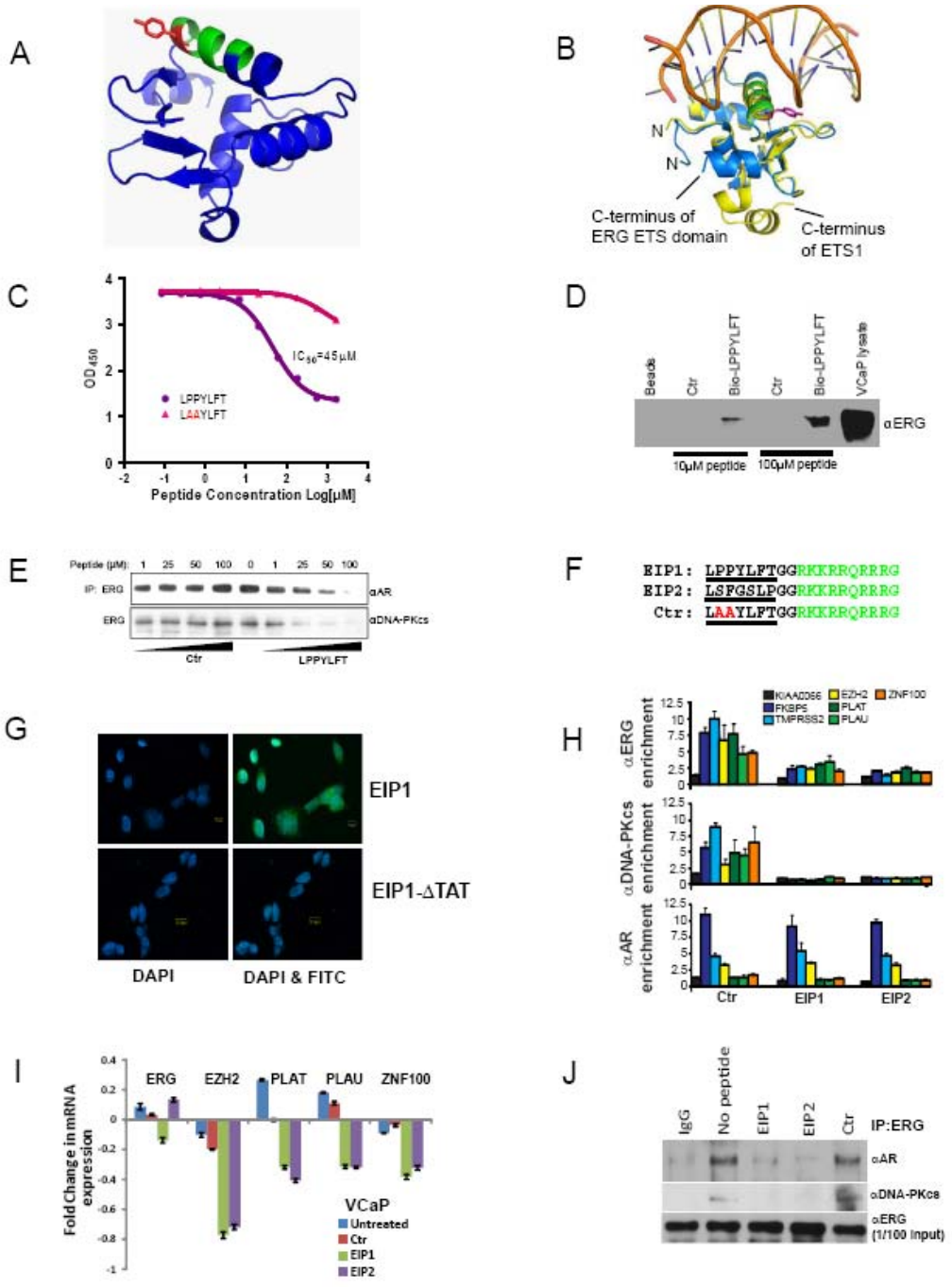



**Figure 5.5: ERG binding peptides specifically inhibit ERG activities**. (**A**) Crystal structure of the ERG ETS DNA binding domain. Recombinant human ERG spanning amino acids 307-395 of the ETS domain was crystallized and the structure was determined. The region of the ERG ETS domain shown to interact with ERG binding peptides is highlighted in green (residues RALRYYYDK), while tyrosine 373 is highlighted in red. (**B**) The ETS DNA binding domain of ETS1 bound to DNA (known structure) overlaid with the ERG ETS domain as determined from this study. The ERG ETS domain structure is shown in blue, the

ETS1 ETS domain bound to DNA (PDB ID: 1K7A) is shown in yellow with DNA shown in orange. The N- and C-termini are labeled. The ERG binding peptide interaction residues of ERG, as well as tyrosine 373, are highlighted as in **A**. (C) Competitive transcription factor binding assay demonstrates that ERG binding peptide LPPYLFT can disrupt the interaction of ERG with DNA. (D) Interaction of ERG binding peptide with ERG expressed in cells. Biotinylated-LPPYLFT or Biotinylated-Ctr (control) peptide were incubated with VCaP cell lysates. Elutes from the pull-downs were then Western blotted and probed with an anti-ERG antibody. (E) ERG binding peptide LPPYLFT disrupts ERG interaction with AR and DNA-PKcs. Halo-tagged ERG protein was expressed using an *in vitro* transcription/translation system. Pull-down Western blot assays were then performed on mixtures of either Halo-ERG and GST-AR or Halo-ERG and DNA-PKcs in the presence of various concentrations of peptide as indicated. (F) Sequence of ERG inhibitory peptides (EIPs) and control used for cell line studies. Permeable peptide sequences with TAT sequence highlighted as green. Consensus sequence with two alanine substitutions in red serves as the negative control. (G) EIPs are cell permeable. VCaP cells were treated with FITC labeled EIP1 or EIP1 without TAT sequence (ΔTAT). Cells were mounted and counterstained with DAPI before imaging with a 40x objective. (H) EIPs block ERG recruitment to target loci. ChIP assays performed using ERG, DNA-PKcs and AR antibodies on VCaP cells treated with peptides for 48hrs prior to crosslinking. QPCRs were performed for several ERG-regulated genomic loci including promoters and enhancers. *KIAA0066* serves as negative control gene. (I) EIPs block ERG-mediated transcription. QPCR analysis of several ERG-regulated genes in VCaP cells treated with EIPs. The y-axis represents the normalized mRNA expression level using β-actin as internal reference. (J) EIPs block the ERG interactions with AR and DNA-PKcs in VCaP cells. IP-Western blot analysis of VCaP cells treated as in (I) with either EIPs or Ctr peptides. Cell lysate (1/100 IP input) are loaded as positive control.

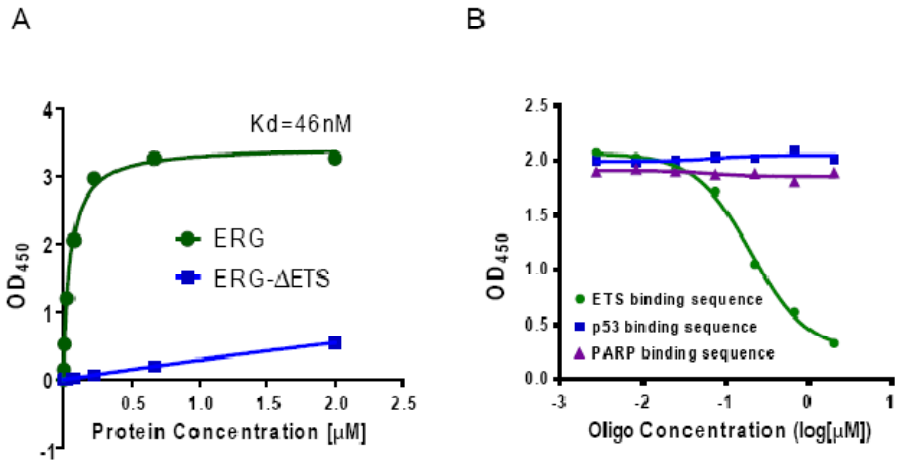



**Figure 5.6: ELISA-based transcription factor binding assay.** (**A**) Wide type ERG, but not a deletion mutant lacking the ETS DNA binding domain, specifically binds to ETS binding sequence (EBS) with high affinity. Microtiter plates coated with biotinylated EBS were incubated with increasing concentrations of proteins. The bound proteins were determined by a polyclonal antibody against ERG followed by a peroxidase-conjugated secondary antibody as described in detail in Materials and Methods. The apparent dissociation constants (Kd) represents the ERG concentration yielding half-maximal binding. (**B**) Competitive DNA-binding assays were performed in the presence of free EBS as competitor oligonucleotides and p53 or PARP1 binding sequences as non-competitor oligonucleotides using a fixed concentration of ERG protein at 0.2μM. Absorbance signal decrease only in the presence of the EBS oligonucleotide confirms that the assay specifically measures binding of ERG to its target site.

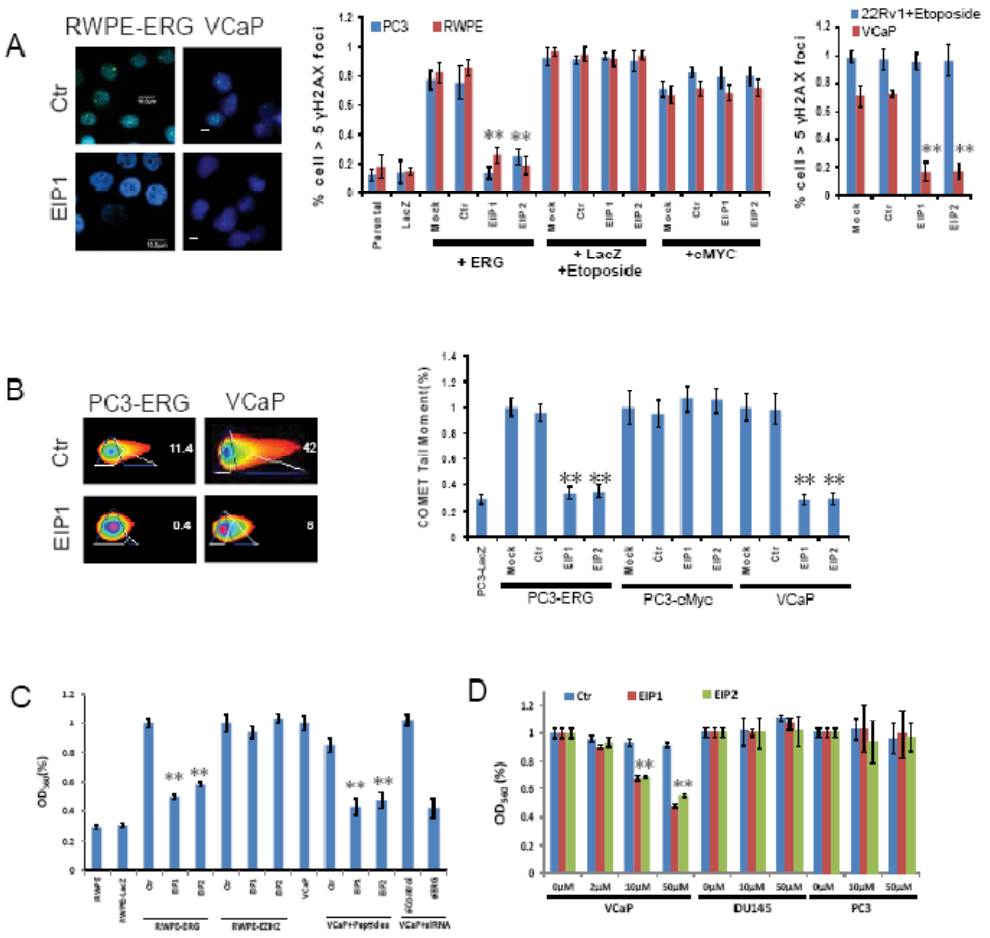



**Figure 5.7: EIP blocks ERG-mediated DNA damage and cell invasion.** (A) Representative γ-H2A.X immunofluorescence staining shows that EIP1 specifically inhibits the formation of γ-H2A.X foci in RWPE-ERG and VCaP cells (images on left). Quantification of γ-H2A.X immunofluorescence staining is shown in the bar graphs. (B) Single cell COMET assays performed in a neutral buffer to assess relative levels of DNA double strand breaks. Representative images show quantification of head and tail height, length, and fluorescence intensity (images on left) with calculated COMET tail moments shown in white. Quantification of average COMET tail moments following treatment as indicated is shown in the box plot on the right. (C-D) Boyden chamber transwell invasion assays were performed on chambers pre-coated with Matrigel. Cells were pre-treated as indicated for 48 hours and then allowed 48 hours to invade before fixation, imaging and quantification. For all experiments mean ± SEM is shown (**$P < 0.01$).

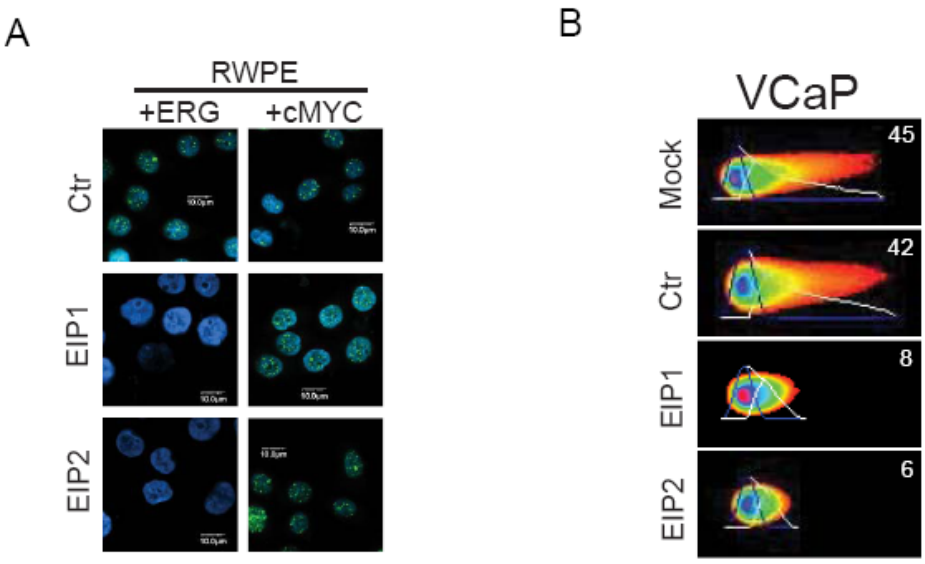



**Figure 5.8: EIPs block ERG-driven DNA damage.** (**A**) γ-H2A.X immunofluorescence staining (Alexafluor 488, counterstained with DAPI) of stable RWPE-*TMPRSS2:ERG or -cMYC* cells treated with EIPs as indicated. Representative photomicrographs were taken using an Olympus confocal with 100x objective lens. (**B**) EIPs attenuate DNA damages as assessed by neutral COMET assay in VCaP cells. Cells with DNA damage have an extended "tail moment" of fragmented DNA shown in white. Representative images showing quantification of head and tail height, length, and fluorescence intensity using automated software are shown.

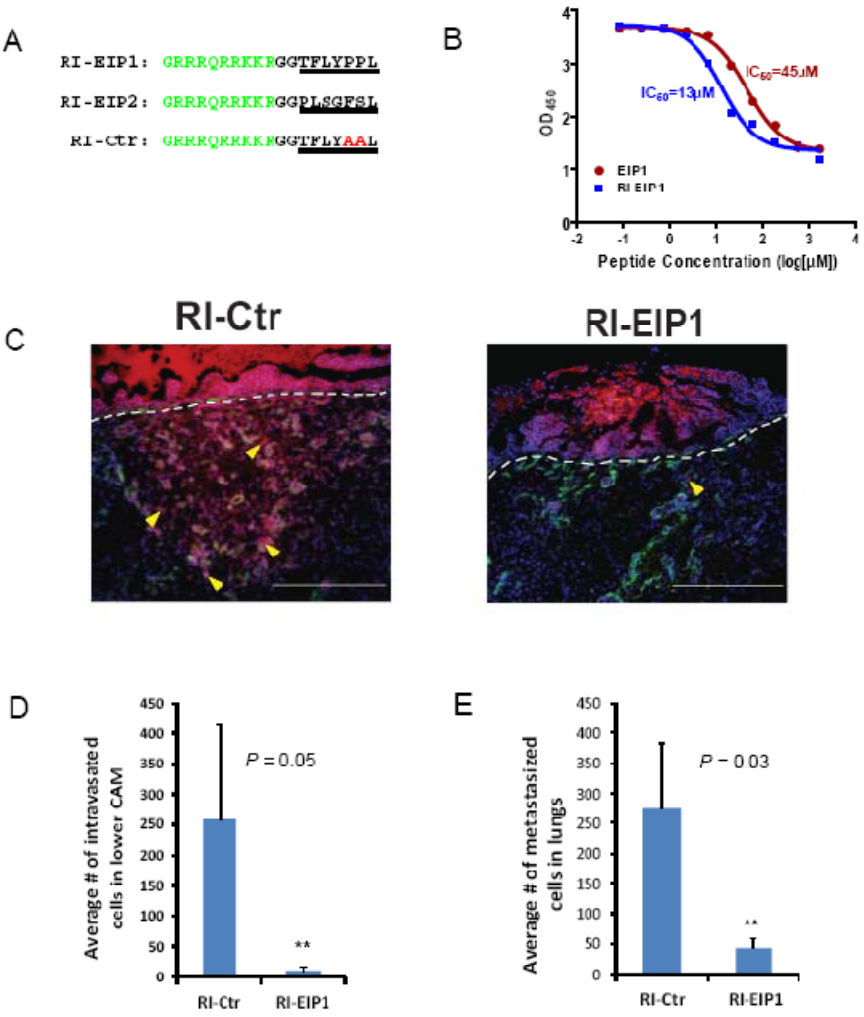



**Figure 5.9: Retroinverso EIPs block ERG-mediated tumor invasion and metastasis *in vivo*.** (A) Sequences and names of retroinverso peptidomimetics. Except glycine, all amino acids are D-isomer. TAT sequences are in green. (B) Competitive transcription factor binding assay shows superior binding ability of the retroinverso peptidomimetic than its parental L-peptide (also see Figure 2C). (C) Chicken chorioallantoic membrane (CAM) invasion assays were performed using stable VCaP cells that overexpress Cherry Red (red fluorescence emission). The xenografts were treated with a single dose of peptide (25mg/kg) as indicated. Frozen sections of upper CAM were stained for chicken-specific type IV collagen (green fluorescence). Yellow arrowheads indicate cells invaded through the upper CAM. White dotted line indicates CAM basement membrane. Scale bars, 200 μM. (D-E) CAM intravasation and lung metastasis assays were performed on VCaP CAM xenografts. qPCR was performed on total DNA isolated from lower CAM and chicken lung using human-specific Alu-PCR primers. Total cell number was determined by comparing to a standard curve created using varying amounts of VCaP cells as input.

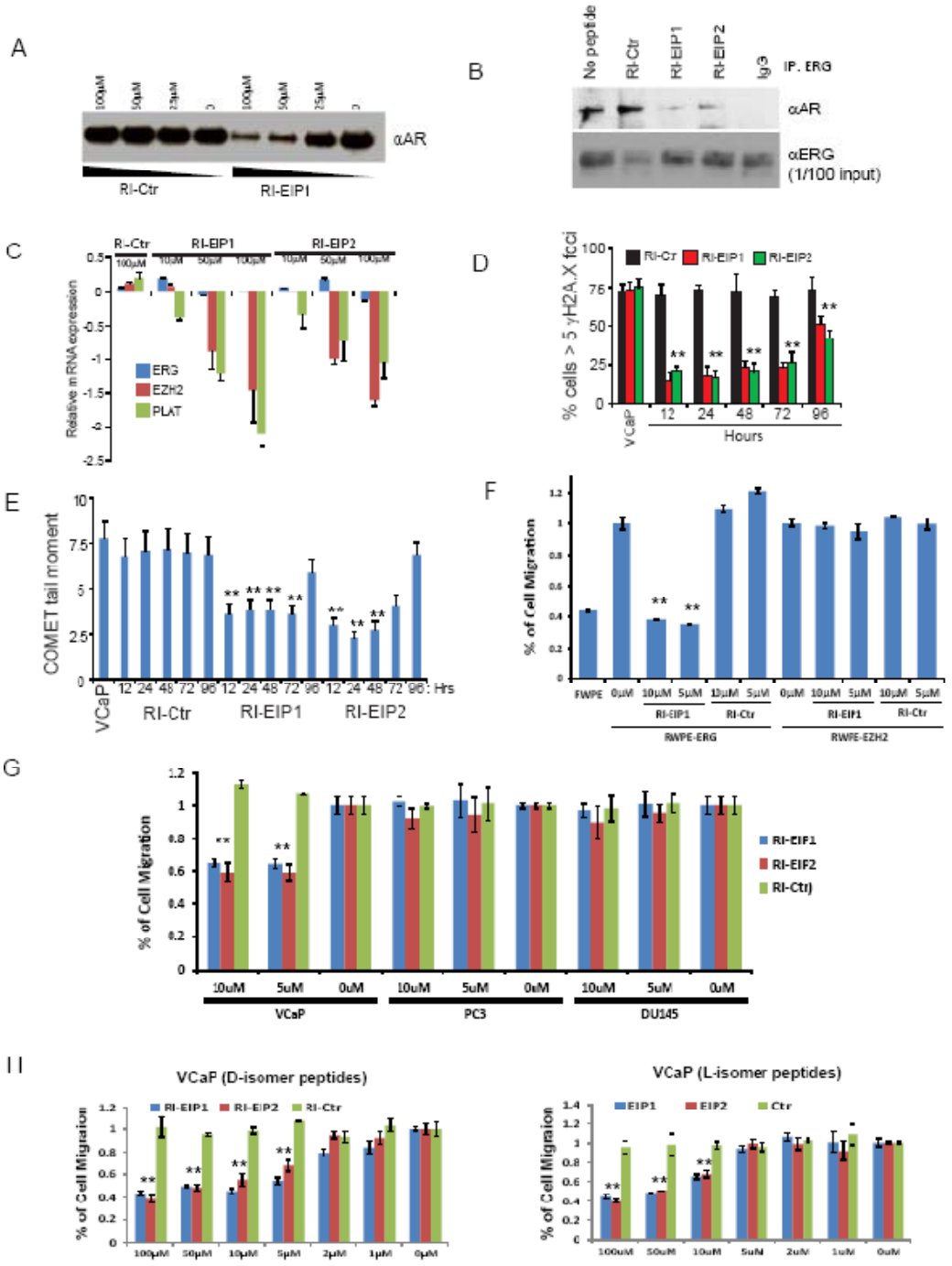



**Figure 5.10: RI-peptides retain the specific anti-ERG properties.** (A) RI-EIP disrupts the ERG-AR interaction *in vitro* as in Figure 2E. (B) RI-EIPs block ERG interaction with AR in VCaP cells as in Figure 2J. (C) RI-EIPs block ERG mediated transcription as in Figure 2I. (D) RI-EIP blocks DNA damage in VCaP cells by either γH2AX foci formation assays, or, **(E)** neutral COMET tail moment assays. Cells were treated with peptide for various amounts of time as indicated. Histograms denote average number of foci per cell or average COMET tail

moment, respectively. (**F-G**) Boyden chamber transwell migration assays showed RI-EIPs specifically block ERG-, but not EZH2-driven, cell migration in transfected RWPE cells at 5μM as well as inhibits VCaP (*TMPRSS2:ERG*), but not PC3 or DU145 (both ETS negative) cell migration. (H) Comparison of cell migration in VCaP cells treated with RI-peptides and its parental peptides at different doses as indicated. Approximately half as much RI-EIP is required to achieve comparable inhibitory effects as the corresponding L-isomer peptide. For all experiments mean ± SEM is shown (**$P < 0.01$).

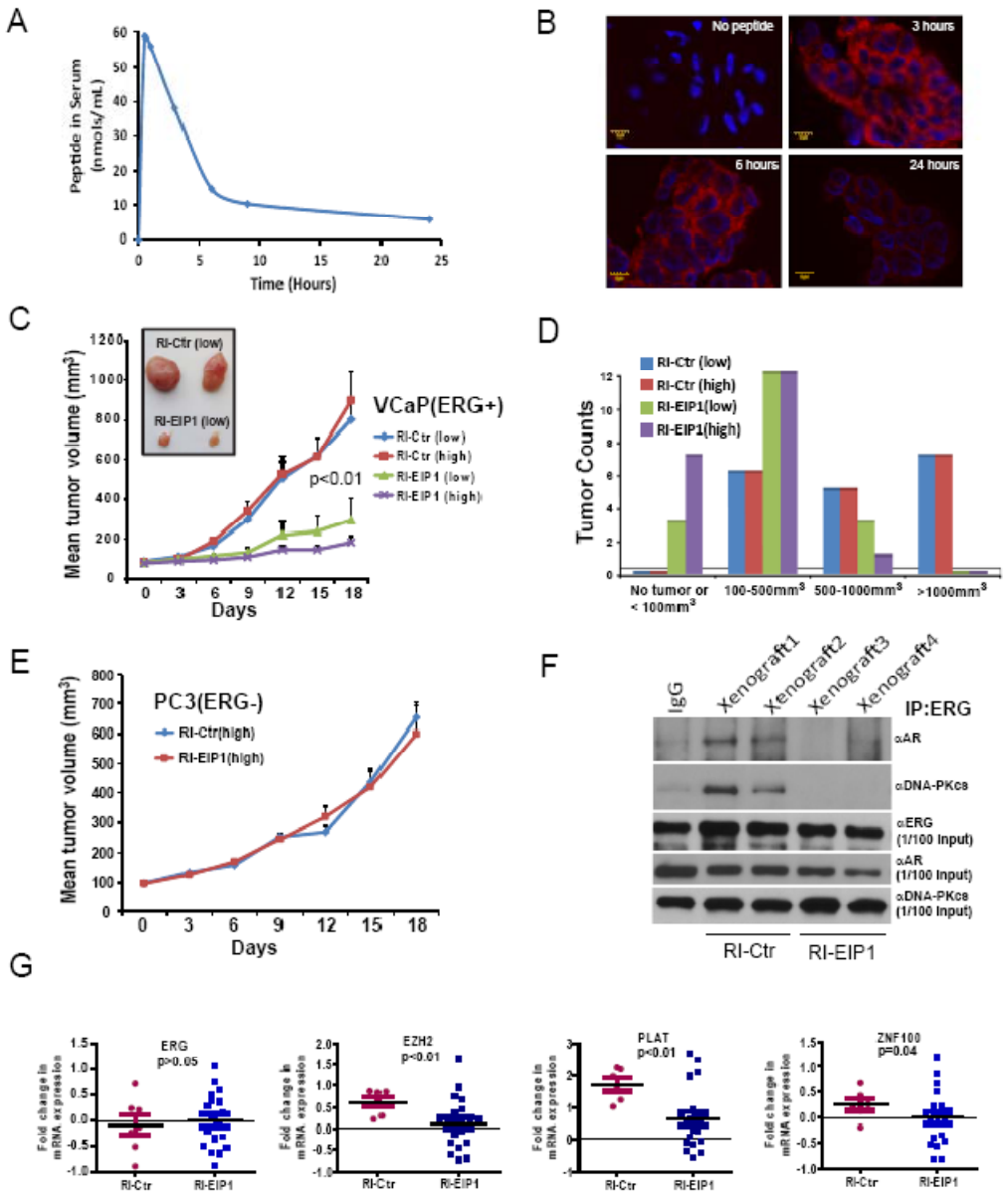



**Figure 5.11: Retroinverso EIPs suppress tumor growth *in vivo*.** (A) The serum concentration of biotinylated RI-EIP1 after intraperitoneal administration of 25mg/kg in mice harboring VCaP xenografts was assessed by colorometric assay. Serum was taken at indicated time points (x-axis), and the concentration of biotinylated peptides was determined by chemical reaction with avidin-HRP (y-axis). (B) Immunofluorescence staining of VCaP xenografts injected with biotinylated RI-EIP1. Slides were stained using Alexa 594–streptavidin and counterstained with DAPI. (C) VCaP xenografted mice treated with RI-EIP1 or Ctr at 25mg/kg or 10mg/kg for 18 consecutive days. Caliper measurements were taken every three days and average tumor volume (in $mm^3$) was plotted. Inset shows the representative VCaP tumors treated by RI-EIP1 or Ctr. (D) Distribution of VCaP xenograft tumors at day 18 categorized by tumor size. (E) PC3

xenografts were treated with 25mg/kg RI-EIP1 or Ctr and average tumor volume plotted as in (D). (F) VCaP xenograft tumors that were treated with 25mg/kg RI-EIP1 or RI-Ctr were used to perform IP-Western blot with an anti-ERG antibody. Tumor lysate (1/100 input) were loaded as positive controls for ERG, AR and DNA-PKcs. (G) Quantitative PCR was used to assess the relative expression of several ERG target genes in VCaP xenograft tumors treated as in (F). For all experiments mean ± SEM is shown.

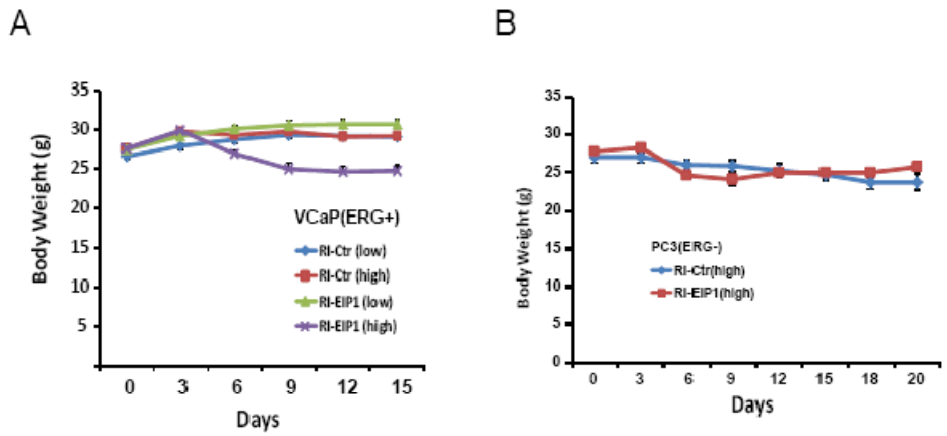



**Figure 5.12: VCaP xenografts show sustained therapeutic response following RI-EIP.** (A and B) Mice from VCaP or PC3 cell xenograft experiments, respectively, were weighed at the time of caliper measurement. Groups of mice treated with high or low dose peptide are indicated. Graphs are shown with +/- SEM unless otherwise indicated.

| Peptides | Nucleotide Sequences | Clone Repeat |
|---|---|---|
| FSASSPA | AGCAGGCGAAGAAGCAGAAAA | 1 |
| FSFGSLP | CGGCAGAGAACCAAACGAAAA | 1 |
| | AGGAAGAGAACCAAAACTAAA | 2 |
| FTFGSLP | AGGCAGAGAACCAAACGTAAA | 1 |
| | CGGAAGCGAACCAAACGTAAA | 2 |
| LAFGSLP | AGGCAGCGACCCAAAAGCAAG | 1 |
| LRFGTLP | AGGCAAAGTCCCAAAACGCAA | 1 |
| LSFGSFP | AGGAAACGAACCAAAAGACAA | 2 |
| LSFGSLP | AGGAAGAGAACCAAAAGACAG | 6 |
| | AGGAAGAGACCCAAAAGAAAG | 4 |
| | AGGCAGAGAACCAAACGACAA | 3 |
| | AGGCAGAGACCCAAACGAAAG | 2 |
| | CGGAAGAGACCCAAACGACAG | 14 |
| LSFGTFP | CGGAAAAGTACCAAAAGAAAG | 1 |
| LTFGSLP | AGGCAGACTACCAAAAGTAAG | 11 |
| MTFGTLP | AGGCAGCGTACCAAAAGTCAT | 2 |
| YTFGTLP | CGGCAAAGTACCAAACGTATA | 3 |
| LPPYLFT | CGTAAAAAGATAAGGAGGTAG | 2 |

**Table 5.1: Sequences of 59 phage clones sequenced following final bio-pan.** "Peptides" are the encoded amino acid residues translated from the "Nucleotide Sequences" corresponding to the unique open reading frame in the phage vector DNA backbone. "Clone Repeat" indicates the number of replicate phage clones enriched for the unique DNA sequence.

## REFERENCES


1. S. A. Tomlins *et al.*, *Nature* **448**, 595 (Aug 2, 2007).
2. S. A. Tomlins *et al.*, *Science* **310**, 644 (Oct 28, 2005).
3. S. A. Tomlins *et al.*, *Cancer Res* **66**, 3396 (Apr 1, 2006).
4. C. Kumar-Sinha, S. A. Tomlins, A. M. Chinnaiyan, *Nat Rev Cancer* **8**, 497 (Jul, 2008).
5. B. E. Helgeson *et al.*, *Cancer Res* **68**, 73 (Jan 1, 2008).
6. S. A. Tomlins *et al.*, *Neoplasia* **10**, 177 (Feb, 2008).
7. C. Sun *et al.*, *Oncogene* **27**, 5348 (Sep 11, 2008).
8. J. Wang *et al.*, *Cancer Res* **68**, 8516 (Oct 15, 2008).
9. O. Klezovitch *et al.*, *Proc Natl Acad Sci U S A* **105**, 2105 (Feb 12, 2008).
10. K. G. Hermans *et al.*, *Cancer Res* **68**, 7541 (Sep 15, 2008).
11. J. C. Brenner *et al.*, *Cancer Cell* **19**, 664 (May 17, 2011).
12. B. S. Carver *et al.*, *Nat Genet* **41**, 619 (May, 2009).
13. J. C. King *et al.*, *Nat Genet* **41**, 524 (May, 2009).
14. Y. Zong *et al.*, *Proc Natl Acad Sci U S A* **106**, 12465 (Jul 28, 2009).
15. G. Attard *et al.*, *Cancer Res* **69**, 2912 (Apr 1, 2009).
16. J. S. de Bono *et al.*, *N Engl J Med* **364**, 1995 (May 26, 2011).
17. C. Tran *et al.*, *Science* **324**, 787 (May 8, 2009).
18. J. E. Darnell, Jr., *Nat Rev Cancer* **2**, 740 (Oct, 2002).
19. C. Klambt, *Development* **117**, 163 (Jan, 1993).
20. S. Carrere, A. Verger, A. Flourens, D. Stehelin, M. Duterque-Coquillaud, *Oncogene* **16**, 3261 (Jun 25, 1998).
21. R. Fenrick *et al.*, *Mol Cell Biol* **19**, 6566 (Oct, 1999).
22. A. Verger *et al.*, *J Biol Chem* **276**, 17181 (May 18, 2001).
23. L. H. Brand, T. Kirchler, S. Hummel, C. Chaban, D. Wanke, *Plant Methods* **6**, 25 (2010).
24. J. S. Wadia, S. F. Dowdy, *Adv Drug Deliv Rev* **57**, 579 (Feb 28, 2005).
25. A. Efthymiadis, L. J. Briggs, D. A. Jans, *J Biol Chem* **273**, 1623 (Jan 16, 1998).
26. J. Yu *et al.*, *Cancer Cell* **17**, 443 (May 18, 2010).
27. O. Vafa *et al.*, *Mol Cell* **9**, 1031 (May, 2002).
28. Q. Cao *et al.*, *Oncogene* **27**, 7274 (Dec 11, 2008).
29. E. L. Snyder, B. R. Meade, C. C. Saenz, S. F. Dowdy, *PLoS Biol* **2**, E36 (Feb, 2004).
30. L. C. Cerchietti *et al.*, *Blood* **113**, 3397 (Apr 9, 2009).
31. J. Kim, W. Yu, K. Kovalski, L. Ossowski, *Cell* **94**, 353 (Aug 7, 1998).
32. I. S. Jeon *et al.*, *Oncogene* **10**, 1229 (Mar 16, 1995).
33. P. H. Sorensen *et al.*, *Nat Genet* **6**, 146 (Feb, 1994).
34. S. A. Shurtleff *et al.*, *Leukemia* **9**, 1985 (Dec, 1995).
35. T. Oikawa, T. Yamada, *Gene* **303**, 11 (Jan 16, 2003).
36. A. Albihn, J. I. Johnsen, M. A. Henriksson, *Adv Cancer Res* **107**, 163 (2010).

# Chapter VI

# Discussion: Mechanistically-Informed Personalized Therapy for Transcription Factor Driven Disease

## SUMMARY

A primary goal of this thesis was to develop a clinically viable therapeutic approach for cancers driven by ETS transcription factors. The rationale for choosing ETS transcription factor-driven cancers as a model for transcription factor-driven disease was due to the high prevalence of rearrangement in multiple tissue types, the ability to assess ETS status in a clinically meaningful time frame, and the basic mechanistic understanding that cancers harboring ETS gene fusions are dependent on their continued expression for viability (reviewed in (*1*)). In this thesis, I presented evidence that hypothesis-based experiments can lead to the identification and development of a clinically viable drug regimen, which specifically disrupts tumor cell growth in xenograft models.

In this chapter, I will integrate the primary data published for this thesis with emerging data published by other groups with a focus on: 1) the ETS complex using recently published DNA-PKcs structural information as well as the model for the DNA-PKcs:PARP1 interaction, 2) the implications of ETS-

mediated DNA damage for cancer development and therapeutics as well as highlight the key similarities and differences between ETS-mediated and cMYC-mediated DNA damage, 3) the current uses of peptides in the clinic and approaches that are being used to convert peptides such as the EIP into clinically useful agents, 4) the lessons from existing PARP inhibitor data that can be applied to using PARP and DNA-PKcs inhibitors in ETS positive cancers, and, finally, 5) the implications of this work for personalized medicine.

## IMPLICATIONS OF EMERGING STRUCTURAL CONCEPTS

Over the past two years several significant structural advances have been made for both DNA-PKcs and PARP1 that can be used to build hypotheses about the ERG:DNA-PKcs complex and ERG:PARP1 interaction. As detailed mechanistic studies further defining the importance of the interaction will be helpful in providing an improved understanding of the functional significance of the ETS:DNA-PKcs complex for ETS-mediated biology, additional structural information must be determined. DNA-PKcs is over 4000 amino acids long and has only recently been crystallized (*2*) and, while the structure of PARP1's catalytic domain has been established, several additional PARP1 complexes are also beginning to emerge. In this section, I will discuss the implications of this data for ETS-mediated biology and some potential future directions for ETS, DNA-PKcs and PARP1 related structural biology.

**ETS structural biology**

Here, we provided the first crystal structure of an ETS DNA binding domain without DNA and the first structure of the ERG ETS DNA binding domain. This suggests that the ETS DNA binding domain is ordered in the absence of DNA, unlike many other transcription factor DNA binding domains (*3, 4*), and suggests that this structure can be used in virtual binding models to discover potential binding small molecules. We also found that mutation of Y373A within the ETS DNA binding domain blocks the ERG:DNA-PKcs interaction. Importantly, the motif containing Y373A is adjacent to the alpha-helix that fits into the major groove of DNA and that this motif is accessible to potential interacting proteins. In the future, studies will need to focus on crystallizing full length ERG as well as the ERG:DNA-PKcs interface.

**Concepts from emerging DNA-PKcs structural data**

In contrast to ETS DNA binding domains, for which some structural information was available at the onset of this thesis, a coarse X-ray crystallography structure of DNA-PKcs was only recently published (*2*). This structure provides the first framework to help build hypotheses about the interaction between DNA-PKcs and ETS transcription factors. In this structure, DNA-PKcs was shown to take on an open circle structure that is postulated to encompass Ku heterodimer-capped DNA ends such that the kinase domain of DNA-PKcs is almost fully exposed and available for potential protein:protein interactions.

Although the model published by Sibanda *et. al.* is the only X-ray crystal structure published thus far, it is important to note some previously published structural data suggesting that DNA-PKcs can adopt substantially different conformations by binding to different DNA structures (*5*), or upon autophosphorylation (*6*). In the future, more detailed structural information may allow for the generation of co-crystals using the ERG ETS DNA binding domain and the DNA-PKcs interaction domain. Because DNA-PKcs conformations can shift, it will also be important to understand if the interaction takes place on normal DNA, broken DNA or in the absence of DNA. Once solved, however, this structural information will enable the generation of competitive peptides and non-binding mutants of DNA-PKcs to specifically study the importance of the ETS:DNA-PKcs interaction. Our peptide data indicates that disrupting the interaction may be a viable therapeutic target, but does not discriminate between DNA-PKcs, AR and other co-factors that may bind to this surface.

**Importance of the DNA-PKcs:PARP1 interaction for ETS-mediated biology**

The first observation that PARP1 and DNA-PKcs function together came from studies of V(D)J recombination in mice lacking both enzymes. In contrast to PARP1 knockout mice, which develop with normal genomes, double knockout mice have an increased recombination activity that drives chromosomal abnormalities and rearrangements (*7*). Soon after this initial report, detailed biochemical analysis led to the identification of PARP1 as a DNA-PKcs interacting enzyme, and it was demonstrated that DNA-PKcs phosphorylates

PARP1 to increase ribosylation activity, and in turn, PARP1 ribosylates DNA-PKcs to increase its phosphorylation activity (*8*). However, the functional effects of the interaction to transcription and DNA repair pathways have not been studied.

Recently, electron microscopy was used to visualize the DNA-PKcs:Ku70:Ku80:PARP1 complex on Y-shaped DNA to an approximate resolution of 36Å (*9*). The model is consistent with our ERG:DNA-PKcs:PARP1 model, in that the DNA-PKcs:PARP1 interaction appears to be DNA-independent (*9*). In contrast to the conformations previously noted for DNA-PKcs, the authors found that DNA-PKcs had many unique structural features when in complex with PARP1 on Y-shaped DNA such that it resembled DNA-PKcs on hairpin DNA. While DNA-PKcs is thought to interact with Ku70 and Ku80 in a dimer-of-trimers complex such that DNA-PKcs subunits phosphorylate the HEAT repeats in *cis,* Spanolo *et. al.* found that 4 protein complex takes on a dimer-of-tetramers conformation and postulated that the DNA-PKcs kinase domain shifts to phosphorylate its HEAT repeats in *trans* (9). Importantly, the model suggests that PARP1 can cause dramatic structural changes to DNA-PKcs and suggests that it may be important to model the ERG:DNA-PKcs interaction with and without PARP1. Furthermore, this paper used docking models to fit the crystal structure of the PARP1 catalytic subunit (chicken protein) into the DNAPK structure (*2*) and found that the PARP1 catalytic subunit occupies a region near the HEAT repeats at the base of DNA-PKcs. This model was consistent with previous reports showing that the PARP1 BRCT domain interacts with Ku (*10*). As postulated by

this model, therefore, PARP1 may have a direct interaction with the base of DNA-PKcs as well as an indirect interaction mediated by the Ku heterocomplex.

This model is important for understanding the possible relationship of ERG with DNA-PKcs:PARP1. Our data is clear that ERG interacts directly with DNA-PKcs in a DNA-independent manner in cell lines, tissues and with purified proteins. As the Ku heterodimer does not interact with DNA-PKcs in the absence of DNA (*11*), the model suggests that ERG can interact with DNA-PKcs:PARP1 in the absence of the Ku heterodimer, which is consistent with our data. Furthermore, we have now mapped the ERG interaction site to the DNA-PKcs crown region containing the PI3/4 catalytic domain (Data not shown) suggesting that ERG forms a trimeric complex with DNA-PKcs and PARP1. Because our mapping data suggests that PARP1 and ERG bind to opposite ends of DNA-PKcs such that they do not physically interfere with one another, the data is consistent with the model by Spanolo *et. al.*. In line with this hypothesis, we found that PARP1 was capable of ribosylating DNA-PKcs as has been previously shown (*8*), but did not ribosylate *in vitro* purified ERG protein (Data not shown).

Taken together, the available structural information suggests multiple functional possibilities for how DNA-PKcs and PARP1 modulate ERG activity or *vice versa*. As we have shown that Y373 is required for ERG to interact with DNA-PKcs, it is clear that DNA-PKcs is interacting with the ETS DNA binding domain of ERG. Consistent with our mapping data, we have also shown that several ETS transcription factors including ETS1, ETV1 and ETV6 can interact with DNA-PKcs. This suggests the possibility that DNA-PKcs modulates ETS

activity by directly phosphorylating the transcription factor. Indeed, DNA-PKcs has recently been shown to phosphorylate ETS1 *in vitro* (*12*), which would be consistent with how DNA-PKcs has been proposed to modulate transcriptional activity of USF1 (*13*). Alternatively, the model suggests the possibility that ERG overexpression disrupts either NHEJ or AEJ via direct inhibition of the DNA-PKcs kinase domain. It is important to note that ERG overexpression does not disrupt the formation of H2A.X foci, and actually imparts resistance to cell lines treated with double strand inducing agents such as radiation and $H_2O_2$ (Data not shown). As such, more biochemical and functional studies need to be performed to fully understand the consequences of the ERG:DNA-PKcs interaction. In the future, studies disrupting the DNA-PKcs:PARP1 interaction will help inform models of the putative role of the ERG:DNA-PKcs:PARP1 trimeric complex in ETS-mediated biology.

**Dissecting the mechanistic role of PARP1 in ETS-biology**

Very recently, structural studies of PARP1 have led to the identification of critical mutants that can modulate individual functions of PARP1, ie. transcription versus DNA repair (*14*). Detailed studies employing mutants altering these individual phenotypes will help further the understanding of the role of PARP1 in ETS-mediated biology. As molecules such as Olaparib compete for the NAD+ binding site in PARP1 (*15*) and, disrupt poly(ADP)-ribosylation of target proteins, controls for these studies will include the E988K mutant, which has been shown to lose the ability to extend (ADP)-ribose chains (*16*) and should mimic the

effects of Olaparib. Likewise, even though the interaction site between DNA-PKcs and PARP1 has not been precisely mapped, it will also be important to study mutants disrupting this interaction for ETS signaling.

Finally, it is possible that PARP inhibitors exert their effects on ETS positive cells by disrupting a poly(ADP) ribosylation-dependent signaling pathway. Thus, it will likely be important to pursue the use mass spectrometry-based studies to identify proteins that are differentially poly(ADP)-ribosylated in ETS positive cancers. Regardless of the means, however, a better functional understanding of the mechanistic role of PARP1 in ETS-mediated biology may lead to additional therapeutic targets for ETS-positive cancers as well as provide a mechanistic basis for potential mechanisms of resistance to PARP inhibitors. These studies have not yet been published in *BRCA1/2*-mutant models.

**Synopsis of structural data**

Taken together, the functional and structural studies that have emerged within the last year will allow for a more detailed analysis of the ERG:DNA-PKcs and ERG:PARP1 interactions. These studies will be important to identify the molecular importance of both DNA-PKcs-mediated phosphorylation and PARP-mediated poly(ADP)-ribosylation. It is likely that this protein modification cascade plays a key role in ETS-mediated transcription through either direct modification of the ETS transcription factor itself, or by modification of other critical interacting proteins. Finally, these studies may shed light on to potential

resistance mechanisms that will develop in ETS positive cancers following treatment with either PARP or DNA-PKcs inhibitors.

## IMPORTANCE OF ETS INDUCED DNA DAMAGE

One of the critical discoveries of this thesis was the observation that ETS gene overexpression drives DNA damage in multiple genetic backgrounds. Here, I will discuss the implications of ETS induced DNA damage for disease progression, the mechanistic insights gained from EIP studies, and finally compare and contrast ETS-induced DNA damage with recent models of DNA damage caused by another transcription factor, cMyc.

### Implications of ETS-induced DNA damage

We have now shown that *wt-ERG*, *ETV1*, *ETV4*, *ETV5* and *TMPRSS2*:*ERG* all drive the accumulation of DNA double strand breaks when overexpressed in cells. In contrast, normal human umbilical vein endothelial cells (HUVECs) which have high endogenous wt-ERG do not have elevated baseline levels of DNA damage suggesting that this phenotype is dependent on expression level (data not shown). In normal cells, DNA damage leads to senescence and apoptosis (*17*) and, in some model systems, a single DNA double strand break can be lethal (*18*). While enforced ERG expression leads to DNA double strand breaks, it does not induce either premature senescence or apoptosis in non-transformed primary human prostate epithelial cells or in primary HUVECs

(Manuscript in preparation Feng and Brenner). This suggests that one consequence of ERG overexpression is to allow a cell to tolerate DNA damage.

In line with this hypothesis, isogenic DU145 and PC3 ERG overexpressing cells were both significantly more resistant to radiation and H2O2-induced DNA damage (both cause double strand breaks) (Manuscript in preparation Feng and Brenner). Interestingly, we found that ERG mutants that lose either the ability to activate transcription or the ability to interact with DNA-PKcs and PARP1 are not resistant to radiation. This suggests that either ERG:protein interactions or the ERG transcriptional program is important for the observed tolerance of DNA damage. In the future, detailed studies will help determine if this is due to an increased ability to repair DNA double strand breaks or a decreased ability to sense DNA damage.

Beyond the precise mechanism that allows ERG positive cancer cells to tolerate DNA damage, the presence of DNA breaks as well as the potential repair mechanisms used to repair these breaks has profound implications for how ETS gene fusions drive cancer progression. Indeed, studies determining the frequency and genomic distribution of DNA damage will be critical for understanding these implications. For example, if DNA damage is occurring at sites of ERG-regulated transcription, as may be expected from recent studies indicating that areas of increased transcription are more likely to undergo translocation (*19, 20*), then it is possible that ERG-regulated promoters and enhancers are sustaining DNA damage. This is important because NHEJ induces minor nucleotide changes during the repair process such that mutations would be introduced at the sites. In

this case, ERG could drive a targeted mutator phenotype. Importantly, whole genome sequencing studies will soon determine if ETS positive cancers have a higher rate of mutation than ETS negative cancers, and if these mutations are correlated with sites of ERG binding.

Whether ETS-induced is targeted or random, it has implications for the clinical course of disease. ETS gene fusions occur early in PCa development in the transition of HG-PIN to PCa; however, ETS positive PCa can remain latent for years only to suddenly become very aggressive. In line with this notion, if one important mechanistic role of ETS gene fusions is to induce DNA damage, then waiting for ERG to generate a spontaneous secondary lesion may explain the long latency often experienced before ETS positive tumors become highly metastatic. In contrast, ESFTs are highly aggressive at the onset and ESFT fusions induce more DNA damage than their counterparts in prostate cancer. This may be due to a variety of mechanisms, but a more active mutator phenotype may explain the rapid propensity for recurrence in this disease.

**Insights from EIP studies**

Mechanistically, this thesis indicates that EIPs or other small molecules that disrupt ERG protein:protein interactions may also block several ERG-mediated functions. Consequently, EIPs are also critical for improving our mechanistic understanding of how ETS gene fusions drive prostate cancer progression. While it is known that ERG drives many oncogenic phenotypes through the induction of a unique transcriptional program, the mechanism leading

to DNA damage is not clear. For example, it is unknown if the DNA breaks are the result of an aberrant transcriptional program (for example, one that leads to increased exonuclease expression) or if ERG itself is mechanistically involved in the generation of breaks. Because disruption of the ERG:DNA-PKcs interaction with the EIP or Y373A mutation led to a decrease in ERG-mediated DNA damage, one hypothesis to be explored is that ERG sequesters DNA-PKcs, thereby, preventing it from rapidly repairing endogenous DNA damage. Alternatively, the EIP and ERG DNA binding domain mutant prompt the hypothesis that ERG-mediated DNA breaks occur in an ERG-targeted manner that requires both DNA binding and ERG-localized DNA-PKcs. As such, the use of EIPs has led to several additional postulates as to how ERG causes DNA double strand breaks.

Because ERG-mediated DNA damage likely leads to the accumulation of mutations, inhibitors that block this phenotype may have the added benefit of impeding the gradual evolution of a cancer cell's genome. Importantly, this study demonstrates that ERG positive xenograft growth can be inhibited in a DNA damage-independent manner with EIPs, which directly contrasts with our previous work demonstrating that inhibition of PARP1 blocks ETS positive xenograft growth through the synergistic induction of DNA damage (*21*). As such, because the EIPs and PARP inhibitors use opposing mechanisms to block cell growth, one would expect combination therapy to be antagonistic indicating that sequential inhibition of ERG with EIPs followed by PARP and/or DNA-PKcs inhibitors may be the most viable therapeutic option.

**Comparison to the cMYC-induced DNA damage model**

cMyc is a transcription factor that has long been known to induce DNA damage. Despite this association, it took many years to unravel the mechanism of cMyc-induced DNA damage. In a critical study by Dominguez-Sola *et. al.*, mammalian cells and *Xenopus* cell-free extracts devoid of RNA transcription were used to demonstrate a non-transcriptional role for c-Myc in the initiation of DNA replication (*22*). Interestingly, this study also showed that activation of cMyc caused frequent replication-induced DNA damage at sites of replication initiation, which then persisted throughout the cell cycle (*22*). This suggested that cMyc partially allowed cells to bypass normal cell cycle checkpoints despite the presence of DNA damage. However, in contrast to ETS-mediated DNA damage, cMyc induced DNA damage is not always tolerated and can drive both senescence and apoptosis in cells (*23, 24*).

This phenotypic difference is due to several potential mechanisms that need to be explored in the future. Because cMyc had long been known to induce DNA damage, I used cMyc overexpression as a positive control in several of the DNA damage experiments. In contrast to cMyc-induced DNA damage, ERG induced less DNA damage per cell as assessed by γH2A.X foci counting in transformed RWPE and PC3 cells. As such, ERG-induced DNA damage may be more "tolerable" due to the fact that it is less abundant. Unfortunately, these studies are complicated by the fact that normal cells senesce after only a few passages regardless of transcription factor overexpression. Likewise, a detailed

and clinically verified mechanism of DNA damage should be first established for both cMyc and ERG before assessing questions about the apparent differences in DNA damage tolerance between the models. Fortunately, as noted above, high throughput sequencing of tumors will soon determine if cancers have mutation "hotspots" that are enriched for specific transcription factor binding sites. Accordingly, if the cMyc hypothesis presented by Dominguez-Sola *et. al.* is correct, then one would expect that cancers with cMyc overexpression will be associated with mutations at replication origins.

**Conclusion of ETS-mediated DNA damage**

Taken together, ETS-mediated DNA damage clearly has implications for both disease pathogenesis and therapeutic design in all ETS-driven cancers. While ETS gene overexpression drives the accumulation of DNA damage, normal ETS gene expression does not cause sustained DNA damage as assessed by γH2A.X foci formation or neutral COMET assays. As such, this transcription factor mediated DNA damage appears to be a cancer-specific phenomena. In the future, therapeutic approaches to further potentiate ETS-induced DNA damage may have a viable benefit, and similar approaches may be beneficial for cancers with cMyc overexpression. Surprisingly, this approach has not yet been explored for cMyc-driven malignancies.

## DEVELOPING EIPs AS CLINICAL THERAPEUTICS

In order to improve the design of the EIP to create a molecule more likely to enter the clinical setting, it is important to understand if and how peptides are currently used in the clinic as well as what methods are used to convert peptides into "drug-like" molecules.

### Peptides in the clinic

While peptide-based therapies are currently underrepresented in the list of FDA approved drugs, 50 peptide therapeutics have been approved for clinical use in various diseases (*25*). Studies using these peptide based therapeutics have given insight into both the bioavailability and physiological consequences of administering peptides in general. For example, peptides have a short biological half-life as they are rapidly cleaved by proteases and lost during renal filtration. Thus, potential peptide therapeutics require molecular modifications to both avoid degradation and glomerular filtration (*26*). Peptides are also not generally orally available compounds, so treatment protocols are dependent on injections, controlled-release formulations or osmotic pumps (*27*). Of the peptide-based therapeutics that have achieved clinical success, the two most prominent are insulin, used to treat diabetes, and Fuzeon, which is used for the treatment of multi-drug resistant HIV. Fuzeon is postulated to function by competitively blocking viral fusion with the host cell (*28*). Consequently, while very few of these therapies target intracellular proteins, the data indicates that large molecular

peptide-based therapies can be implemented successfully in the clinic for various diseases.

In terms of cancer therapeutics, however, only a few peptide-based therapies have been approved for cancer treatment. One such molecule that was designed to mimic a di-peptide structure (N-, C-terminal protected pyrazinoic acid-phenylalanine-leucine) is a modified peptide that is used to treat multiple myeloma. This peptide mimetic is a proteasome inhibitor called Bortezomib. Despite its derivation, however, Bortezomib really represents a more classical small molecule-based approach in that the entire molecule is engulfed by the protein target in a hydrophobic pocket rather than binding to an extended protein surface area as required by for function of the EIP. Likewise, several obvious differences are apparent between the di-peptide-based structure and our 7-amino acid-based EIP sequence. Nonetheless, in the future, it is possible that the EIP length could be reduced and/or modified to create molecular properties assimilating more traditional chemotherapeutics such as Bortezomib. As such, molecular modifications to peptide backbones can lead to therapeutics that target intracellular processes and may provide one avenue for the further development of the EIP.

Beyond small molecule (<1kDa) peptide-like therapeutics such as Bortezomib, several groups have also attempted to use larger peptide-based structures to target intracellular proteins. One example of such an approach is the use of BH3 peptide mimetics (though these mimetics still target a hydrophobic pocket). These therapeutics target a protein called Bcl2 that is frequently

amplified in cancer to block a cell death pathway called apoptosis (reviewed in (*29*)). Normally, Bcl2 regulates apoptosis by sequestering the pro-apoptotic proteins Bax and Bak at the surface of the mitochondria. Upon initiation of apoptosis a BH3-only protein such as Bid or Bad (*30*), and possibly a few non-BH3-only proteins like p53 (*31*), translocate to Bcl2 disrupt the Bcl2:Bax and Bcl2:Bak interactions (*32*). This serves as one of the first checkpoints for the apoptosis pathway at the level of the mitochondria. BH3 peptide mimetics, replicate this process by disrupting the Bcl2:Bax and Bcl2:Bak interactions (*33*). The unique modification that Walensky ***et. al.*** made was to incorporate a "hydrocarbon staple" on the backside of the BH3 peptide such that it improved stability of the α-helical peptide structure in a way that did not alter Bcl2 binding (*33*). Interestingly, this modification was also shown to make both the BH3-peptide mimetics as well as Fuzeon resistant to proteolytic cleavage (*34*). Hydrocarbon stapled peptides are still being explored clinically, but represent a promising advance for peptide-based therapeutics. If the crystal structure of the ERG:EIP interaction is solved in the future, "hydrocarbon stapling" the backside of the EIP may be one option to further improve the stability of our ERG therapeutic.

Taken together, peptides used as therapeutics are emerging as potential options for clinical use in many diseases, not just cancer. However, much research is needed to determine if this therapeutic approach will achieve clinical efficacy for intracellular protein surface targets. Consequently, the fastest way into the

clinic may be to pursue alternative screening and hypothesis-driven approaches to identify small molecules that disrupt the ERG:EIP interaction.

## CLINICAL IMPLICATIONS OF PARP AND DNA-PKcs INHIBITORS

With the discovery that ETS gene fusion positive cancers are highly susceptible to PARP inhibitors, clinical trials are emerging or have already been initiated to extend the discovery into the clinic in both localized and castrate resistant PCa as well as in recurrent ESFTs. Thus, it is important to understand the current clinical indication of PARP inhibitors and both the pre-clinical and clinical data that has been generated with this class of drugs. Concepts such as successful drug combinations and resistance mechanisms that drive recurrence may help predict how ETS positive patients will respond to this class of inhibitors. Here, I will discuss the implications of the published PARP inhibitor studies for ETS-driven cancers and I will analyze the perspective use of newly developed DNA-PKcs inhibitors for ETS positive cancers.

### Biomarkers of PARP susceptibility

In designing clinical trials for histopathologic disease sites associated with ETS rearrangements, it is important to understand what other molecular lesions may impart sensitivity to PARP inhibitors. The first genetic lesion to be associated with PARP inhibitor sensitivity was *BRCA1/2* mutation (*35*, *36*). Importantly, a low prevalence of *BRCA1/2* mutations are found in prostate cancer (*37*) and, more recently, targeted screening of men from families with *BRCA1/2*

mutations has demonstrated a high prevalence of aggressive PCa in these patients (*38*). As such, a subset of patients in both PCa clinical trial settings would be expected to respond based on *BRCA1/2* mutation status alone. In contrast, no reports of *BRCA1/2* mutations in ESFTs have been published and it is unclear if these mutations play a role in ESFT development or progression.

In addition to germline mutations in in the *BRCA1/2* genes, the original discovery of *BRCA1/2* mutation-induced PARP sensitivity led to a search for additional "*BRCA*-like" cancers. Not surprisingly, the first "*BRCA*-like" phenotype discovered was actually a set of cancers that have loss of *BRCA1* expression due to promoter hypermethylation (*39*). Beyond this association, more recent studies have indicated that PTEN loss can lead to PARP inhibitor sensitivity in some models (*40-42*). This is of note because PTEN is mutated or lost in 30-40% of PSA screened prostate cancers (reviewed in (*43*)). These studies have been perplexing in relation to the data presented here, as some of the PTEN mutated and null prostate cancer cell lines such as PC3 cells (*44*) were not at all sensitive to PARP inhibition by either siRNA or Olaparib until an ETS gene was overexpressed. It should also be mentioned that VCaP cells have intact PTEN (*44*). In line with the PTEN hypothesis, a patient with metastatic ovarian cancer (BRCA1/2 WT, PTEN null) responded to a single of cycle of PARP inhibition, but recurred within 8 months (*45*). Unfortunately, however, the tumor was not sequenced and additional genetic information could not be gained from the anecdotal report. Recently, a paper has come out refuting the mechanistic hypothesis linking PTEN with PARP inhibitor sensitivity (*46*). Consequently, this

genetic relationship requires additional research to be validated, but should be monitored as a potential biomarker for response in PARP inhibitor trials. To date, no other specific genetic lesion has been reported to correlate with PARP inhibitor sensitivity.

**Lessons from preclinical and clinical PARP inhibitor data**

In addition to monitoring additional genetic lesions that may predict for response, clinical designs in PCa and ESFTs can incorporate much of the information that has been published surrounding PARP inhibitor development for other indications in both pre-clinical and clinical models. In fact, given the high sensitivity of *PARP1*-/- mice to alkylating agents and γ radiation (*47, 48*) as well as the reports linking *BRCA1/2*-deficiency to PARP inhibitors (*35, 36*), many drug companies have developed PARP inhibitors. At this point, the primary PARP inhibitors that have been used in the clinic are Olaparib (AZD2281) and ABT-888 (Veliparib).

Olaparib was developed by Kudos pharmaceuticals from a 4-benzyl-2H-phthalazin-1-ones backbone as a PARP1 and PARP2 inhibitor ($IC_{50}$ 0.005μM and 0.001μM, respectively, 1.5μM for tankyrase-1) and chosen for its pharmacokinetic profile, oral bioavailability and *in vivo* efficacy against SW620 *BRCA1*-deficient colon cancer xenografts (*15*). Olaparib is thought to make hydrogen bonds with Ser904 and Gly863 in the nicotinamide substrate binding pocket (*15*), but the crystal structure of PARP1-olaparib has not yet been published. Olaparib has been shown to slow the growth of many *BRCA1/2*-

deficient cell lines after 5 days *in vitro* (*49*) and *in vivo* (*50, 51*). Olaparib concentration peaks in mouse serum 30 minutes after IP injection (50mg/kg) corresponding to maximum inhibition of PARP activity after injection (*51*). In these *in vivo* models, Olaparib has been shown to potentiate the effects of cisplatin and carboplatin (*50, 51*). Olaparib has been shown to delay cell growth in a transplant xenograft model of primary human *BRCA2*-mutant ovarian cancer (*52*), which showed a similar magnitude of effect as our ETS positive human primary prostate cancer xenograft models. Prior to our papers, Olaparib had not been shown to potentiate the effects of temozolomide, and we have not tested the combination with platinum-based regiments in ETS positive cancers. Of note, PARP2 has 5-fold lower mRNA expression than PARP1 in VCaP cells and was not detectable by Western blot (Data not shown).

Similar to Olaparib, ABT-888 was derived from a lead structure that was known to hydrogen bond with Ser904 and Gly863 and form π-stacking interactions with Tyr907 in the nicotinamide substrate binding pocket (*53*). ABT-888 was also chosen for its inhibition of PARP1 and PARP2 ($K_d$ 5.2nM and 2.9nM, respectively), pharmacokinetic profile, oral bioavailability and *in vivo* efficacy against murine melanoma and breast tumor models (*54*). Crystal structures of ABT-888 in complex with murine PARP2 confirmed binding to the nicotinamide pocket and suggested a rationale for greater PARP2 specificity; PARP2 has a Glu335 residue that may interact with the inhibitor whereas PARP1 has a shorter Asp776 residue that is less likely to interact (*55*). In contrast to Olaparib, ABT-888 has been most frequently studied with alkylating agents and

has been shown to potentiate the effects of this class of drug in many preclinical (*54, 56-59*) and clinical models (*60*). Of note, ABT-888 in combination with TMZ was shown to have a significant affect in T-ALL cell lines known to have high ERG expression, but not in pre-B-cell ALL cell lines known to have low ERG expression (*57*).

The first phase I clinical trial of a PARP inhibitor employed Olaparib and demonstrated that 100-400mg b.i.d. is well tolerated and sufficient to block PARP activity without causing significant side effects (*61*). This observation has subsequently been recapitulated in many phase I trials in different organ sites (*62-67*). Initial observations indicated that some *BRCA1/2*-mutant tumors regress completely (*61*), and that platinum-resistant *BRCA1/2*-mutant tumors are not sensitive to PARP inhibition whereas platinum-sensitive tumors are responsive (*63*). In more recent phase II studies of patients with *BRCA1/2* mutations, single agent Olaparib has been associated with an approximate 50% response rate (*64-66, 68*) and was been shown to enhance the effects of pegylated liposomal doxorubicin in one study of this patient population (*65*). Published studies have not extensively analyzed histo-pathologic populations that are known to harbor ETS fusions.

In addition to PARP inhibitor monotherapy, several combination studies using various chemotherapeutics or radiation have been completed in both the Phase I and II trial settings. For example, a combination regimen of Olaparib and bevacizumab (Avastin) was well tolerated in patients with advanced solid tumors (*62*). Some clinical studies evaluate low dose Olaparib with high dose of a

complementary alykylating agent. In these studies, 100mg Olaparib bid was well tolerated with either 1 mg/m$^2$ topotecan (*69*) or 600 mg/m$^2$ dacarbazine (focused on refractory melanoma) (*70*). Likewise, ABT-888 has been shown to enhance the effects of topotecan in adults with refractory solid tumors or lymphomas (*60*). Although some evidence has suggested that PARP inhibitors sensitize cells to radiation, this combination has not yet been attempted clinically (*71-75*). As such, careful design of clinical trials incorporating additional therapeutics may be required for all biomarker driven studies including both *BRCA1/2*-defiencey or ETS rearrangement.

Despite the early success of PARP inhibitors, one clinical trial worth noting because of its extraordinary failure occurred in the PARP inhibitor field using the Sanofi-aventis lead drug iniparib (BSI-201). In fact, several concerns were raised about PARP inhibitors in general as a result of this trial (*76*). With the success of the AstraZenica compound Olaparib in early clinical trials (*61*), several PARP inhibitors were rapidly being moved into the clinic to determine individual efficacy. Of these, Iniparib was being evaluated in triple negative breast cancer in combination with gemcitabine and carboplatin and improved average patient survival by approximately 5 months in a phase II study (*77*). When expanded to phase III, however, the overall survival benefit was lost (Sanofi-aventis press release). While questions remain about the phase II results published in NEJM, follow up studies completed independent of Sanofi-aventis demonstrated that BSI-201 is not a PARP inhibitor, but generally disrupts all proteins by non-specifically forming cysteine adducts (*78*). While most PARP inhibitors are

designed to compete for NAD+ binding in the PARP1 catalytic domain (*15, 53*), at best, BSI-201 weakly disrupted PARP1 activity because of the cysteine residues found in the DNA binding and catalytic domains (*78*). As such, a drug was approved for clinical trials without extensive, independent verification of the mechanism of action. Although the trial yielded a significant amount of information about BSI-201's general efficacy that would be valuable for the company, it also temporarily led to a high level of skepticism in the PARP inhibitor field. This is also another example of how randomized approaches to population-based medicine that lack biomarker-driven hypotheses are usually ineffective, even if the therapeutic has a safety profile capable of succeeding in phase I/II trials.

**Resistance to PARP inhibitors**

As PARP inhibitor trials move forward using *bona fide* PARP inhibitors, understanding the mechanisms driving resistance in both *BRCA1/2*-deficient tumors and ETS positive cancer cells is important. Several studies have now analyzed PARP inhibitor resistant in the former subtype in great detail. In both spontaneous *BRCA1*-/-*p53*-/- tumors as well as a transplant model of *BRCA2*-deficient cells, tumors initially regress, but eventually PARP inhibitor resistant tumors recur (*51, 79*). By continuously culturing the *BRCA2* mutant pancreatic cancer cell line, CAPAN1, in KU0058948 (an early PARP inhibitor), it was shown that cells can reactivate homologous recombination-mediated repair as seen by RAD51 foci experiments (*80*). Similar observations were reported for

platinum-resistant CAPAN1 cells, which were also resistant to the PARP inhibitor Olaparib (*81*). Surprisingly, sequencing of *BRCA2* revealed that PARP inhibitor or platinum based selection led to secondary deletions in *BRCA2*, which restored *BRCA2* function by removing the baseline frameshift mutation 6174delT and restoring most of the open reading frame to proper orientation (*80*). It is worth noting that this may represent a general mechanism for platinum resistance as both carboplatinum resistant patients as well as HCC1428 (*BRCA2*-mutant) platinum resistant cells harbor secondary, restorative mutations in *BRCA2*. Furthermore, in these studies, those *BRCA1/2*-mutant patients who recur following treatment with platinum based regimens may not be susceptible to PARP inhibition. Increased homologous recombination based repair has also been shown to cause resistance to ABT-888 *in vitro* (*82*).

As an alternative mechanism of resistance in *BRCA2*-mutant tumors, Rottenberg *et. al.* demonstrated in their *in vivo* model that upregulation of p-glycoproteins, which are multidrug resistance protein capable of pumping xenobiotics out of cells (*83*), are upregulated in PARP inhibitor resistant tumors (*51*). Consistent with this hypothesis, a p-glycoprotein inhibitor called tariquidar was able to restore PARP inhibitor sensitivity in this model (*51*). Of note, high levels of ADP-ribose, which would present in cells with an activated PARP1-PARG signaling axis (*84*), have been shown to inhibit the activity of several ABC transporters including p-glycoproteins (*85*). Despite these findings in cell lines, clinical evidence has not yet been published that upregulation of p-glycoprotein

expression or activity correlates with PARP inhibitor resistance in *BRCA1/2*-mutant human tumors.

In contrast to *BRCA1/2* mutant cancer cells, however, ETS positive cancer cells appear to have intact homologous recombination (*21*) and it is unlikely that ETS positive cancers will share the precise *BRCA*-restoration mechanism acquired in BRCA1/2 deficient cells. It is possible that PARP inhibitor resistance will occur due to either increased HR-efficiency in general or due to increased expression or activity of an ABC transporter such as one of the p-glycoproteins. Importantly, however, resistance studies may also lead to a better understanding of the mechanism of cell death following PARP inhibition in ETS positive cancers and will be important to understand the appropriate treatment course in these cancers.

**DNA-PKcs inhibitors in ETS positive cancers**

As DNA-PKcs directly interacts with ERG in prostate cancer cells and tissues, inhibitors of this enzyme may also have therapeutic potential in ETS positive cancers. Accordingly, while DNA-PKcs inhibition with either siRNA or NU7026, a first generation DNA-PKcs inhibitor, blocked ETS-mediated transcription and invasion in both PCa and ESFT cell line models, limitations on the availability of second line DNA-PKcs inhibitors prevented extensive analysis of this class of inhibitors. Of note, treatment of VCaP xenografted mice with NU7026 resulted in a significant decrease in tumor volume with limited effects on animal weight (Data not shown). In the future, DNA-PKcs inhibitors should be

explored as a potential therapeutic for ETS positive cancers and it will be interesting to determine if inhibition of DNA-PKcs is additive to PARP inhibition or EIPs.

**The future of PARP and DNA-PKcs inhibitors in ETS positive cancers**

As noted above, PARP inhibitors are being actively evaluated in the clinic. In total, thirty one phase I and II clinical trials have been opened using Olaparib as either a monotherapy or in combination with other agents; of these, only seven have been completed, and ten are already accruing patients (clinicaltrials.gov). As with Olaparib, fifty four clinical trials have been opened using ABT-888; of these, only five trials have been completed (clinicaltrials.gov). Finally, trials are open for the emerging PARP inhibitors Rucaparib (3) and MK4827 (1) (clinicaltrials.gov). These trials most frequently evaluate PARP inhibitors alone or in combination with an alykylating agent, usually in breast and ovarian cancers. Thus far, two trials have been officially opened to specifically test PARP inhibitors in castrate resistant prostate cancer as a result of this thesis work: 1) a phase II testing Olaparib as a monotherapy (*86*), and, 2) a phase I of ABT-888 in combination with temozolomide (NCT01085422). Both of these trials use a classic histo-pathologic design and are set to accrue all comers, and to later evaluate the correlation of response with ETS status. To date, no clinical trial has been opened to evaluate PARP inhibitors for locally advanced prostate cancer or in combination with the emerging anti-androgen therapies such as abiraterone acetate or radiation.

In contrast to the high prevalence of PARP inhibitors and PARP inhibitor trials that are currently ongoing, DNA-PKcs inhibitors with potential clinical efficacy have only begun to enter the clinic. The drug CC-115 is the first in class DNA-PKcs inhibitor, which is currently being evaluated in a phase I trial for Advanced Solid Tumors, Non-Hodgkin's Lymphoma and Multiple Myeloma (clinicaltrials.gov). While no trials have been opened to focus on DNA-PKcs inhibitors in histopathologic sites with high frequencies of ETS positive cancers, it is possible that this class of patients could be one to benefit from this type of therapy. If the initial phase I suggests that the CC-115 can inhibit DNA-PKcs activity at a tolerable dose, it is possible that emerging trials with this drug could focus on using it in combination with PARP inhibitors and/or second generation anti-androgens in subsequent phase I trials. It is important to note that in a recent study, we identified a single patient with metastatic prostate cancer that harbored a mutation in DNA-PKcs (Quist *et. al.* in review). Although it is unknown if this mutation is activating, it further suggests the importance of DNA-PKcs for prostate cancer pathogenesis.

Taken together, both PARP and DNA-PKcs are emerging therapeutic targets that are actively being pursued by a number of drug companies. These therapeutics are in clinical development and some trials have now been focused on determining efficacy in prostate cancer. While our initial studies have indicated that biomarker driven selection of patients may help predict the initial response of patients to PARP inhibitors, the clinical impact of my findings will soon be determined.

## A MOVE TO PERSONALIZED MEDICINE

From these emerging prostate cancer PARP inhibitor trials, it is interesting that the design of ongoing clinical trials still divides patients by histo-pathology rather than by molecular lesion. In one way, we do not know how any patient will actually respond to a chemotherapeutic. By taking all comers, it presents an opportunity to discover new genomic lesion-drug sensitivity combinations in order to develop additional biomarkers for any given therapy. Correspondingly, if a biomarker truly stratifies patients, this trial design generates a "control" group to determine biomarker validity. However, given the recent emergence of high throughput sequencing, one question that needs to be asked for the future is when does trial design need to be altered (if ever)? High throughput sequencing presents the opportunity for a hypothesis-guided treatment decision at any stage of progression. Without knowing how a patient will truly respond to a treatment protocol, at what point does a molecular subtype trump a histo-pathologic-driven selection?

Cancer treatment options are complex decisions based on years of research experience, histo-pathologic grading and occasionally molecular biomarkers; in the future, the hope is that therapeutic course will be less complicated and driven completely by molecular biomarkers (*87*). One recent study from our lab attempted to simplify this decision by performing personalized molecular medicine for two patients. In this study, Roychowdhury ***et. al.*** sequenced the tumors of two patients, one with metastatic colorectal cancer and a second with metastatic melanoma, both of which had bypassed all standard of care therapies

for their respective histologically-defined courses to become eligible for the study (*88*). In both cases a simple exome and transcriptome-coupled analysis, which has become standard for next generation sequencing, led to the identification of several unique mutations in each patients cancer genome. Unfortunately, no clinical trials were being run for any of the mutations discovered in either patient (*88*).

Consequently, the study highlighted some of the problems with academic attempts to pursue personalized medicine and demonstrates some of the fundamental problems the cancer research community will face in the near future. First, the field has identified very few genetic lesions that are established biomarkers of a *bona fide* clinical response. As such, few "targeted" drugs are readily available to the academic community for rapid clinical use. Second, cancer mutations appear to be of two types: those that are generally found in the tumor at high frequency and those that are seemingly random. A significant amount of biological validation needs to be completed to fully understand the consequences of these new mutations. As such, the tumor board – which makes the decision on clinical course based on all available information – is led to make a clinical decision based largely on a hypothesis for an N=1 trial. This is not necessarily a bad thing as it is a partially informed decision guided by an ethical committee (*88*), but if this is way to the future of medicine, then new methods are going to need to be developed to interpret and standardize individualized biomarker-trials for low frequency aberrations.

With this new trial design in mind, it is important to remember the experiences of using a targeted therapy called Trastuzamab, a humanized monoclonal antibody against human epidermal growth factor receptor 2 (Her2), in various cancers. *HER2* is overexpressed in 20-30% of invasive breast carcinomas (*89*), 10% of ovarian cancers (*90*) and 20% of gastric cancers (*91*) – usually through genomic amplification. As a receptor tyrosine kinase, Her2 dimerizes upon ligand binding and the subsequent tyrosine kinase activation drives downstream signaling through the RAS-MAPK pathway (reviewed in (*92*)). In cancer, Her2 can adopt a fixed conformation resembling a ligand-activated state (*93*) and high concentrations in the cell membrane due to overexpression can also be sufficient to induce dimerization (*94*). Several clinical trials in all three of these disease sites have assessed the impact of Trastuzamab in combination with current chemotherapeutic protocols. In *HER2* amplified breast cancers, Trastuzamab significantly improved the overall response rate (*95*) and, while the optimal use is still under investigation, the impact on patient care has been described as "profound" (*92*). Surprisingly, however, the impact of Trastuzamab in *HER2* amplified ovarian and gastric cancers has been relatively minor (*90, 91*) highlighting a contradictory principle that is beginning to emerge in the field of personalized medicine – while the therapies directed at specific molecular lesions hold promise for individualized therapy, clinical protocols need to be developed based on an understanding of the global mutational state and possibly disease site. In this thesis, I have not addressed the cancer stem cell model, but it is possible that individual disease sites have differences in oncogene dependency or

chemotherapeutic response due to uniqueness in stem cell populations or cells-of-origin that are not yet fully understood. Collectively, these issues serve as a reminder that histopathologic assessments should not yet be completely abandoned in the determination of personalized clinical course.

Because of the clinical experience with HER2, therefore, it is evident that personalized medicine strategies and trials will have to be carefully designed. Given the concerns that were raised by Roychowdhury *et. al.*, it is also evident that biomarker-based therapeutics research is now the rate limiting step for the development of personalized medicine rather than gross mutation detection (*88*). Furthermore, that study supports the rationale that emerging research must continue to focus on developing approaches for high frequency genomic aberrations that are known to drive cancer development and progression in order to identify clinically relevant biomarker driven therapeutic protocols. In this way, as sequencing approaches become standardized, more clinical trials will be open for other common molecular lesions to support individualized, biomarker-driven medicine. Likewise, because of the experiences with Trastuzamab, it may enable researchers to tease apart differences that may arise based on either cell lineage or genetic background.

## CONCLUSION

In an optimistic view of the future of medicine, cancer patients will come to clinic and treatment course (whether it be surgery, radiation or chemotherapy) will be determined by molecular subtype in addition to staging and histological

site. Of course this will not come without challenges such as the ability to identify the true drivers in each cancer, especially in those cancers that present with a high mutation load such as melanoma and lung, or those with genomic instability. In the near future, our understanding of the genomic landscape of cancer will improve our catalog of disruptive events that drive cancer progression will increase. This will lead to a focused set of cancer-specific targets for which the research community will need to develop clinically viable inhibitors and treatment protocols. Currently, we may already know many of the cancer-specific oncoproteins that are deregulated in very high percentages of tumors regardless of histological subtype such as ETS and cMYC transcription factors, k-RAS GTPase and a variety of kinases, etc. However, as a community, we have yet to develop methods to identify clinically viable drugs specific for many of these molecular lesions.

In this thesis, I presented a series of papers that can be logically linked in a similar hypothesis-driven manner to identify therapeutics for several common molecular lesions thought to be chemically intractable. Specifically, I focused on building treatment protocols for cancers driven by the previously intractable class of proteins called transcription factors. Using ETS transcription factors as a model system, I identified two ETS protein interacting enzymes, PARP1 and DNA-PKcs, that are critical for ETS transcription factor activity. Both of these enzymes have drugs in clinical development, some of which have already been moved into trials for prostate cancer, and will soon be moved into similar trials for other cancers driven by ETS proteins such as ESFTs. Likewise, we developed a proof-

of-principle model for a system to disrupt oncoprotein function by identify surface-binding peptides that disrupt protein:protein interactions. Taken together, by creating a viable therapeutic strategy that has already been translated to the clinic for what was considered a class of chemically intractable targets, this thesis continues to build on the work of many generations of researchers towards the ultimate goal of personalized medicine.

## REFERENCES


1. J. C. Brenner, A. M. Chinnaiyan, *Biochim Biophys Acta* **1796**, 201 (Dec, 2009).
2. B. L. Sibanda, D. Y. Chirgadze, T. L. Blundell, *Nature* **463**, 118 (Jan 7, 2010).
3. T. Berg, *Curr Opin Chem Biol* **12**, 464 (Aug, 2008).
4. S. K. Nair, S. K. Burley, *Cell* **112**, 193 (Jan 24, 2003).
5. M. Hammel *et al.*, *J Biol Chem* **285**, 1414 (Jan 8, 2011).
6. E. P. Morris *et al.*, *Nucleic Acids Res* **39**, 5757 (Jul, 2011).
7. C. Morrison *et al.*, *Nat Genet* **17**, 479 (Dec, 1997).
8. T. Ruscetti *et al.*, *J Biol Chem* **273**, 14461 (Jun 5, 1998).
9. L. Spagnolo, J. Barbeau, N. J. Curtin, E. P. Morris, L. H. Pearl, *Nucleic Acids Res* (Jan 5, 2011).
10. M. N. Paddock *et al.*, *DNA Repair (Amst)* **10**, 338 (Mar 7, 2011).
11. M. Yaneva, T. Kowalewski, M. R. Lieber, *Embo J* **16**, 5098 (Aug 15, 1997).
12. S. Choul-li, H. Drobecq, M. Aumercier, *Biochem Biophys Res Commun* **390**, 839 (Dec 18, 2009).
13. R. H. Wong *et al.*, *Cell* **136**, 1056 (Mar 20, 2009).
14. M. F. Langelier, J. L. Planck, S. Roy, J. M. Pascal, *J Biol Chem* **286**, 10690 (Mar 25, 2011).
15. K. A. Menear *et al.*, *J Med Chem* **51**, 6581 (Oct 23, 2008).
16. F. Le Page, V. Schreiber, C. Dherin, G. De Murcia, S. Boiteux, *J Biol Chem* **278**, 18471 (May 16, 2003).
17. R. H. te Poele, A. L. Okorokov, L. Jardine, J. Cummings, S. P. Joel, *Cancer Res* **62**, 1876 (Mar 15, 2002).
18. C. B. Bennett, A. L. Lewis, K. K. Baldwin, M. A. Resnick, *Proc Natl Acad Sci U S A* **90**, 5613 (Jun 15, 1993).
19. R. Chiarle *et al.*, *Cell* **147**, 107 (Sep 30, 2011).
20. I. A. Klein *et al.*, *Cell* **147**, 95 (Sep 30, 2011).
21. J. C. Brenner *et al.*, *Cancer Cell* **19**, 664 (May 17, 2011).
22. D. Dominguez-Sola *et al.*, *Nature* **448**, 445 (Jul 26, 2007).
23. D. S. Askew, R. A. Ashmun, B. C. Simmons, J. L. Cleveland, *Oncogene* **6**, 1915 (Oct, 1991).
24. G. I. Evan *et al.*, *Cell* **69**, 119 (Apr 3, 1992).
25. K. C. Brown, *Curr Pharm Des* **16**, 1040 (2010).
26. D. P. McGregor, *Curr Opin Pharmacol* **8**, 616 (Oct, 2008).
27. D. K. Malik, S. Baboota, A. Ahuja, S. Hasan, J. Ali, *Curr Drug Deliv* **4**, 141 (Apr, 2007).
28. J. P. Lalezari *et al.*, *Aids* **17**, 691 (Mar 28, 2003).
29. M. H. Kang, C. P. Reynolds, *Clin Cancer Res* **15**, 1126 (Feb 15, 2009).
30. T. Kuwana *et al.*, *Mol Cell* **17**, 525 (Feb 18, 2005).
31. J. E. Chipuk *et al.*, *Science* **303**, 1010 (Feb 13, 2004).
32. E. H. Cheng *et al.*, *Mol Cell* **8**, 705 (Sep, 2001).
33. L. D. Walensky *et al.*, *Science* **305**, 1466 (Sep 3, 2004).
34. G. H. Bird *et al.*, *Proc Natl Acad Sci U S A* **107**, 14093 (Aug 10, 2010).
35. H. E. Bryant *et al.*, *Nature* **434**, 913 (Apr 14, 2005).
36. H. Farmer *et al.*, *Nature* **434**, 917 (Apr 14, 2005).
37. D. J. Gallagher *et al.*, *Clin Cancer Res* **16**, 2115 (Apr 1, 2010).
38. A. V. Mitra *et al.*, *BJU Int* **107**, 28 (Jan, 2011).

39. I. Ibragimova, P. Cairns, *Methods Mol Biol* **780**, 277 (2011).
40. K. J. Dedes *et al.*, *Sci Transl Med* **2**, 53ra75 (Oct 13, 2010).
41. B. McEllin *et al.*, *Cancer Res* **70**, 5457 (Jul 1, 2010).
42. A. M. Mendes-Pereira *et al.*, *EMBO Mol Med* **1**, 315 (Sep, 2009).
43. D. Sarker, A. H. Reid, T. A. Yap, J. S. de Bono, *Clin Cancer Res* **15**, 4799 (Aug 1, 2009).
44. X. Jiang, S. Chen, J. M. Asara, S. P. Balk, *J Biol Chem* **285**, 14980 (May 14, 2010).
45. M. D. Forster *et al.*, *Nat Rev Clin Oncol* **8**, 302 (May, 2011).
46. M. Fraser *et al.*, *Clin Cancer Res* **18**, 1015 (Feb 15, 2011).
47. J. M. de Murcia *et al.*, *Proc Natl Acad Sci U S A* **94**, 7303 (Jul 8, 1997).
48. Z. Q. Wang *et al.*, *Genes Dev* **9**, 509 (Mar 1, 1995).
49. B. Evers *et al.*, *Clin Cancer Res* **14**, 3916 (Jun 15, 2008).
50. T. Hay *et al.*, *Cancer Res* **69**, 3850 (May 1, 2009).
51. S. Rottenberg *et al.*, *Proc Natl Acad Sci U S A* **105**, 17079 (Nov 4, 2008).
52. U. Kortmann *et al.*, *Clin Cancer Res* **17**, 783 (Feb 15, 2011).
53. T. D. Penning *et al.*, *Bioorg Med Chem* **16**, 6965 (Jul 15, 2008).
54. T. D. Penning *et al.*, *J Med Chem* **52**, 514 (Jan 22, 2009).
55. T. Karlberg, M. Hammarstrom, P. Schutz, L. Svensson, H. Schuler, *Biochemistry* **49**, 1056 (Feb 16, 2010).
56. M. J. Clarke *et al.*, *Mol Cancer Ther* **8**, 407 (Feb, 2009).
57. T. M. Horton *et al.*, *Mol Cancer Ther* **8**, 2232 (Aug, 2009).
58. X. Liu *et al.*, *Mol Cancer Res* **6**, 1621 (Oct, 2008).
59. J. P. Palma *et al.*, *Clin Cancer Res* **15**, 7277 (Dec 1, 2009).
60. S. Kummar *et al.*, *Cancer Res* **71**, 5626 (Sep 1, 2011).
61. P. C. Fong *et al.*, *N Engl J Med* **361**, 123 (Jul 9, 2009).
62. E. Dean *et al.*, *Br J Cancer* **106**, 468 (Jan 31, 2012).
63. P. C. Fong *et al.*, *J Clin Oncol* **28**, 2512 (May 20, 2010).
64. K. A. Gelmon *et al.*, *Lancet Oncol* **12**, 852 (Sep, 2011).
65. S. B. Kaye *et al.*, *J Clin Oncol* **30**, 372 (Feb 1, 2012).
66. A. Tutt *et al.*, *Lancet* **376**, 235 (Jul 24, 2010).
67. N. Yamamoto *et al.*, *Cancer Sci* (Dec 7, 2011).
68. M. W. Audeh *et al.*, *Lancet* **376**, 245 (Jul 24, 2010).
69. J. Samol *et al.*, *Invest New Drugs* (May 18, 2011).
70. O. A. Khan *et al.*, *Br J Cancer* **104**, 750 (Mar 1, 2011).
71. J. C. Barreto-Andrade *et al.*, *Mol Cancer Ther* **10**, 1185 (Jul, 2011).
72. E. V. Efimova *et al.*, *Cancer Res* **70**, 6277 (Aug 1, 2010).
73. D. Nguyen *et al.*, *Cell Cycle* **10**, 4074 (Dec 1, 2011).
74. S. Nowsheen, J. A. Bonner, E. S. Yang, *Radiother Oncol* **99**, 331 (Jun, 2011).
75. J. M. Senra *et al.*, *Mol Cancer Ther* **10**, 1949 (Oct, 2011).
76. M. Guha, *Nat Biotechnol* **29**, 373 (May, 2011).
77. J. O'Shaughnessy *et al.*, *N Engl J Med* **364**, 205 (Jan 20, 2011).
78. X. Liu *et al.*, *Clin Cancer Res* **18**, 510 (Jan 15, 2012).
79. N. Issaeva *et al.*, *Cancer Res* **70**, 6268 (Aug 1, 2010).
80. S. L. Edwards *et al.*, *Nature* **451**, 1111 (Feb 28, 2008).
81. W. Sakai *et al.*, *Nature* **451**, 1116 (Feb 28, 2008).
82. X. Liu *et al.*, *Mol Cancer Res* **7**, 1686 (Oct, 2009).
83. C. A. Lee, J. A. Cook, E. L. Reyner, D. A. Smith, *Expert Opin Drug Metab Toxicol* **6**, 603 (May, 2010).

84. L. Rossi, M. Denegri, M. Torti, G. G. Poirier, A. Ivana Scovassi, *Biochimie* **84**, 1229 (Dec, 2002).
85. I. E. Dumitriu *et al.*, *Cell Death Differ* **11**, 314 (Mar, 2004).
86. R. A'Hern, DeBono J., Sandhu S, Kalaitzaki E, Usdin M, Hall EE, *Trials* **12**, A88 (October 4, 2011, 2011).
87. I. J. Majewski, R. Bernards, *Nat Med* **17**, 304 (Mar, 2011).
88. S. Roychowdhury *et al.*, *Sci Transl Med* **3**, 111ra121 (Nov 30, 2011).
89. D. J. Slamon *et al.*, *Science* **244**, 707 (May 12, 1989).
90. M. A. Bookman, K. M. Darcy, D. Clarke-Pearson, R. A. Boothby, I. R. Horowitz, *J Clin Oncol* **21**, 283 (Jan 15, 2003).
91. Y. J. Bang *et al.*, *Lancet* **376**, 687 (Aug 28, 2010).
92. C. A. Hudis, *N Engl J Med* **357**, 39 (Jul 5, 2007).
93. H. S. Cho *et al.*, *Nature* **421**, 756 (Feb 13, 2003).
94. P. J. Brennan, T. Kumagai, A. Berezov, R. Murali, M. I. Greene, *Oncogene* **19**, 6093 (Dec 11, 2000).
95. N. Robert *et al.*, *J Clin Oncol* **24**, 2786 (Jun 20, 2006).

# APPENDIX I

Oligonucleotide Primers

| Assay | Gene/Region | Sequence 5' to 3' |
|---|---|---|
| Expression qPCR | DNALIG4 | AGCTGGGATTCTCTGGTTCA |
| Expression qPCR | DNALIG4 | TGCAAAAGGAACGTGAGATG |
| Expression qPCR | PARP1 | GCGTGAAGGCGAATGCCAGC |
| Expression qPCR | PARP1 | TGTAGCCTGTCACGGGCGCT |
| Expression qPCR | XRCC4 | CCTCAGGAGAATCAGCTTCAA |
| Expression qPCR | XRCC4 | GTCTTCTGGGCTGCTGTTTC |
| Expression qPCR | ATM | GCTGTGAGAAAACCATGGAAG |
| Expression qPCR | ATM | AGTTTCATCTTCCGGCCTCT |
| Expression qPCR | ATR | AAGCGCCACTGAATGAAACT |
| Expression qPCR | ATR | AACGGCAGTCCTGTCACTCT |
| Expression qPCR | DNA-PKcs | TGCCAATCCAGCAGTCATTA |
| Expression qPCR | DNA-PKcs | TGAAAGCCCACTCTCTGGTT |
| Expression qPCR | *ERG* | CGCAGAGTTATCGTGCCAGCAGAT |
| Expression qPCR | *ERG* | CCATATTCTTTCACCGCCCACTCC |
| Expression qPCR | *ETV1* | ATAGCAGCTACCCCATGGAC |
| Expression qPCR | *ETV1* | TCAGACATCTGGCGTTGGTA |
| Expression qPCR | *EZH2* | AGAATGGAAACAGCGAAGGA |
| Expression qPCR | *EZH2* | TCAATGAAAGTACCATCCTGATCT |
| Expression qPCR | *BRCA2* | GCCGTACACTGCTCAAATCA |
| Expression qPCR | *BRCA2* | TTTGAAGTCATCTGGGCTGA |
| Expression qPCR | *DNMT3a* | CACCGGCCATACGGTGGAGC |
| Expression qPCR | *DNMT3a* | TCGTGGTCTTTGGAGGCGAGAGT |
| Expression qPCR | *ZNF100* | TGGAGGAGTGGCAATGCCTGGA |
| Expression qPCR | *ZNF100* | TGGCTTAGTGAGAGCAATACCTGCCAA |
| Expression qPCR | *PBX1* | AGCCGGACCAGGCCCATCTC |
| Expression qPCR | *PBX1* | CTTCCGCCGCGCATCCAGAA |
| Expression qPCR | *ZNF618* | CTCCTCCCGCACGGACCCAT |
| Expression qPCR | *ZNF618* | TTCAAGCGCTCCCTGCTCGC |
| Expression qPCR | *KLK3 (PSA)* | GAGCACCCCTATCAACCCCCTATT |
| Expression qPCR | *KLK3 (PSA)* | AGCAACCCTGGACCTCACACCTAA |
| Expression qPCR | *β-Actin* | CGCGAGAAGATGACCCAGAT |
| Expression qPCR | *β-Actin* | GAGTCCATCACGATGCCAGT |
| Expression qPCR | *ACCN1* | TTCGAGGGAATCAGCGCCGC |
| Expression qPCR | *ACCN1* | TAAACACGGAGGAGAAGTTGTGCG |
| Expression qPCR | *TCF12* | CGGCAGAAGTGGCCGAAGATGAA |

| | | |
|---|---|---|
| Expression qPCR | *TCF12* | GGGAAAACATCGCACTGAAGTCCA |
| Expression qPCR | *PFTK1* | CCAGCTCGCCAACAAGTCCCAA |
| Expression qPCR | *PFTK1* | CCTGCAGCCTGATCACCTTCAGAG |
| Expression qPCR | *MK-STYX* | AGGGAGCCTGGCCATAGTCGC |
| Expression qPCR | *MK-STYX* | AGCAAACCAGGCATCCTTTCCG |
| Expression qPCR | *CDK7* | TGGAGTCGGGCTTTACGGCG |
| Expression qPCR | *CDK7* | AAACGGTGGCAAACTGTCCCTCC |
| ChIP-qPCR | *PLA1A* | TGGCCACCCAGAGATGCAGGA |
| ChIP-qPCR | *PLA1A* | ACACACTGTCCCTCTTTGAGCCA |
| ChIP-qPCR | *TMPRSS2* | TGGAGCTAGTGCTGCATGTC |
| ChIP-qPCR | *TMPRSS2* | CTGCCTTGCTGTGTGAAAAA |
| ChIP-qPCR | *FKBP5* | GGTTCCTGGGCAGGAGTAAG |
| ChIP-qPCR | *FKBP5* | AACGTGGATCCCACACTCTC |
| ChIP-qPCR | *KLK3 (PSA)* | CCTAGATGAAGTCTCCATGAGCTACA |
| ChIP-qPCR | *KLK3 (PSA)* | GGGAGGGAGAGCTAGCACTTG |

# APPENDIX II

Genes Differentially Regulated by PARP1 or DNA-PKcs siRNA

| Gene ID | PARP1 | DNA-PKcs |
|---|---|---|
| MGA | 0.264462 | 0.4686345 |
| TBX15 | 0.270868 | 0.2553663 |
| STARD4 | 0.303519 | 0.3408556 |
| SUV39H2 | 0.307705 | 0.4381685 |
| CD556746 | 0.33708 | 0.3847578 |
| KITLG | 0.337534 | 0.3872908 |
| NEIL2 | 0.364262 | 0.4383824 |
| LAMA1 | 0.380545 | 0.3314308 |
| AOC3 | 0.381653 | 0.4914786 |
| E2F8 | 0.392452 | 0.4470068 |
| HPGD | 0.395736 | 0.4250418 |
| CCT8 | 0.401878 | 0.3645652 |
| TDRD9 | 0.408429 | 0.4014726 |
| BTG3 | 0.415805 | 0.4533361 |
| FAM49B | 0.415849 | 0.48009 |
| C2orf14 | 0.420487 | 0.4325688 |
| AA631847 | 0.420546 | 0.4824154 |
| LOC641467 | 0.4296 | 0.4994357 |
| LOC440295 | 0.43172 | 0.3096605 |
| SEMA3C | 0.432329 | 0.4787954 |
| MTHFD2 | 0.435626 | 0.4779902 |
| GNB4 | 0.442082 | 0.2707471 |
| ZC3HAV1L | 0.44313 | 0.4022851 |
| CDCA7 | 0.443146 | 0.3559481 |
| C6orf167 | 0.444557 | 0.352313 |
| C14orf151 | 0.445016 | 0.4596204 |
| KCNJ2 | 0.447251 | 0.4597726 |
| TTTY15 | 0.447368 | 0.4376759 |
| DNMT3A | 0.447503 | 0.3743971 |
| RBL1 | 0.451895 | 0.4610096 |
| CYCS | 0.453801 | 0.3872844 |
| RUNDC2B | 0.454633 | 0.4509508 |
| MARS2 | 0.454741 | 0.472233 |
| SEH1L | 0.459784 | 0.4252461 |
| ZNF100 | 0.460427 | 0.3834184 |
| ARNTL2 | 0.460437 | 0.4956873 |
| EZH2 | 0.461723 | 0.3947172 |
| ABCB10 | 0.466043 | 0.458422 |
| SFRS1 | 0.4705 | 0.43528 |

| | | |
|---|---|---|
| ATAD2 | 0.472069 | 0.3675277 |
| KLHL2 | 0.472132 | 0.4781275 |
| RIF1 | 0.481235 | 0.4771625 |
| A_32_P108420 | 0.481256 | 0.4600318 |
| DLEU2 | 0.484133 | 0.3488587 |
| FAM29A | 0.484236 | 0.4372668 |
| MNS1 | 0.487367 | 0.3852707 |
| PFKFB2 | 0.493362 | 0.4422727 |
| ARHGAP11A | 0.496593 | 0.498053 |
| DKFZP586B0319 | 0.496614 | 0.3504658 |
| CENPK | 0.497873 | 0.4569846 |
| CDCA7 | 0.499617 | 0.3733235 |
| PTPRT | 7.700304 | 12.388148 |
| WNK4 | 5.192412 | 4.9384666 |
| ART3 | 4.974834 | 7.1878289 |
| SLC7A14 | 4.878306 | 6.2776191 |
| ZNF536 | 4.772624 | 4.533891 |
| FOS | 4.619362 | 6.8309435 |
| ANKRD38 | 4.586529 | 3.7157074 |
| PTPRE | 4.220544 | 7.2166772 |
| FAM135B | 4.171088 | 4.0039167 |
| SATB1 | 3.853212 | 5.2788281 |
| CXCL11 | 3.834556 | 6.0433518 |
| AI674800 | 3.774656 | 3.3599359 |
| PCDH19 | 3.726251 | 5.6496715 |
| CLIP3 | 3.592404 | 4.2223238 |
| SNAP25 | 3.586151 | 3.459766 |
| DGKG | 3.560303 | 3.7277485 |
| OR51E1 | 3.552674 | 3.6436745 |
| MDK | 3.533256 | 4.051834 |
| TFF3 | 3.4889 | 5.8561434 |
| PGF | 3.414377 | 4.0232471 |
| WNK4 | 3.393173 | 3.602689 |
| BMX | 3.379299 | 5.0145222 |
| RAB42 | 3.338727 | 3.8161748 |
| C21orf88 | 3.330321 | 4.6726438 |
| GUCA2A | 3.322561 | 5.8424265 |
| GPR120 | 3.243831 | 3.2216067 |
| SOX30 | 3.233005 | 3.1596345 |
| KCNB1 | 3.198907 | 5.300129 |
| YPEL3 | 3.183858 | 3.732482 |
| DIO3 | 3.13993 | 3.0791304 |
| MYO1G | 3.103429 | 2.9154131 |
| CD200 | 3.097168 | 3.119397 |
| PTCHD2 | 3.058486 | 2.0607478 |
| C9orf4 | 3.025913 | 2.8213145 |
| CUEDC1 | 3.01445 | 3.4799769 |
| BMX | 2.976712 | 5.6889022 |

| | | |
|---|---|---|
| FAM14A | 2.95608 | 5.9476405 |
| SLFN5 | 2.955286 | 3.7314277 |
| C1QTNF1 | 2.906643 | 3.2318844 |
| LOC202134 | 2.870954 | 3.1441018 |
| FOXN2 | 2.861073 | 2.841886 |
| DENND2A | 2.855316 | 3.2996164 |
| LOC647502 | 2.818921 | 2.3359048 |
| C12orf51 | 2.811883 | 2.5676902 |
| NTF5 | 2.809855 | 2.9490018 |
| Kua-UEV | 2.805097 | 2.8114755 |
| SLC2A10 | 2.793324 | 2.81208 |
| NEURL | 2.781188 | 3.5854038 |
| C20orf74 | 2.77641 | 2.7811885 |
| TCN2 | 2.769731 | 2.4674434 |
| KLHL24 | 2.768599 | 2.7073712 |
| CXCL11 | 2.762173 | 4.6493967 |
| MRAS | 2.746989 | 4.8921041 |
| PTPRT | 2.711533 | 2.955711 |
| KREMEN1 | 2.708394 | 2.0486227 |
| ABCC3 | 2.707233 | 6.689315 |
| LOC730091 | 2.701996 | 5.8574316 |
| DDX17 | 2.6961 | 2.1963041 |
| NAP5 | 2.695599 | 3.4332303 |
| HSD17B2 | 2.687121 | 5.0416176 |
| GRIA2 | 2.686483 | 2.9473426 |
| ARRDC3 | 2.681253 | 2.053636 |
| KLC4 | 2.680831 | 2.9301589 |
| TMEM30B | 2.668045 | 2.090522 |
| GRK4 | 2.662371 | 4.4908928 |
| TMEM169 | 2.650214 | 5.5549314 |
| KIAA1853 | 2.649369 | 3.2768672 |
| CIB2 | 2.644662 | 3.3931109 |
| HRNBP3 | 2.641761 | 2.0605058 |
| PPFIA2 | 2.632261 | 4.2412463 |
| ANGPTL2 | 2.629201 | 2.6510222 |
| KIAA0152 | 2.6289 | 2.8886974 |
| RASSF4 | 2.627694 | 3.3216571 |
| TPCN1 | 2.62708 | 2.2936089 |
| SLC9A3 | 2.606581 | 2.2781992 |
| IGFBP7 | 2.605145 | 5.4180489 |
| RC3H1 | 2.602608 | 2.6797909 |
| SLC4A4 | 2.596962 | 2.7903569 |
| FSTL4 | 2.586459 | 3.4240408 |
| KCNC2 | 2.573538 | 3.1137801 |
| C15orf48 | 2.569925 | 3.6820703 |
| TLN2 | 2.566908 | 2.402091 |
| KIAA1946 | 2.558402 | 2.2126587 |
| ABTB1 | 2.556338 | 2.5252223 |

| | | |
|---|---|---|
| LIX1 | 2.554647 | 4.473526 |
| KIAA0319L | 2.540977 | 2.6312192 |
| NY-REN-7 | 2.5143 | 2.7164223 |
| RPL10 | 2.508925 | 2.390845 |
| AMT | 2.504305 | 2.0054397 |
| LOXL4 | 2.498175 | 2.8813389 |
| SHC2 | 2.493786 | 2.9600594 |
| SV2B | 2.490237 | 4.4659351 |
| HAB1 | 2.486546 | 2.5995588 |
| IGFBP6 | 2.486081 | 2.8296924 |
| LGICZ1 | 2.478803 | 2.4044603 |
| HIP1 | 2.475584 | 2.2487108 |
| NAV1 | 2.466657 | 2.4622555 |
| ABTB1 | 2.458763 | 2.4409994 |
| HADHA | 2.457656 | 2.1974669 |
| CCDC3 | 2.454739 | 2.7294989 |
| PCMTD1 | 2.454076 | 2.4330002 |
| C20orf102 | 2.441737 | 3.2191623 |
| KIAA1212 | 2.440569 | 3.2867962 |
| HOXC8 | 2.44031 | 2.5886973 |
| SNIP | 2.439988 | 2.8650915 |
| FOXO4 | 2.438701 | 2.8365793 |
| AMOTL1 | 2.436569 | 2.1208436 |
| EXT1 | 2.435215 | 2.2079491 |
| LOC642852 | 2.434432 | 2.7085603 |
| NEFM | 2.434017 | 2.9892428 |
| FCGRT | 2.431673 | 2.6428801 |
| SLC24A3 | 2.422566 | 4.0182544 |
| NAP5 | 2.412048 | 3.5940285 |
| CLSTN3 | 2.407237 | 2.082551 |
| HRASLS3 | 2.399854 | 2.5334957 |
| GRIA2 | 2.397384 | 3.6523105 |
| MALAT1 | 2.396116 | 3.1975703 |
| LPP | 2.3922 | 2.9491364 |
| PHF2 | 2.392137 | 2.2825612 |
| EIF2AK2 | 2.389473 | 2.0270739 |
| CTGF | 2.387586 | 3.991631 |
| SH3KBP1 | 2.382395 | 3.6211105 |
| TSPAN1 | 2.37654 | 2.3963208 |
| NLRP1 | 2.364919 | 4.2066233 |
| SCN8A | 2.355579 | 2.0300312 |
| MSRB3 | 2.351674 | 2.4471562 |
| PRSS2 | 2.348646 | 3.129274 |
| ERC1 | 2.332685 | 2.1971823 |
| SNIP | 2.328384 | 3.2670252 |
| MYST4 | 2.325062 | 2.3276905 |
| MYO5A | 2.320072 | 2.1964059 |
| RYR1 | 2.318346 | 2.73224 |

| | | |
|---|---|---|
| PBX1 | 2.318004 | 3.0127011 |
| PCMTD1 | 2.31746 | 2.2801816 |
| VANGL1 | 2.316719 | 2.5121505 |
| CASC4 | 2.316537 | 2.181143 |
| MYO1G | 2.316142 | 2.851615 |
| Kua-UEV | 2.313308 | 2.3668857 |
| KIF12 | 2.307216 | 2.4089092 |
| MAP3K10 | 2.303609 | 2.5190009 |
| JAM3 | 2.303427 | 3.0948891 |
| VTN | 2.300979 | 3.3706898 |
| TOX | 2.300881 | 2.5447356 |
| KCNK13 | 2.29587 | 2.5556048 |
| HSPG2 | 2.28488 | 2.800787 |
| PLEKHB1 | 2.280364 | 3.5807866 |
| ENO2 | 2.278387 | 2.2340388 |
| CPM | 2.276534 | 3.2787182 |
| C21orf34 | 2.275674 | 2.7871111 |
| MIER1 | 2.274298 | 2.093848 |
| CCDC125 | 2.270932 | 2.0925212 |
| SCN4B | 2.262779 | 2.9434376 |
| CDH22 | 2.258926 | 4.6378437 |
| ITM2A | 2.252059 | 3.074315 |
| LOC157860 | 2.250772 | 3.1568912 |
| SERPINA1 | 2.250485 | 2.379129 |
| TMEM45A | 2.245967 | 3.2042487 |
| CCDC113 | 2.245396 | 3.0730312 |
| ADPRH | 2.24524 | 2.279578 |
| C20orf23 | 2.242991 | 2.4382029 |
| CACNG6 | 2.239689 | 3.3737924 |
| S100A6 | 2.239337 | 2.9262959 |
| PRSS2 | 2.239243 | 2.9429149 |
| MB | 2.236818 | 2.7770793 |
| TMC8 | 2.23542 | 2.6788409 |
| PDE6D | 2.230405 | 2.4508248 |
| PIK3AP1 | 2.228892 | 2.3159561 |
| VASP | 2.228652 | 2.4766343 |
| SLC13A3 | 2.2285 | 3.5557934 |
| COL6A2 | 2.22705 | 2.0120694 |
| FLOT2 | 2.226659 | 2.3473858 |
| DUSP6 | 2.224399 | 3.6011119 |
| SVOPL | 2.222723 | 4.6732694 |
| ERG | 2.22227 | 2.2982431 |
| NEURL | 2.222052 | 2.7925728 |
| KCNJ3 | 2.218304 | 3.4604556 |
| ZNF467 | 2.214526 | 2.1368629 |
| DLGAP4 | 2.214316 | 2.2843229 |
| LOC253264 | 2.213666 | 2.8022558 |
| ACPP | 2.212993 | 2.784264 |

| | | |
|---|---|---|
| LY6G5C | 2.212215 | 2.2579312 |
| FBXO11 | 2.204666 | 2.4251787 |
| ACPP | 2.203873 | 2.2053424 |
| LOC388242 | 2.202646 | 2.8128431 |
| YPEL2 | 2.198153 | 3.1780016 |
| PGPEP1 | 2.197898 | 3.0644878 |
| GOLPH4 | 2.197794 | 2.340983 |
| PHF2 | 2.194053 | 2.1648568 |
| NACAD | 2.192686 | 3.8865518 |
| COL14A1 | 2.192002 | 4.3514795 |
| RGPD5 | 2.191394 | 2.3685578 |
| FLJ31356 | 2.190107 | 3.2847933 |
| ACTA2 | 2.188752 | 3.4232389 |
| LPIN2 | 2.188006 | 2.5164606 |
| ZNF658 | 2.184715 | 2.2196328 |
| TPCN1 | 2.178778 | 2.4960427 |
| KCNK3 | 2.177352 | 2.2995379 |
| GPR120 | 2.173727 | 2.2663394 |
| NOV | 2.172924 | 2.5451494 |
| RALB | 2.172864 | 2.158878 |
| LIG3 | 2.165928 | 2.1756869 |
| SELV | 2.165579 | 2.8841506 |
| MMP7 | 2.162323 | 3.6641495 |
| WDR78 | 2.162309 | 2.1997904 |
| SLC18A1 | 2.162184 | 2.155281 |
| SHE | 2.162113 | 2.1899338 |
| OTUD5 | 2.160873 | 2.0160938 |
| GPR161 | 2.160175 | 2.0487018 |
| ENC1 | 2.156227 | 2.6534892 |
| EMID1 | 2.153495 | 2.611273 |
| H6PD | 2.150518 | 2.1848243 |
| WASF2 | 2.148445 | 2.0942485 |
| TSPAN2 | 2.146796 | 2.7472294 |
| NPAL2 | 2.145942 | 2.3712818 |
| PSD3 | 2.1444 | 2.1119799 |
| BTC | 2.142341 | 3.4224326 |
| SPTBN4 | 2.141977 | 3.2597521 |
| SLC43A2 | 2.137532 | 2.4837827 |
| COPZ2 | 2.137326 | 2.4567599 |
| CCDC88 | 2.136328 | 2.4978851 |
| CDH7 | 2.131872 | 2.4943887 |
| FBXO27 | 2.127484 | 2.0806225 |
| MEA1 | 2.125923 | 2.0836844 |
| KCTD16 | 2.125459 | 2.4736287 |
| GPR143 | 2.123931 | 2.2037603 |
| NMNAT2 | 2.119644 | 2.0657551 |
| ZDHHC17 | 2.116382 | 2.1737883 |
| ZNF618 | 2.115425 | 2.2267129 |

| | | |
|---|---|---|
| SLC35F3 | 2.11274 | 2.6318907 |
| LIMD2 | 2.110468 | 2.4427887 |
| SVOP | 2.109529 | 2.8322151 |
| BACE1 | 2.107283 | 2.6119843 |
| ERC2 | 2.102507 | 2.9285853 |
| CHRDL2 | 2.102148 | 2.2867542 |
| ABLIM3 | 2.101433 | 2.2550255 |
| APCDD1 | 2.099848 | 2.7592146 |
| TMEM16B | 2.099184 | 3.1659446 |
| NPR2 | 2.098484 | 2.5732691 |
| PALLD | 2.091088 | 3.417215 |
| KIAA1913 | 2.088103 | 2.4711191 |
| BCMO1 | 2.086638 | 2.2911382 |
| NEU4 | 2.085766 | 2.9678593 |
| SORBS2 | 2.085169 | 2.307659 |
| APOD | 2.084254 | 2.3765494 |
| PTPRE | 2.08137 | 4.0419887 |
| GDAP1L1 | 2.0796 | 2.7306004 |
| TNFSF15 | 2.074546 | 2.3445837 |
| RAB3C | 2.072999 | 3.0342 |
| RAPGEF1 | 2.068393 | 2.2191448 |
| KIAA0152 | 2.066501 | 2.1955042 |
| ZNF776 | 2.064925 | 2.257451 |
| RENBP | 2.063194 | 2.769438 |
| LOC439914 | 2.056252 | 2.9290694 |
| FLOT2 | 2.054778 | 2.1789271 |
| MCAM | 2.05444 | 2.3928049 |
| B3GALT4 | 2.05037 | 3.3795742 |
| B3GALT4 | 2.048325 | 3.1843139 |
| LIM2 | 2.047933 | 2.2296261 |
| CRABP2 | 2.044886 | 2.2886842 |
| PRRC1 | 2.044645 | 2.1034513 |
| SCARNA17 | 2.044009 | 2.3704371 |
| CDRT4 | 2.043087 | 4.7420169 |
| ACIN1 | 2.03952 | 2.01136 |
| CCDC88 | 2.035789 | 2.5555261 |
| RBP1 | 2.035457 | 2.5229638 |
| ZNF704 | 2.034631 | 2.2145682 |
| ChGn | 2.03149 | 2.2721582 |
| FNDC4 | 2.028601 | 2.2236075 |
| C20orf194 | 2.027046 | 2.6911343 |
| TNNC2 | 2.026079 | 2.3617158 |
| AMT | 2.025644 | 2.2535052 |
| UFM1 | 2.025116 | 2.1345964 |
| LOXL4 | 2.020069 | 2.0116867 |
| CXCL9 | 2.019576 | 3.668813 |
| SFTPD | 2.019477 | 3.1270064 |
| HIP1 | 2.019431 | 2.5213178 |

| | | |
|---|---|---|
| SMARCD3 | 2.015494 | 2.62263 |
| NDRG2 | 2.013154 | 2.3490005 |
| PRSS3 | 2.011519 | 2.5113238 |
| RNF152 | 2.008668 | 2.9526871 |
| MAP1LC3A | 2.008225 | 2.3188173 |
| ICAM1 | 2.003071 | 2.8082907 |