Development of ERG Inhibitors as Potential Drugs for the Treatment of Metastatic

Prostate Cancer

by

Mani Roshan-Moniri

B.Sc., The University of British Columbia, 2008
M.Sc., The University of British Columbia, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Experimental Medicine)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

February 2021

© Mani Roshan-Moniri, 2021
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Development of ERG Inhibitors as Potential Drugs for the Treatment of Metastatic Prostate Cancer

submitted by Mani Roshan-Moniri in partial fulfilment of the requirements the degree of DOCTOR OF PHILOSOPHY in Experimental Medicine

Examing Committee:

Dr. Artem Cherkasov, Department of Urologic Sciences
Supervisor

Dr. Michael Cox, Department of Urologic Sciences
Supervisory Committee Member

Dr. William Lockwood, Department of Pathology and Laboratory Medicine
University Examiner

Dr. Jörg Gsponer, Department of Biochemistry and Molecular Biology
University Examiner

Additional Supervisory Committee Members:

Dr. Ralph Buttyan, Urologic Sciences
Supervisory Committee Member

Dr. Lawrence McIntosh, Biochemistry and Molecular Biology
Supervisory Committee Member
Abstract

Prostate cancer is one of the leading causes of cancer-related death in men worldwide. If diagnosed early, prostate cancer can be treated by surgery and/or radiotherapy. In cases where the cancer has returned or is more aggressive and has metastasized, hormone therapy is the standard treatment. While initially effective, resistance to hormone therapy often occurs. Therefore, there is a pressing demand for new therapeutics to be developed to treat this disease. Previous studies have established that in up to 50% of all prostate cancer cases, a genomic irregularity involving the ETS-related gene (ERG) is present. This alteration results in the aberrant production of predominantly amino-terminal truncated ERG proteins in the prostate where it is linked to disease development and progression. This thesis tested the hypothesis that direct, small molecule targeting of ERG DNA binding could result in inhibition of the metastatic potential of PCa through the following specific aims: a) develop and apply in vitro assays to validate inhibitory activities/mechanisms of lead anti-ERG compounds, and b) determine the therapeutic effects of the lead compounds based on their effects and activity in in vivo xenograft models. The results demonstrate the direct binding of a novel small molecule, VPC-18005, with the ERG-ETS domain using biophysical approaches. This was further supported by reduced migration and invasion rates of ERG expressing prostate cancer cells, and reduced metastasis in a zebrafish xenograft model following exposure to VPC-18005. These results support the concept that small molecules targeting the ERG-ETS domain that suppress transcriptional activity and reverse transformed characteristics of prostate cancers aberrantly expressing ERG can be developed. It is hoped that these approaches might lead to identification of small molecules that can be further developed as drug candidates as alternatives to, or in combination with, current therapies for prostate cancer patients harboring ERG fusions.
Lay Summary

Previous studies have established that in about 50% of all prostate cancer cases a genomic irregularity involving the ETS-related gene ERG is present. This abnormality is thought to occur early in development of the disease and may be part of the reason prostate cancers can invade other organs and become lethal. Evidence to date suggests that when ERG is made in prostate cells, it can give the cells the ability to migrate away from the prostate gland. This type of cell migration, called metastasis, is what allows tumors to invade other organs and kill patients. There are as yet no therapies that target ERG. This thesis has identified novel drug candidates that target ERG and have validated their activity in different laboratory assays. Targeting ERG is a major unmet opportunity for new prostate cancer therapies for the half of prostate cancer patients whose cancers carry these ERG genetic abnormalities.
Preface

- In chapter 1, section 1.2, a slightly different version of this has been previously published: Dalal K, Moniri MR, Sharma A, Li H, Ban F, Hessien M, Hsing M, Singh K, LeBlanc E, Dehm S, Cherkasov A, Rennie PS. Selectively Targeting the DNA Binding Domain of the Androgen Receptor as a Prospective Therapy for Prostate Cancer. J Biol Chem. 2014 Sep 19;289(38):26417-29. KD (70%) and MRM (30%) designed and performed the experiments, derived the models, and analyzed the data. KD (80%) and MRM (20%) wrote the manuscript with support from EL, SD, AC and PR. LH and FB performed the computational modeling. AS and KS provided additional experimental support. EL, SD, AC and PR helped supervise the project and provided funding.

- A slightly reformatted version of chapter 2 has been previously published: Miriam S. Butler*, Mani Roshan-Moniri*, Michael Hsing*, Desmond K.W. Lau*, Ari Kim, Paul Yen, Marta Mroczek, Mannan Nouri, Scott Lien, Peter Axerio-Cilies, Kush Dalal, Clement Yau, Fariba Ghaidi, Yubin Guo, Takeshi Yamazaki, Sam Lawn, Martin E. Gleave, Cheryl Y. Gregory-Evans, Lawrence P. McIntosh, Michael E. Cox, Paul S. Rennie, and Artem Cherkasov. Discovery and characterization of small molecules targeting the DNA-binding ETS domain of ERG in prostate cancer. Oncotarget. 2017 Jun 27;8(26):42438-42454. (* co-first authors). MRM (40%), MSB (30%), MH (30%) devised the project and the main conceptual ideas. MSB (40%), MRM (45%), MH (5%) and DL (10%) designed and performed the experiments, derived the models, and analyzed the data as follows: MH (assisted by TY) performed the computational modeling, DL performed the physical chemistry experiments, MSB and MRM performed in vitro experiments, and MRM designed and performed in vivo studies with assistance from SL, FG, and CL. AK, PY, MM, MN, SL, PAC, KD, CY, FG, YG, SL provided additional experimental support. MSB
(40%), MRM (40%), and MH (20%) wrote the manuscript with editorial support from LPM and MEC. MEG, CYG, LPM, MEC, PR, AC helped supervise the project and provided funding support.

- All chapters in the thesis were written by me. The results described throughout this thesis is based on work conducted at Vancouver Prostate Centre. I was responsible for carrying out the in vitro/in vivo experiments with the supervision and guidance of Dr. Butler, except for the NMR assays that were performed by Dr. Lau and the in silico work that was carried out by Dr. Hsing. The in silico work was supervised by Dr. Cherkasov and the in vitro/in vivo studies by Drs. Cox and Rennie.

- This thesis was conducted according to the guidelines approved by the UBC Biosafety committee (B13-0100) and Animal Care Centre (ACC) Ethics committee: A11-0337 (pharmacokinetics and toxicity), A11-0275 (kidney implants), A14-0006 (xenograft efficacy); and my ACC certificates: RBH-994-10 (ACC’s Rodent Biology and Husbandry), RA-548-10 (ACC’s Rodent Anesthesia), RSx-376-10 (ACC’s Principles or Rodent Surgery).
Table of Contents

Abstract ........................................................................................................................................ iii
Lay Summary ................................................................................................................................ iv
Preface ........................................................................................................................................ v
Table of Contents ........................................................................................................................ vii
List of Tables ................................................................................................................................ xi
List of Figures ............................................................................................................................... xii
List of Abbreviations .................................................................................................................... xiv
Acknowledgments ........................................................................................................................ xviii
Dedication ...................................................................................................................................... xxi

Chapter 1: Introduction ................................................................................................................. 1

1.1 The Prostate .............................................................................................................................. 2
1.2 Androgen Signaling in the Prostate ............................................................................................ 4
1.3 Prostate Cancer: diagnosis and current therapies ................................................................. 7
1.4 Management of metastatic complications in prostate cancer ............................................ 8
  1.4.1 Castration ............................................................................................................................. 9
  1.4.2 Anti-androgens .................................................................................................................. 9
1.5 Metastatic castration-resistant prostate cancer ...................................................................... 9
1.6 Changing patterns of metastatic CRPC ................................................................................ 12
1.7 Metastasis and cellular plasticity ............................................................................................ 15
  1.7.1 Targeting cellular plasticity to improve patient outcome ............................................... 16
1.8 ERG expression is linked to EMT in prostate cancer ............................................................. 17
1.9 ERG Biology/Pathology .......................................................................................................... 19
2.4.7  Electrophoretic mobility shift assay (EMSA) ...................................................... 66
2.4.8  Analyses of gene expression ........................................................................... 67
2.4.9  Real time cell analysis (xCELLigence) ............................................................. 68
2.4.10 Spheroid invasion assay .................................................................................. 68
2.4.11 Zebrafish ....................................................................................................... 69
2.4.12 Statistics ......................................................................................................... 69
2.4.13 Compound solubility and stability ................................................................... 70
2.4.14 Cell cycle analysis .......................................................................................... 71
2.4.15 Proliferation/cell viability assay (Incucyte generated growth curves) .......... 71
2.4.16 Bioinformatics and statistical analyses on gene expression datasets from PCa
        patients .............................................................................................................. 71
2.4.17 Chemical synthesis of VPC-18005 (Figure 2-14C) ......................................... 72

Chapter 3: Utilization of the small molecule screening pipeline to identify inhibitors of ERG
in the Prestwick chemical library .............................................................................. 79

3.1  Introduction ......................................................................................................... 79
3.2  Results and Discussion ...................................................................................... 79
3.3  Materials and Methods ...................................................................................... 87
    3.3.1  Cell culture .................................................................................................. 87
    3.3.2  Dual reporter luciferase assay ..................................................................... 87
    3.3.3  NMR spectroscopy: ERG-ETS domain expression and purification .......... 88
    3.3.4  Electrophoretic mobility shift assay (EMSA) .............................................. 89
    3.3.5  In silico modeling ....................................................................................... 90

Chapter 4: Discussion ............................................................................................... 91
4.1 Future Directions ........................................................................................................................................ 94

4.1.1 Characterize the impact of molecular probes on ERG-Mediated Metastasis .... 94

4.1.1.1 Murine metastasis assays .................................................................................................................. 95

4.1.1.2 Xenograft growth and metastasis ...................................................................................................... 95

4.1.2 Functional Characterization of ERG chemical probes ................................................................. 97

4.1.2.1 Transcriptome analysis of VPC-18005-treated ERG PCa models .......... 97

4.1.2.2 Validation of ERG transcriptional targets and identify novel ERG partners 97

4.2 Conclusion .................................................................................................................................................. 99

Bibliography ..................................................................................................................................................... 101
List of Tables

Table 1-1  Guidelines for the Management of CRPC .......................... 14

Table 2-1 Analysis of the serum concentration curves produced in Supp. Figure 2.16A .......... 74

Table 2-2 Overexpressed genes common in the VPC and TCGA gene expression sets.......... 75

Table 2-3 Overexpressed genes common in the MSKCC and TCGA gene expression sets....... 77

Table 2-4 A list of active VPC-18005 derivatives......................................................... 78

Table 3-1 A summary of top 15 compounds in the Prestwick library that resulted in
approximately 60 percent or lower activity after compound treatment in ERG+ cells and at 10
and 1 μM concentrations.......................................................................................... 86
List of Figures

Figure 1-1 The anatomy of the prostate gland in relation to the pelvic cavity. ......................... 3
Figure 1-2 Coronal section through the prostate gland, depicting its three glandular zones. ....... 3
Figure 1-3 The hypothalamic–pituitary–gonadal axis and testosterone production. ............... 6
Figure 1-4 Metastatic cascade and molecular pathways involved in PCa metastasis............. 18
Figure 1-5 Sequence conservation of SAM and ETS domains across the ETS family members. 20
Figure 2-1 ERG as a drug target and discovery of VPC-18005. ........................................... 30
Figure 2-2 VPC-18005 is stable and soluble. ......................................................................... 32
Figure 2-3 VPC-18005 selectively suppresses ERG-mediated transcriptional activity. .......... 33
Figure 2-4 Characterization of VPC-18005 binding to the ERG-ETS domain. ................. 36
Figure 2-5 VPC-18005 directly binds to the ERG-ETS domain. ........................................... 39
Figure 2-6 VPC-18005 disrupts binding of the ERG-ETS domain to DNA. ......................... 42
Figure 2-7 VPC-18005 disrupts binding of the ERG-ETS domain to DNA. ......................... 43
Figure 2-8 VPC-18005 inhibits SOX9 gene expression. ......................................................... 44
Figure 2-9 VPC-18005 inhibits migration and invasion of prostate cell lines in vitro and in vivo .......................................................................................................................... 48
Figure 2-10 VPC-18005 suppresses migration, but not growth of RWPE-1-ERG cells. ....... 49
Figure 2-11 ERG is overexpressed in prostate cancer ........................................................... 55
Figure 2-12 Mutually exclusive binding of VPC-18005 and DNA with the ERG-ETS domain. 56
Figure 2-13 Preliminary SAR studies using derivatives of VPC-18005. ............................. 57
Figure 2-14 .......................................................................................................................... 58
Figure 2-15 NMR analysis for titration of VPC-18005 with ETV4 or PUI ETS domain. ....... 59
Figure 2-16 Preliminary data for in vivo study of VPC-18005 .............................................. 60
Figure 3-1  Discovery of parthenolide as an inhibitor of the ERG-ETS domain and its transcriptional activity. .................................................................................................................. 82

Figure 3-2  Binding mode of parthenolide to the ERG-ETS domain. ........................................... 83

Figure 3-3  Parthenolide disrupts binding of the ERG-ETS domain to DNA. .............................. 84

Figure 4-1  The future of ERG inhibitors in the clinic ................................................................. 100
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTH</td>
<td>Adrenocorticotropic hormone</td>
</tr>
<tr>
<td>AD</td>
<td>Androgen dependent</td>
</tr>
<tr>
<td>ADT</td>
<td>Androgen deprivation therapy</td>
</tr>
<tr>
<td>AF2</td>
<td>Activation function-2</td>
</tr>
<tr>
<td>AP</td>
<td>Anterior pituitary</td>
</tr>
<tr>
<td>APC</td>
<td>Antigen presenting cell</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen receptor</td>
</tr>
<tr>
<td>ARPI</td>
<td>Androgen receptor (AR) pathway inhibitors</td>
</tr>
<tr>
<td>AR-V</td>
<td>Androgen receptor variant</td>
</tr>
<tr>
<td>BF3</td>
<td>Binding function 3</td>
</tr>
<tr>
<td>BRCA</td>
<td>Breast cancer gene</td>
</tr>
<tr>
<td>CD49f</td>
<td>Integrin α6, stem cell surface marker</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>CRPC</td>
<td>Castration-resistant Prostate Cancer</td>
</tr>
<tr>
<td>CSC</td>
<td>Cancer stem-like cell</td>
</tr>
<tr>
<td>CSS</td>
<td>Charcoal stripped serum</td>
</tr>
<tr>
<td>CUA</td>
<td>Canadian urologic association</td>
</tr>
<tr>
<td>CXCR4</td>
<td>C-X-C motif chemokine receptor 4</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA binding domain</td>
</tr>
<tr>
<td>DHT</td>
<td>Dihydrotestosterone</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's modified eagle medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DRE</td>
<td>Digital rectal exam</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electrophoretic mobility shift assay</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial-to-Mesenchymal Transition</td>
</tr>
<tr>
<td>ERG</td>
<td>ETS-related gene</td>
</tr>
<tr>
<td>ETS</td>
<td>E26 transformation specific</td>
</tr>
<tr>
<td>ETS1</td>
<td>E26 oncogene homolog 1</td>
</tr>
<tr>
<td>ETV</td>
<td>ETS translocation variant</td>
</tr>
<tr>
<td>EWS</td>
<td>Ewing's sarcoma</td>
</tr>
<tr>
<td>EWS-FLI1</td>
<td>Ewing's sarcoma - Friend leukemia integration 1</td>
</tr>
<tr>
<td>EZH2</td>
<td>Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>FDA</td>
<td>The U.S. Food and Drug Administration</td>
</tr>
<tr>
<td>FEV</td>
<td>Fifth Ewing Variant, ETS transcription factor</td>
</tr>
<tr>
<td>FSH</td>
<td>Follicle-stimulating hormone</td>
</tr>
<tr>
<td>FUS</td>
<td>RNA-binding protein FUS/TLS</td>
</tr>
<tr>
<td>FZD4</td>
<td>Frizzled Class Receptor 4</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>Granulocyte-macrophage colony stimulating factor</td>
</tr>
<tr>
<td>GnRH</td>
<td>Gonadotropin-releasing hormone</td>
</tr>
<tr>
<td>GPCR</td>
<td>G protein-coupled receptors</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylases</td>
</tr>
<tr>
<td>HGPIN</td>
<td>High-grade prostatic intraepithelial neoplasia</td>
</tr>
<tr>
<td>IC50</td>
<td>Half maximal inhibitory concentration</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ILK</td>
<td>Integrin-linked kinase</td>
</tr>
<tr>
<td>LBD</td>
<td>Ligand binding domain</td>
</tr>
<tr>
<td>LEF1</td>
<td>Lymphoid enhancer binding factor 1</td>
</tr>
<tr>
<td>LH</td>
<td>Luteinizing hormone</td>
</tr>
<tr>
<td>LHRH</td>
<td>Luteinizing hormone-releasing hormone</td>
</tr>
<tr>
<td>Luc</td>
<td>Luciferase</td>
</tr>
<tr>
<td>mCRPC</td>
<td>Metastatic castration-resistant prostate cancer</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>MMPs</td>
<td>Matrix metalloproteinases</td>
</tr>
<tr>
<td>mpMRI</td>
<td>Multiparametric MRI</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>NCI</td>
<td>National cancer institute</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NTD</td>
<td>N-terminal domain</td>
</tr>
<tr>
<td>PAP</td>
<td>Prostatic acid phosphatase</td>
</tr>
<tr>
<td>PARP1</td>
<td>Poly(ADP-ribose) polymerase 1</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PCa</td>
<td>Prostate cancer</td>
</tr>
<tr>
<td>PCA3</td>
<td>Prostate cancer antigen 3</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PI3K</td>
<td>Phosphatidylinositol-4,5-bisphosphate 3-kinase</td>
</tr>
<tr>
<td>Plau</td>
<td>Plasminogen Activator, Urokinase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PrEC</td>
<td>Prostatic epithelial cells</td>
</tr>
<tr>
<td>PSA</td>
<td>Prostate-specific antigen</td>
</tr>
<tr>
<td>PTEN</td>
<td>Phosphatase and tensin homolog</td>
</tr>
<tr>
<td>PTL</td>
<td>Parthenolide</td>
</tr>
<tr>
<td>PU.1</td>
<td>Protein that is encoded by the SPI1 gene</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RPMI</td>
<td>Roswell Park Memorial Institute medium</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription-polymerase chain reaction</td>
</tr>
<tr>
<td>SAM</td>
<td>Sterile alpha motif</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>shRNA</td>
<td>Short hairpin ribonucleic acid</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering ribonucleic acid</td>
</tr>
<tr>
<td>SPOP</td>
<td>Speckle-type POZ protein</td>
</tr>
<tr>
<td>TBS-T</td>
<td>Tris-buffered saline-Tween 20</td>
</tr>
<tr>
<td>TGFβ</td>
<td>Transforming growth factor beta</td>
</tr>
<tr>
<td>TMPRSS2</td>
<td>Transmembrane protease, serine 2</td>
</tr>
<tr>
<td>USP9</td>
<td>Ubiquitin specific protease 9</td>
</tr>
<tr>
<td>WNT</td>
<td>Wingless/integrated</td>
</tr>
<tr>
<td>ZEB</td>
<td>Zinc finger E-box-binding homeobox</td>
</tr>
</tbody>
</table>
Acknowledgments

I have been fortunate to collaborate with many individuals that have not only taught me many new research perspectives but have been an integral part of my life over the past few years. There is no way I can thank every single individual who has helped in the development and progression of this work; however, I would like to extend my gratitude to several indispensable individuals that have made this work possible.

I was fortunate to work closely with Drs. Miriam Butler, Michael Hsing, and Desmond Lau over the course of my PhD. I have learned so much from them and now have the honor to call them my friends. Importantly, not only were they there for me during times of happiness and prosperity, but they were also pillars of strength during the more difficult times. I also must thank Dr. Kush Dalal, for his friendship and mentorship. Very early on he took me on as a collaborator on his projects and we generated some great science which helped to propel my career in the lab. Overall, I cherish the friendships I have made during my PhD and I wish my friends and colleagues success in their future.

I also had the privilege of mentoring several undergraduate students during my PhD. Thank you to Dennis Ma, Ari Kim, Scott Lien, Paul Yen, Marta Mroczek, and Alice Tai for their dedication and patience during the development and troubleshooting of assays used in this thesis. I hope that they were able to gain valuable scientific experience during their time in the lab and I wish them all success in their chosen career paths.

It is not every day that you meet a mentor that will change your life. Throughout my work as a PhD candidate, I have been very fortunate to be mentored by several prominent faculty members who have influenced me deeply. I appreciate the support and guidance of my supervisors, Drs. Cherkasov and Rennie. I also have tremendous respect for the members of my PhD committee.
Dr. Ralph Buttyan, Dr. Lawrence McIntosh, and Dr. Michael Cox, all of whom supported me as if I were one of their own students. In particular, I have special gratitude for Dr. Michael Cox. I had the opportunity to work very closely with Dr. Cox during the development of this project. He took on a major mentorship role throughout the course of my PhD degree and also helped me to expand my mind with his perspectives on not just science but life in general. I am forever indebted to his kindness.

I would like to send my warm regards and thank you to Dr. Cyrus Ghajar at Fred Hutch Cancer Center who kindly hosted me in his lab for the period of 6 months during my Friedman Scholar visit. Although none of the work I performed there was included in this thesis, I learned plenty from him and his wonderful lab crew. The experience shaped me into a much better scientist and helped me expand my knowledge in the field of breast cancer and the metastatic niche.

I would like to extend my gratitude to my funding sources throughout the course of my PhD. Not only did they provide me with confidence to proceed with my work, but they have shaped me into a much better scientist by providing opportunities to excel that would have not been possible otherwise. Worth noting are: The Frederick Banting and Charles Best Canada Graduate Scholarships Doctoral Awards (CGS-D), UBC Four Year Fellowship, The Roman Babicki scholarship, Prostate Cancer Foundation BC (PCFBC), Friedman Scholars program, and the Michael Smith Foundation for Health Research supplement. These awards not only provided me with salary support but enabled me to travel the world and visit top tier scientists in the field.

Finally, I would like to extend my never-ending gratitude towards my family and friends for their support throughout this work. Even though, I missed many family events due to lab schedule conflicts, they all supported my cause and enabled me to be a better scientist every day. The past
few years have been a roller coaster and despite facing some hardship I know I am now on a much better and more exciting trajectory. I would like to thank Matilda, my furry friend, who joined my life five years ago and has helped me stay healthier by pushing my limits on many hikes and providing me with a better work life balance. Last but not least I would like to thank my partner and Aiden. As Michael J. Fox said, ‘Family is not an important thing. It’s everything.’ Thank you to my entire family from the bottom of my heart for helping me to pursue my dreams.
Dedication

I would like to dedicate this work to all cancer survivors and those being affected by the adversity of this disease. I hope that this work has contributed to the understanding in the field and one day we can celebrate the eradication of this disease.
Chapter 1: Introduction

Prostate cancer (PCa) is the most commonly occurring non-cutaneous cancer in Canadian men, with over 23,300 cases and 4,200 deaths estimated in 2020. When the cancer is diagnosed in its early phases, it can often be eradicated via surgery or radiation or both. It is worth noting that in about 30% of cases disease recurrence or metastasis occurs and they are primarily managed via drugs targeting the androgen receptor (AR) pathway. Unfortunately, the effectiveness of this therapy is temporary due to reactivation of the AR pathway or development of AR independent phenotypes. While several chemo- and immunotherapies are available for managing such advanced PCa, they offer only modest survival benefits. Thus, their limited effectiveness and significant side effects necessitate the development of new therapeutics to combat aggressive and resistant PCa. In breast cancer, the identification of distinct subtypes (i.e. determination of BRCA and hormone receptor status) has led to an increased understanding of the disease and allowed for the prediction of prognosis and response/outcome to subtype targeted treatments. Given that recent genomic analysis of prostate cancer has determined different molecular subtypes targeted therapies hold great promise in the future treatment of prostate cancer.

Normal AR functioning is altered in cancer cells by a plethora of mechanisms including mutations and aberrant gene expressions. A prime example is the fusion of the AR DNA response element of transmembrane protease serine 2 (TMPRSS2) and the ETS-related gene (ERG), adjacently coded on chromosome 21. The TMPRSS2-ERG fusion is the most common genomic rearrangement in PCa occurring in 50% of Caucasian, 20 – 30% of African American and less than 20% of Asian PCa patients. ERG is a transcription factor in the ETS family whose members are implicated in many cancers. Normally ERG is not expressed in prostate epithelial cells, but its fusion with the TMPRSS2 promoter causes AR to drive ERG expression.
Thus, ERG is one of the most commonly overexpressed genes in PCa. Being an oncogenic hub that activates multiple cancer-inducing pathways, ERG can effectively promote epithelial-mesenchymal transition (EMT) and transform normal prostate cells into cancerous and invasive forms. This chapter outlines the role of androgens in PCa, the current PCa treatments, and the significant potential for targeting ERG to improve treatment options.

1.1 The Prostate

The prostate is a composite muscular gland that contains glandular, muscular, fibrous, lymphatic, and nervous tissues. It is a walnut shaped gland that normally weighs around 18 g with an average volume of 24 cm³ in the adult male. As depicted in Figure 1-1, anatomically, it is located in the pelvic region, inferior to the bladder neck and anterior to the rectum. The urethra runs through the body of the gland connecting the bladder to the penis. The prostate is formed during the 10th week of gestation and it is an outgrowth of the urogenital sinus epithelium. The main function of the prostate is to secret nutrient and enzymes for the survival and enrichment of spermatozoa. The human prostate is organized into three distinctive glandular zones (Figure 1-2): 1) the peripheral, which makes up 70% of prostate gland and is the most prone to the development of cancer; 2) the central, which is the most similar pathologically to the murine dorsolateral prostate lobe; and 3) the transitional zone.

On a cellular level, prostate is composed of two major cellular compartments: the epithelium and stroma. The epithelium consists of basal, intermediate, neuroendocrine and luminal secretory cells and the stroma contains smooth muscle cells, connective tissue, and fibroblasts.
Figure 1-1 The anatomy of the prostate gland in relation to the pelvic cavity.
The illustration depicts sagittal section through the male pelvic. Major structures surrounding the prostate superiorly are the bladder and seminal vesicles, posteriorly prostate is related to the rectum and at its base is connected to the urethra and penis through the urinary sphincter.

Figure 1-2 Coronal section through the prostate gland, depicting its three glandular zones.
The figure depicts the prostate and its three glandular zones (central, peripheral, transition). Most of the prostate is derived from the urogenital sinus except for the central zone that is derived from Wolffian duct. The peripheral zone is the most common site of cancer and the transitional zone is the almost always the site of benign prostatic hyperplasia.
1.2 Androgen Signaling in the Prostate

Androgen, the primary male sex hormone, is responsible for growth and development of the prostate and its actions are governed by the hypothalamus-pituitary-gonads axis, or the androgen axis. As portrayed in Figure 1-3, gonadotropin-releasing hormone (GnRH) released by the hypothalamus in response to low androgen levels in the body, stimulates the anterior pituitary (AP) gland. The stimulation of the AP leads to the production of luteinizing hormone (LH) and follicular stimulating hormone (FSH) that acts on the Leydig cells present in the testes to produce testosterone. Alternatively, stimulated AP leads to the production of adrenocorticotropic hormone (ACTH) that works on the adrenal glands of the kidneys to produce androgen derivatives. While testosterone is the most produced androgen in men, it is dihydrotestosterone (DHT), which is produced when testosterone is reduced by the enzyme 5α-reductase, that is the most active form. Finally, the androgen produced from either of the routes exert their effects through binding to the androgen receptor (AR).

The AR is activated by binding to androgen and this results in a conformational change that allows it to translocate into the nucleus of the cell. The AR protein is made up of an N-terminal domain (NTD), followed by the DNA binding domain (DBD) and the ligand binding domain (LBD) domains. X-ray crystal structures of the LBD and DBD have assisted in defining surface-exposed regions on the AR that facilitate ligand, DNA, and co-factor binding. Thus, it has been well characterized how androgens bind to a ligand binding pocket on the surface of the LBD termed the androgen-binding site. The LBD site is the best understood target for anti-androgen compounds (such as enzalutamide) that compete with testosterone for binding. In

---

1 A version of this work has been published in J Biol Chem. 2014 Sep 19;289(38):26417-29. doi: 10.1074/jbc.M114.553818. Reprinted with permission from JBC
addition, the LBD contains alternative surface-exposed pockets such the activation-function 2 (AF2) region, important for co-regulator recruitment \(^{28}\), and the binding-function 3 (BF3) site of unknown function, located near the androgen-binding site \(^{29,30}\). The AR-DBD contains P-box and D-box amino acid motifs that are involved in nucleic acid binding and DBD-mediated dimerization. The AR acts as a transcription factor that drives the expression of genes that contribute to the growth and maintenance of the normal prostate and the development, progression, and recurrence of PCa \(^{31-33}\).
Figure 1-3 The hypothalamic–pituitary–gonadal axis and testosterone production.
The production of the androgen starts at the hypothalamus and with the activation of the AP, leads to the activation of FSH/LH and ACTH which activate the production of androgens in testes (~80-90% of body’s total androgen) and adrenal glands (~10-20% of body’s total androgen). The growth of the prostate is mainly regulated by the more active version of the testosterone (DHT). GnRH, gonadotropin-releasing hormone; AP, anterior pituitary; LH, luteinizing hormone; FSH, follicle stimulating hormone; ACTH, adrenocorticotropic hormone; DHT, dihydrotestosterone; AR, androgen receptor; P, prostate; +, positive influence; –, negative feedback.
1.3 Prostate Cancer: diagnosis and current therapies

The first recorded case of PCa was diagnosed in 1853 by J. Adams, a surgeon in London. At the time, this phenomenon was reported as a “rare disease”, but we now know that PCa poses as a significant health problem worldwide. In Canada, PCa is the most commonly diagnosed non-cutaneous cancer in men and the third leading cause of death. Since this first recorded case, our perspective in the field has evolved to the point where we now understand that disease incidence and outcome are influenced by multiple factors: predominantly age, race, and family history, with an ever growing appreciation for specific genetic predisposition.

A multifaceted approach is used in the diagnosis of PCa however treatment options are also influenced by the patients age, life expectancy, risk of treatment, and the patient’s personal preference and as such are decided on a case by case basis. Generally, diagnosis is made according to two main tests: the digital rectal exam (DRE) and blood testing for Prostate Specific Antigen (PSA). The Canadian Urologic Association (CUA) recommends PSA blood testing for high-risk individuals with greater than 10 years life expectancy to begin at age 45. PSA is produced by a gene that is under direct transcriptional control by the AR and so is secreted by the normal and malignant prostate. In the instance that PSA levels are above 3 ng/ml it is recommended that the patient undergoes more frequent PSA testing to determine if levels are rising. In addition, PSA testing is used in conjunction with a DRE where the prostate is palpitated to feel for any irregularities (i.e. hardness, lumps) which may be indicative of cancer. Should both PSA testing and the DRE both return positive results additional tests should be carried out and these include ultrasound-guided core tissue biopsies to determine the histological and morphological properties of the prostate cells. More recently the PCa antigen 3 (PCA3) urine test and multiparametric magnetic resonance imaging (mpMRI) have been included in PCa diagnosis as they are less invasive.
With advancements in early diagnosis and detection, nearly 90% of all PCa cases are diagnosed as a localized disease at which time the treatment options of radical prostatectomy and radiation therapy are potentially curative. In addition, treatment sometimes includes neoadjuvant (prior to therapy) or adjuvant (in conjunction with therapy) hormonal therapy. PCa, like the non-malignant prostate, is dependent on androgens for proliferation and survival. Neo-adjuvant hormonal therapy is used to reduces the size of the tumor before primary treatment and there is some evidence that the use of adjuvant hormonal therapy, used directly after treatment to target any residual cells, confers a survival benefit following radiotherapy \(^{38}\).

Additionally, it is worth noting that up to 67% of men to be diagnosed with PCa have clinically inconsequential PCa that will not be associated with illness or death \(^{39-41}\). In such cases active surveillance, where the patient receives regular PSA monitoring and DRE, helps reduce the over diagnosis of clinically significant PCa. However, it is still unclear how to stratify low risk patients from those at high risk and active surveillance itself is associated with lowering in the quality of life \(^{42}\), which makes this one of the more contentious issues currently facing the field of urology.

### 1.4 Management of metastatic complications in prostate cancer

In about 30% of cases, prostate cancer patients will relapse after their primary therapy due to the detection of metastasis or a rise in PSA (blood) or PCA3 (urine) levels \(^2\). While local relapse may allow for a second therapy (also known as salvage therapy, e.g. radiation therapy) after radical prostatectomy, most patients will be diagnosed with incurable advanced disease where the mainstay of treatment is androgen deprivation therapy (ADT). ADT aims to eliminate androgen production and action in the body.
1.4.1 Castration

Testicles are the main source of androgens (e.g. testosterone) in the male body. Castration aims to reduce the level of testosterone in the blood which, in the case of prostate cancer, should relieve the patient of their pain symptoms or difficulty passing urine. Medical castration is normally carried out because surgical castration (orchidectomy), whilst effective, is irreversible and causes significant psychological distress to the patient. Medical castration is achieved though drugs designed to interfere with androgen production via the hypothalamic-pituitary-gonadal axis (figure 1.3). Examples include luteinizing hormone-releasing hormone (LHRH) agonists or LHRH antagonists that suppress luteinizing hormone leading to suppression of the production of testosterone in the testes.

1.4.2 Anti-androgens

Anti-androgens, also known as AR antagonists, are competitive inhibitors of androgen action through direct competition for binding to the LBD of the AR. Androgen Receptor Pathway Inhibitors (ARPIs) include the steroidal antagonists cyproterone acetate and abiraterone and the non-steroidal antagonists enzalutamide, apalutamide, and darolutamide. While bicalutamide has high affinity for AR, it is still 50-fold less than DHT. Enzalutamide is a second generation AR antagonist that has increased potency and significantly prolongs progression-free and overall survival in men with metastatic CRPC +/- testosterone suppression. Anti-androgens are usually given in combination with LHRH agonist or LHRH antagonists and this is referred to as combined androgen therapy.

1.5 Metastatic castration-resistant prostate cancer

Although initially effective, resistance to hormonal therapies occurs within five years of treatment initiation and the disease eventually progresses to castration resistant PCa (CRPC).
CRPC is a result of emerging resistance mechanisms that are related to the AR, including: amplification of AR protein expression, mutations at the androgen-binding site on the AR, and/or elevated production of AR variants \(^4,48,49\). Significant efforts are underway to try to develop drugs to target these AR variants. For example, AR splice variants arise from messenger RNA lacking the LBD coding sequence. The best characterized AR splice variants are AR-V7 and AR-v567es and these are implicated in several studies to contribute to reactivation of AR signaling in castration-resistant tumors \(^4,50-57\). The absence of an LBD would prevent access to the androgen-binding site where most traditional anti-androgens target. This leaves only the NTD and DBD as viable domains. Inhibition of splice variant transcriptional activity would be a significant breakthrough in the development of a new class of anti-AR drugs and several molecules have been identified and are progressing to clinical studies \(^58\).

Currently, the CUA recommends clinical therapies for CRPC disease according to the subtype of disease present in the patient, as detailed in Table 1-1. These recommendations include several chemo- and immunotherapies to manage advanced PCa, but ultimately they offer only modest survival benefits and have significant side effects \(^3\). For example, docetaxel and cabazitaxel are two approved chemotherapies that bind tubulin to disrupt cell division. Both taxane therapies have shown benefit to CRPC patients however the individual patient responses are variable, i.e. some patients will experience a significant response while others will experience no response \(^59\). Thus, to extend survival, more potent agents targeting androgen signaling have been developed. Most notable is the AR antagonist enzalutamide, which is effective in patients who have failed ADT and docetaxel-based chemotherapy \(^60\). However, resistance to enzalutamide also occurs and so additional AR antagonists are in development. For example, EPI-506 targets the AR transactivation domain (NTD) and was the first NTD inhibitor to be tested in a phase I clinical
trial. It was well tolerated but highly metabolized and so more potent and stable next generation molecules are in development. Additionally, our lab has identified a compound that targets the AR DNA binding domain which will retain activity against AR splice variants and so has the potential to be an effective therapy in enzalutamide resistant or AR-variant driven prostate cancer. Alternative non-hormonal therapies are also available to slow the growth of the cancer and prolong life. Sipuleucel-T is a personalized immunotherapy treatment in which the patient’s own antigen presenting cells (APC) are extracted from their blood and activated to target their cancer. Briefly, the patient’s isolated APCs are incubated with a fusion protein that contains antigen prostatic acid phosphatase (PAP), which is present on most prostate cancer cells, and granulocyte-macrophage colony stimulating factor (GM-CSF), which is an immunostimulant that helps APC maturation. The activated blood is then reinfused into the patient where it can enhance the patient’s immune response against PAP. It has shown survival benefit in several clinical trials and is an FDA approved treatment. In addition, Radium 233 (Alpharadin), is a calcium mimetic that can accumulate in the bone around metastatic deposits where it is able to expose the surrounding area to radiation and thus specifically target bone metastasis. Radium 233 has recently been FDA approved as it is able to reduce the severity of skeletal related events and improve overall survival. Poly-ADP-ribose polymerase (PARP) inhibitors, which work through inhibition of single-strand DNA repair that leads to DNA damage and cell death, have recently been approved to treat men with metastatic CRPC that have progressed following prior treatment with enzalutamide or abiraterone. Olaparib has been approved for men with homologous recombination repair gene mutations and rucaparib has been approved for men with BRCA gene mutations.
1.6 Changing patterns of metastatic CRPC

In addition to the development of resistance or lack of response to current therapies, current treatments have also altered the metastatic patterns of prostate cancer. Traditionally, the most common site for metastatic CRPC is the skeletal system. In a study that investigated approximately 19,000 patients it was found that cases of non-bone metastasis increased 1.6% per year between 1990 and 2012. Since there have been an increasing number of approved therapeutic agents available over this period, particularly since the year 2000, it was suggested that this was the reason for the increased incidence of non-bone metastasis in patients who have undergone chemotherapy. In contrast, the rate of bone metastasis decreased by 0.5% every year during the same period. This reduction in bone metastasis may be due to better screening technologies (e.g. positron emission tomography (PET) scan) and thus the elimination of false positive diagnoses. Alternatively, patients with bone metastasis develop skeletal complications that are associated with reduced quality of life and increased mortality. In prostate cancer, ADT leads to bone loss which further increases the risk for skeletal complications. Thus, treatment regimens are also increasingly including the use of the bisphosphonate derivative zoledronic acid or denosumab to maintain bone health and reduce skeletal complications related to bone metastasis. More recently, studies also suggest that zoledronic acid and denosumab treatment are also able to delay time to first bone metastasis. Overall, these results suggest that the selective pressure of treatments has altered the natural history of mCRPC. While the study by Doctor et al. did not demonstrate increased incidence of liver metastasis, a study by Singh et al. found liver metastases to be a more aggressive subtype that did not respond well to hormone therapy or chemotherapy.
In addition, approximately 25% of the men who die of prostate cancer have tumors that have a neuroendocrine phenotype which is associated with low AR signaling and poor prognosis. The increased incidence of neuroendocrine prostate cancer (NEPC) is correlated with the use of new generation potent AR pathway inhibitors, such as abiraterone and enzalutamide. The biology of NEPC is still unclear and currently there is no effective treatment for NEPC. However, it is thought that prostate cancer cells undergo trans-differentiation into NEPC as a mechanism of adaptive resistance to AR-targeted therapies and due to their lack of AR there is continued growth of the cancer due to stimulation by the paracrine system. Furthermore, there is a recently identified subtype, double negative PCa, which is AR and NEPC marker negative. Overall, while these changing forms of PCa are currently rare they highlight the changing face of the disease and the need to develop novel therapies to combat aggressive and resistant PCa and prevent and/or inhibit metastasis.
Table 1-1 Guidelines for the Management of CRPC

<table>
<thead>
<tr>
<th>Patient with CRPC</th>
<th>Management option</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑PSA/ Non-metastatic CRPC</td>
<td>No approved regimen</td>
</tr>
<tr>
<td></td>
<td>Secondary hormonal therapy</td>
</tr>
<tr>
<td>Metastatic CRPC without symptoms</td>
<td>discontinue antiandrogens</td>
</tr>
<tr>
<td></td>
<td>Abiraterone</td>
</tr>
<tr>
<td></td>
<td>Enzalutamide</td>
</tr>
<tr>
<td></td>
<td>Docetaxel</td>
</tr>
<tr>
<td>Metastatic CRPC with symptoms</td>
<td>Docetaxel</td>
</tr>
<tr>
<td></td>
<td>Radium-233</td>
</tr>
<tr>
<td>Metastatic CRPC progress post docetaxel-based chemotherapy</td>
<td>Abiraterone</td>
</tr>
<tr>
<td></td>
<td>Enzalutamide</td>
</tr>
<tr>
<td></td>
<td>Cabazitaxel</td>
</tr>
<tr>
<td></td>
<td>Radium-223</td>
</tr>
<tr>
<td>CRPC and bone metastases</td>
<td>denosumab</td>
</tr>
<tr>
<td></td>
<td>zoledronic acid</td>
</tr>
<tr>
<td></td>
<td>calcium &amp; vitamin D</td>
</tr>
</tbody>
</table>
1.7 Metastasis and cellular plasticity

Metastasis, simply put, is the transition of cancer cells from one location in the body to another. For epithelial-derived cancers (adenocarcinomas), this requires a complex series of events that enables tumor cells to invade into the surrounding stoma, enter the circulation (i.e. blood or lymph), spread to distant sites, exit the circulation, and establish growth in a secondary location (Figure 1-4A). One way cancer cells can achieve this by taking advantage of a normal physiological process that is required for wound healing and organogenesis called EMT. Environmental factors, such as soluble secretions (e.g. matrix metalloproteinases (MMPs)) from the stroma and infiltrating immune cells are able to trigger several pathways, including transforming growth factor-beta, Wnt/beta-catenin, fibroblast growth factor, epidermal growth factor, and Notch that in turn lead to activation of members of the snail, twist, and zinc finger E-box-binding homeobox (ZEB) pathways (Figure 1-4B). This results in loss of epithelial factors, such as: E-cadherin, epithelial adhesion molecule, cytokeratin’s and a gain of mesenchymal factors, such as vimentin, fibronectin and N-cadherin, which facilitates dissemination of the tumor cells by reducing cell adhesion, changing cell shape and polarity, and providing the ability to migrate from the primary tumor, invade into the stroma, and enter the bloodstream. Significantly, as metastases often resemble the primary tumor it corresponds that cells are also capable of the reverse transition, i.e. mesenchymal-epithelial transition (MET), which restores the proliferative capacity of the cells and enables disseminated tumor cells to form metastases at distant sites. However, this is a simplified version of cellular plasticity and there are likely multiple transitional states between epithelial and mesenchymal states and correspondingly multiple cellular properties. For example, cells that have gone through partial EMT have cancer stem-like
cell (CSC) properties. CSC are able to resist apoptosis, self-renew, survive and colonize at the distant site, and also are more resistant to therapy.

1.7.1 Targeting cellular plasticity to improve patient outcome

In prostate cancer progression, EMT is associated with high Gleason score, a shortened time to biological recurrence, and the presence of bone metastasis. As discussed, the AR is a major driver of prostate cancer growth. However, there appears to be a context dependent relationship between androgen signaling and cellular plasticity. AR signaling inhibits EMT and ADT can relieve this inhibition. For example, ADT increased expression of mesenchymal markers N-cadherin and cadherin-11 in patients with prostate cancer. ADT also increases the expression of AR splice variants, including AR-V7, which is constitutively active and can induce expression of EMT and CSC-associated genes. However, the AR is also able to upregulate expression of Slug that in turn interacts directly with the AR to coactivate gene expression with or without androgen. Thus, the Slug-AR interaction results in the androgen independent growth that enhances castration resistance. Thus, it appears that normal AR signaling, i.e. normal prostate or treatment naïve tumors, inhibits EMT. Whereas, in cases of resistance AR signaling can promote EMT. Therefore, drugs that target the epithelial phenotype, such as enzalutamide or abiraterone (i.e. androgen signaling or synthesis inhibitors), as discussed earlier are part of standard prostate cancer therapeutic management. However, in cases of resistance when constitutive AR is active alternative approaches are needed. For example, when there is increased expression of mesenchymal markers, such as N-cadherin, an N-cadherin targeted antibody is capable of inhibiting EMT and delays time to resistance in vivo. Alternatively, targeting the stroma may indirectly reduce epithelial mesenchymal plasticity. Prodrugs have been developed that target fibroblast specific protein and in animal models have been shown to inhibit growth with an anti-
tumor effect comparable to the effect of docetaxel but without the toxicity. Overall, these studies suggest that combination therapies which include mesenchymal or stromal targeted agents may ultimately produce a more enhanced effect of current therapies. While there are additional targeting strategies available toward cellular plasticity, such as HDAC inhibition and targeting of the TGFβ or Notch pathways, there are very few molecules that have made it to or passed clinical trial. Thus, there is a need for additional or improved strategies to inhibit epithelial plasticity and subsequent metastasis.

1.8 ERG expression is linked to EMT in prostate cancer

Genome wide analytics has uncovered some of the most prevalent genomic alterations that are correlated with metastatic CRPC (mCRPC), including: AR overexpression and mutations, loss/mutations of key tumor suppressors, p53 and phosphatase tensin homologue (PTEN), and, with respect to this thesis, gene fusions that direct aberrant expression of members of the ETS family, such as TMPRSS2-ERG. Being an oncogenic hub that activates multiple cancer-inducing pathways, ERG can effectively promote cell migration and invasion, and can promote malignant transformation of PrECs. ERG expression has been shown to suppress prostatic epithelial differentiation, and to promote EMT by activation of FZD4mediate WNT/LEF1 signaling, TGFβ signaling, ZEB1/ZEB2 and ILK signaling (Figure 1-4B). The resulting suppression of E-cadherin expression, and enhanced motility and invasion of ERG transformed PrECs has been linked to expression of vimentin, matrix metalloproteinases (MMPs) and CXCR4. These events imply that ERG expression promotes epithelial dedifferentiation and expression of stem cell surface markers such as CD49F indicating that ERG has the potential for facilitating adaptive responses to therapies and allowing tumors to acquire further aggressive characteristics.
Figure 1-4 Metastatic cascade and molecular pathways involved in PCa metastasis.
A) Depiction of metastatic cascade starting with primary tumor that via EMT can enter the circulation to a distant site where once out of the vasculature is able to form micro-metastasis and subsequent distant metastasis. (BioRender program was utilized to produce this illustration)
B) The extracellular molecules and pathways involved in mediating epithelial mesenchymal transition. Highlighting the molecules/pathways targeted by ERG (bold) to promote EMT.
1.9 ERG Biology/Pathology

The E26 transformation-specific (ETS)-gene family is highly conserved in eukaryotes and includes 28 human members organized into 10 subfamilies \(^{119,120}\) (Figure 1-5B). All members possess a conserved winged helix-turn-helix motif defined as the ETS domain and regulate transcription of a plethora of growth and regulatory genes \(^{121}\). ERG expression is tightly restricted to cells of specific mesodermal lineages during development. In particular, ERG is expressed in the early phases of hematopoietic, chondrocytic and endothelial differentiation \(^{119}\). ERG is also expressed in the mesenchyme of the developing genital ridge and subsequently during renal and urogenital tract differentiation. The only known example of ERG expression in the epithelial lineage is transient expression in mouse E8.5 migrating neural cells \(^{122}\). Therefore, recognition of ERG as a regulator of chondrogenesis \(^{123,124}\), angiogenesis \(^{124}\), and of endothelial cell differentiation and survival \(^{125}\), suggests that aberrant expression in the prostatic epithelia might mediate transformation by activation of these cellular processes.

The ERG subgroup within the ETS family of transcription factors contains two additional members, FLI1 and FEV. FLI1 shares 98% homology with ERG \(^{126}\). In Ewing’s sarcomas (EWS) and some primitive neuroectoderm tumor, translocation of FLI1 and ERG with EWS results in fusion proteins that promote disease pathogenesis \(^{127,128}\). The EWS-ERG translocation variants involve the fusion of EWS with the carboxy-terminal of ERG that includes the signature ETS DNA binding domain (Figure 1-5C). The resulting chimeric proteins affect intrinsic RNA splicing function of EWS \(^{129}\), and to alter gene expression profiles via ERG-mediated transcriptional activity \(^{130}\). Similarly, in certain acute myeloid leukemias, chimeric fusion proteins resulting from ERG translocations with TLS/FUS include the ERG amino-terminal SAM/Pointed domain and the ETS domain, and mediate malignant transformation by virtue of sustained ERG transcriptional
More than 200 ERG target genes are thus activated to regulate mitogenesis and migration. 

Figure 1-5 Sequence conservation of SAM and ETS domains across the ETS family members.
(A) One dimensional representation of the SAM/PNT and ETS domains on the full-length ERG protein sequence (residue numbering based UniProt P11308-4). The domain boundary positions are defined according to NCBI CDD. (B) Left panel: A phylogenetic tree constructed based on a multiple alignment of full-length protein sequences of all 28 ETS factors using Clustal Omega. Pair-wise percent identities are shown for SAM domain (% on the left) and ETS domain (% on the right) between each ETS factor and ERG, respectively. (*ETV7 is conserved in human, but not in mouse). Right panel: Two-dimensional heat map representation of the corresponding multiple sequence alignments among the 28 ETS factors at the SAM and ETS domains, visualized by Jalview. A darker color represents a residue that matches the consensus sequence (more conserved), while a lighter color represents a different residue but with a positive Blosum62 score (less conserved). A non-conserved residue or a gap is colored in white. Consensus sequences are shown at the bottom. (C) Three-dimensional representation of the SAM domain (PDB ID: 1SXE) and ETS domain (PDB: 4IRI) of the ERG protein. Conserved residues are color-coded according to the heat maps in B) This figure was modified from 134.
1.10 TMMR2-ETS rearrangements in PCa

Recurrent chromosomal rearrangements that lead to the expression of fusion genes or the dysregulation of oncogenes are key events in the development of several hematologic malignancies and sarcomas. Similar rearrangements that involve overexpression of ETS transcription family members, ETV1 and ERG also contribute to PCa etiology. The most prevalent example, occurring in about 40-50% of all PCa’s is the recurrent, non-random chromosomal fusion of the promoter of androgen-regulated serine protease, TMMR2, and full-length or 5’ exon-truncated open reading frames of ERG. The TMMR2:ERG fusion can occur by interstitial deletion or chromosomal translocation. Whole genome sequencing of patients with primary prostate tumors revealed that structural rearrangements are highly prevalent with around 90 rearrangements detected per genome. Interestingly, in the case of TMMR2:ERG tumors, DNA breakpoints are specific and occurred in transcriptionally active chromatin enriched with androgen receptor regulated regions. Thus, because prostate cancer is a highly androgen regulated disease it makes sense that TMMR2-ERG fusions are the most common genomic rearrangements in PCa.

Reports indicate that approximately 20% of high-grade prostatic intraepithelial neoplasia’s (HGPIN) carry TMMR2-ERG fusions, which suggests that these translocations occur early in disease. However, the clinical implications of TMMR2-ERG rearrangements with PCa progression and outcome remains controversial. Several studies report no association, numerous others have linked TMMR2-ERG rearrangements with poor disease outcome, especially when in conjunction with amplification of the translocated locus, ERG protein expression, and/or loss of PTEN. The resulting implication that TMMR2-ERG rearrangements could be used to identify patients at risk for developing clinically relevant disease.
has spurred efforts to establish non-invasive screens for TMPRSS2-ETS translocations alone or in conjunction with other prostatic biomarkers to improve risk stratification of men diagnosed with PCa 174-179.
1.11 ERG as a potential driver of PCa initiation, progression and metastasis

Normal AR functions are altered in PCa by a variety of mechanisms\textsuperscript{8-10}. A prime example is AR being usurped to drive ERG expression in tumors harboring TMPRSS2-ERG rearrangements\textsuperscript{11,180}. ERG is not normally expressed in prostatic epithelial cells (PrECs)\textsuperscript{181}, but its fusion with the TMPRSS2 promoter causes AR to drive ERG expression. Initial studies reported that expression of ETV1 or ERG in normal or immortalized PrECs increased the migration/invasion phenotype but did not alter mitotic rate or transformation\textsuperscript{111,152,182}. However, more recent studies have implicated expression of different TMPRSS2-ERG fusion isoforms in the transformation and suppression of differentiation of immortalized PrEC models\textsuperscript{104,106,113,160,183}. Furthermore, mice expressing ETV1 or ERG in prostatic epithelia exhibit loss of cell adhesion and polarity, gain of nuclear polymorphism and hyperplasia\textsuperscript{111,152}, and progress to overt adenocarcinoma when crossed with animals systemically hemizygous for PTEN, in the context of PI3K activation, p53 loss, or AR over-expression\textsuperscript{113,118,184}. Additionally, ERG can promote phenotypic transition in PCa by disrupting AR signaling events critical for maintenance of lineage-specific differentiation of prostatic epithelia through physically interacting with the AR, binding to the AR promoter, and promoters of AR-regulated genes\textsuperscript{10,165}. ERG is also implicated in epigenetic reprogramming\textsuperscript{10,185} and promoting genomic instability\textsuperscript{186}. While ERG expression is initially driven by the AR via fusion with androgen-responsive promoters, self-driven, feed-forward regulation of the remaining wild-type ERG allele has also been reported\textsuperscript{187}. Also, inactivating mutations in SPOP, decrease ERG degradation, and lead to elevated levels of wild-type ERG\textsuperscript{188,189}. Thus, accumulation of the mutant and wild-type ERG proteins, initially driven by AR, that later become self-sustained when in conjunction with mutations in SPOP, presents a new mechanism that may contribute to ARPI-resistance and disease progression in CRPC.
However, recent genomic analysis suggest that ERG expression does not occur in cancers that have SPOP mutations (~ 10% of prostate cancers)\textsuperscript{101,190}.

1.12 Lack of ERG-targeted therapies

Although ERG is strongly implicated as a critical factor driving PCa development, progression, and metastasis\textsuperscript{102,103,191}, there is no approved drug directly targeting ERG, or any other ETS proteins\textsuperscript{14,192}. Inhibition of VCaP xenograft growth with ERG-targeted siRNA\textsuperscript{104,193} and shRNA\textsuperscript{183} indicates the utility of ERG-targeting. ERG activity has also been targeted by compound DB1255, which binds to ETS binding DNA motifs\textsuperscript{194}. A peptidomimetic, retroinverso ERG-inhibitory peptides (RI-EIP; OncoFusion Therapeutics Inc.), has been shown to inhibit the ERG protein by mimicking the peptide by binding to the ETS domain and disrupting protein-protein interactions\textsuperscript{195}. In addition, ERG has been targeted indirectly through inhibition of its binding proteins PARP1\textsuperscript{187} and USP9X\textsuperscript{196}, as well as the ERG downstream target gene, YAP1\textsuperscript{197}. Among the various anti-ERG strategies, the only small molecule demonstrated to directly bind ERG (Although the exact binding mechanism is unknown) is an experimental compound, YK-4-279. Initially developed to inhibit FLI1 in Ewing sarcoma, YK-4-279 was later shown to inhibit ERG-mediated PCa cell invasion\textsuperscript{198,199}. The exact ERG binding site is unknown, however in Ewing sarcoma YK-4-279 acts to disrupt the interaction between EWS-FLI1 and RNA helicase A blocking the transcriptional activity of EWS-FLI1. While YK-4-279 has issues concerning specificity, toxicity, oral bioavailability and pharmacokinetics\textsuperscript{200,201}, its clinical derivative TK216 is currently in a Phase 1, dose escalation study in patients with relapsed or refractory Ewing’s Sarcoma\textsuperscript{202}. This study shows that targeting a specific ETS-associated tumor is possible even with a small molecule that interacts with other members of the ETS family.
1.13 Hypothesis and Aims

Small molecules directly targeting the ERG ETS domain will suppress transcriptional activity resulting in inhibition of the metastatic potential of ERG-expressing PCas and will lead to an entirely new generation of novel therapeutics for the treatment of patients with mCRPC.

1. To develop and apply *in vitro* assays to assess inhibitory potential of the lead anti-ERG compound using biophysical, biochemical, and cell-based reporter assays.

2. Assess potential biologic efficacy of the lead compounds using *in vitro* and *in vivo* viability and phenotypic assays.
Chapter 2: Discovery and characterization of small molecules targeting the DNA-binding ETS domain of ERG in PCa

2.1 Introduction

Although confounded by disease heterogeneity, the emergence of genome-wide analytics has begun to reveal the spectrum of recurrent genomic alterations that may directly affect PCa disease progression and outcome. The TMPRSS2:ERG fusion is the most prevalent genetic irregularity. ERG has been implicated as an oncogenic hub that drives PCa development, progression, and metastasis and makes it a promising drug target.

The feasibility of direct small molecule targeting a member of the ETS factor family has been demonstrated by YK-4-279 which is in clinical development for Ewing’s sarcoma. This molecule was identified from surface plasmon resonance screening of a small-molecule collection from the National Cancer Institute Drug Targeting Program. YK-4-279 disrupts the binding of the transcriptional cofactor RNA Helicase A, to ETS factor, FLI1, in Ewing’s sarcoma. In addition, YK-4-279 has also been reported to antagonize ERG activity although its exact ERG binding mode has yet to be determined. As an alternative to targeting cofactor interactions, we hypothesized that the use of rational drug design approach, supported by in vitro and in vivo screening methods, could identify small molecules that directly target the DNA-binding ETS domain of ERG.

Here we report use of an established drug discovery pipeline that combines in silico prediction with in vitro and in vivo experimentation to identify a new class of anti-ERG compounds. We demonstrate that a lead anti-ERG compound, VPC-18005, inhibits ERG-induced transcription and interacts directly with the ERG-ETS domain, and disrupts the ERG binding to

---

2 Chapter 2 has been previously published. In the journal of Oncotarget. 8(26) 2017: 42438–42454. Reprinted with permission from Oncotarget
DNA. In addition, the compound reduces migration and invasion rates of ERG-overexpressing cells and inhibits metastasis in zebrafish xenograft models. These results demonstrate that the discovered compound and its derivatives can be developed as therapeutic options for mitigating disease progression in men with ERG-expressing PCa and ultimately lead to improved survival for men with advanced disease.

2.2 Results

2.2.1 Discovery of small molecules that target the DNA-binding ETS domain of ERG protein

There are numerous TMPRSS2-ERG fusions that encode for ERG transcripts. Whereas, the majority produce amino terminal-truncated ERG proteins, all retain the C-terminal DNA-binding ETS domain. We chose to target the DNA-binding ETS domain because it is essential for the ability of ERG to function as a direct transcriptional regulator. There is also structural data available for its complex with DNA. Thus, we reasoned that we could use in silico approaches to identify small molecules targeting the ETS domain. Such molecules should therefore inhibit transcriptional activity of all functional ERG mutant proteins by antagonizing their ability to interact with DNA. This in turn might disrupt ERG-mediated transformational events involved in PCa disease development and progression.

A structure-based virtual screening approach, previously established for targeting protein-DNA and protein-protein interaction interfaces, was applied to the 1.7 Å resolution ERG-ETS domain crystal structure [PDB ID: 4IRG]. The DNA binding interface was identified from a 2.8 Å resolution crystal structure of the corresponding ERG-DNA complex [PDB ID: 4IRI] (Figure 2-1A). The ERG-ETS domain contains a winged helix-turn-helix motif, with helix α3
positioned within the major groove of the DNA containing a cognate GGAA sequence \(^{204}\). A top-ranked druggable surface pocket was identified by virtual atomic probes to partially overlap this ERG-DNA interface (Figure 2-1B). The identified pocket is adjacent to the DNA recognition helix (\(\alpha3\)), and thus it was predicted that a small molecule bound at this site will competitively block DNA binding. Three million chemical structures derived from the ZINC database \(^{205}\) were individually docked into this pocket. Combining the docking scores, binding poses, consensus voting and drug-like properties (detailed in Materials and Methods), an initial set of 48 compounds, representing 45 different chemical classes, were selected for \textit{in vitro} analysis.

To evaluate the biological anti-ERG activity of the compounds identified above, we first assessed ERG expression in a panel of prostate cell lines (Figure 2-1C). We confirmed expression of the ERG protein in VCaP (endogenous overexpression) and PNT1B-ERG cells (stable ERG overexpression \(^{106}\)). In contrast, PC3, PNT1B and PNT1B-Mock cells were negative for ERG expression \(^{106,206}\). Each of the compounds was first evaluated in PNT1B-ERG cells at concentrations of 10 \(\mu\)M and 25 \(\mu\)M for its ability to inhibit ERG transcriptional activation of a transiently transfected, endoglin E3 promoter-derived \(^{207}\), ETS-responsive firefly luciferase reporter (pETS-luc) construct containing 3 conserved ETS recognition (GGAA) motifs. A representative example of 5 compounds that showed suppression of the luciferase reporter by 20%-60% are identified in Figure 2-1D. Compound VPC-18005 was identified as the most potent inhibitor of luciferase activity from this initial set. The molecular docking score of VPC-18005 was ranked in the top 0.01\% of all 3 million molecules evaluated in the virtual screening discussed earlier (Figure 2-1E). Before proceeding with in-depth analysis, the media solubility and stability of VPC-18005 were assessed (Figure 2-2). VPC-18005 was soluble in media and remained stable for at least 3 days (93\%). For comparison, the published inhibitor YK-4-279 was soluble but less
stable (60%). A more thorough dose response analysis was performed using both VCaP and PNT1B-ERG cells to evaluate the potency of VPC-18005. VPC-18005 was found to inhibit pETS-luc reporter activity in PNT1B-ERG and VCaP cells with IC50 values of 3 and 6 μM, respectively (Figure 2-1F). For comparison, YK-4-279\textsuperscript{199} exhibited IC50 values of 5 μM and 16 μM in PNT1B-ERG and VCaP cell-based ETS-Luc reporter assays, respectively (Figure 2-1G).

To further assess if VPC-18005 has any non-specific cellular effect, luciferase assays were performed in PNT1B-MOCK and -ERG cells treated with increasing concentrations of VPC-18005 (Figure 2-3A), and VPC-18005 had minimal impact on the reporter signal in PNT1B-MOCK as compared to PNT1B-ERG cells. Furthermore, overexpression of ERG protein through R1881 treatment counteracted VPC-18005 inhibition in the luciferase assay (Figure 2-3B). VPC-18005 was also tested against an androgen receptor luciferase reporter (ARR3tk-luc) and showed no significant effect on the reporter expression (Figure 2-3C). Collectively, these results indicated that VPC-18005 could suppress ERG reporter activity without exhibiting overt cytotoxicity.
Figure 2-1 ERG as a drug target and discovery of VPC-18005.

(A) A ribbon representation of the ERG-ETS domain/DNA complex crystal structure [PDB ID: 4IRI] highlighting the winged helix-turn-helix motif of the ETS domain with helix α3 (red) positioned within the major groove of the DNA (cyan). (B) Left: The ERG-ETS domain pocket (shown as grey molecular surface) that was identified by virtual atomic probes (red spheres) and used to screen 3 million small molecules from the ZINC database. The DNA backbone (cyan) is shown for illustration purposes, but not included in virtual screening. Right: Virtual screening pipeline highlighting the steps taken to identify the top candidates to move forward into in vitro experiments (M = million; K = thousands). (C) Western blot analysis of lysates from the indicative PCa cell lines. Levels of ERG (upper panel) are shown relative to alpha-tubulin as a loading control (lower panel). (D) Luciferase activity of lead candidate VPC-18005 (red bar) at 25 μM is shown against other compounds identified from the virtual screening. Data are presented as the mean ± SEM of 4 technical replicates and expressed as a percentage of luciferase activity relative to DMSO control. (E) A box plot illustrates the distribution of docking scores for 3 million small molecules docked at the ERG-ETS pocket, and VPC-18005 scored in the top 0.01%. (F) Dose response effect of VPC-18005 (media concentration 0.1–100 μM) in PNT1B-ERG (open square) and VCaP (closed circle) cells on ERG-mediated luciferase activity, with IC₅₀ (half-maximal inhibitory concentrations) values of 3 μM and 6 μM, respectively. Data are presented as the mean ± SEM of 4 technical
replicates and expressed as a percentage of luciferase activity (Luciferase/Renilla) relative to DMSO control. Data were fitted using GraphPad Prism 6 software to calculate dose response curves of $\log_{10}$ (inhibitor concentration) vs response. (G) Dose response effect of YK-4-279 (media concentration 0.1 – 100 μM) in PNT1B-ERG (open square) and VCaP (closed circle) cells on ERG-mediated luciferase activity, with IC$_{50}$ values of 5 μM and 16 μM, respectively. Data from 4 technical replicates are presented and fit as explained in Figure 1-1F. (H) Cell viability (MTS) of ERG-expressing cells (PNT1B-ERG (circle) and VCaP (square)) and non-ERG expressing cells (PC3 (triangle)) after treatment with 0.2 to 25 μM VPC-18005 (closed red shape) or published inhibitor YK-4-279 (open blue shape) for 72 h. Impact on viability is presented as the mean ± SEM of 3 technical replicates and expressed as a percentage of absorbance relative to DMSO control.
Figure 2-2 VPC-18005 is stable and soluble. Solubility of VPC-18005 was assessed relative to YK-4-279 by diluting stocks (50 mM in DMSO) 1000x in Methanol (standard), PBS, or media. Resulting solutions were clarified by centrifugation and an aliquot of the supernatant was extracted with 2 volumes acetonitrile. VPC-18005 solubility determined as fraction remaining as quantified by LC-PDA-MS versus the methanol solution as standard.
Figure 2-3 VPC-18005 selectively suppresses ERG-mediated transcriptional activity.
(A) PNT1B-ERG (red) and -MOCK (blue) cells were transfected with ETS responsive luciferase reporter and treated with VPC-18005 at the indicated concentration for 48 h. Data are presented as the mean ± SEM of 4 technical replicates and expressed as luciferase: renilla ratio (activity). (B) Luciferase assay was performed on VCaP cells transfected with ETS responsive luciferase reporter and treated with DMSO (0) or 0.1, 1, or 10 nM R1881 in the presence of 25 μM of VPC-18005. Data are presented as the mean ± SEM of 4 technical replicates and expressed as the luciferase: renilla ratio relative to VPC-18005 alone. Lysates from harvested cells were immunoblotted for ERG (upper panel) or α-tubulin (α-tubulin; lower panel) as a loading control. (C) PC3 cells were transfected with AR and the androgen-responsive promoter, ARR3TK-Luc. Cells were treated with DMSO vehicle (0), or 1 nM synthetic androgen (R1881) ± VPC-18005 at 1, 5 or 10 μM, Data are presented as the mean ± SEM of 4 technical replicates and expressed as the percentage luciferase expression (% Activity) relative to R1881. (results were indistinguishable) (D) VCaP cells were treated for 1 h with cycloheximide (10 μM) then cultured for 24 and 48 h with VPC-18005 at the indicated μM concentrations. Lysates were immunoblotted for ERG (upper panels) and β-actin (lower panels) as a loading control.
2.2.2 Direct binding of VPC-18005 to the ERG-ETS domain

The chemical structure of VPC-18005 is depicted in Figure 2-4A. Using computational modeling methods, the predicted binding pose of VPC-18005 was visualized in more detail inside the target pocket on the ERG-ETS domain (Figure 2-4B and C). VPC-18005 is composed of a hydrophobic isopropyl benzyl group at one end and a negatively charged 5′ carboxyl 4-thiazolidanone group on the other end, linked by an azo moiety with conjugated double bonds. Within the binding pocket on the ERG-ETS domain, VPC-18005 is predicted to form a salt bridge with Lys357, hydrogen bonds with Leu313, Trp351 and Tyr372, and hydrophobic interactions with a number of surrounding amino acid residues, including Gln312, Trp314, Tyr371, Tyr372, Lys375, Ile377, Ile395, Ala398, and Leu399 (residue numbering based on ERG isoform 5, UniProt ID: P11308-4; Figure 2-4C).

We utilized NMR spectroscopy to directly assess the binding of VPC-18005 with the ERG-ETS domain. The $^{15}$N-HSQC spectrum of $^{15}$N-labelled protein (100 μM) was assessed in the presence of increasing concentrations of DMSO-solubilized VPC-18005 (Figure 2-5A and 5B), as well as with a DMSO control (Figure 2-5D and 5E). The spectra demonstrated small dose-dependent chemical shifts changes for a number of amide $^{1}$HN,$^{15}$N groups that occurred upon addition of VPC-18005, but not DMSO. A chemical shift perturbation plot with VPC-18005 at 1:10 molar ratio (i.e. 1 mM) showed that protein residues with changes greater than the mean (0.01 ppm) were mostly located along helix α1, helix α3 and strand β3 (Figure 2-4D). These amides cluster around the predicted binding pocket of VPC-18005 (Figure 2-4E), supportive of its binding pose with the ERG protein. Of note, residues with perturbed amide chemical shifts, including Leu313 on helix α1 and Tyr371, Try372, Lys375 on helix α3, modeled to interact with VPC-18005 through hydrogen bonds and hydrophobic interactions, have also been previously shown to be
involved in ERG-DNA interactions. Fitting of the $^{15}$N-HSQC titration curves to a simple 1:1 binding isotherm yielded a $K_D$ value of $\sim$3 mM for the interaction of VPC-18005 with recombinant ERG-ETS domain (Figure 2-5C). To further localize the binding interactions between VPC-18005 and the ERG-ETS domain, the reverse titration was performed. In this case, the $^1$H-NMR spectrum of VPC-18005 was monitored vs. increasing concentrations of recombinant ERG-ETS domain. Several $^1$H nuclei of VPC-18005 exhibited ERG-dependent chemical shift perturbations. These include the hydrogens on the aromatic ring ($^1$H 7.78 and 7.45 ppm), the methyls on the isopropyl group ($^1$H 1.25 ppm) and the conjugated double bond ($^1$H 8.4 ppm) (Figure 2-4F). Due to the spectral overlap with signals from DMSO, perturbations from the CH$_2$ group near the carboxyl group of VPC-18005 could not be determined. Overall, these two complimentary direct binding assay results are consistent with the proposed model for how VPC-18005 binds to the ERG-ETS protein domain at the interface required for DNA interaction.
Figure 2-4 Characterization of VPC-18005 binding to the ERG-ETS domain.

(A) Chemical structure of VPC-18005, in the isomeric form used for docking. The R-isomer is calculated to have the most favorable binding energy. (Molecular weight = 318 g/mol at pH 7). (B) A space-filling representation of the predicted VPC-18005 binding pose within the ERG-ETS domain pocket (orange = carbon, blue = nitrogen, red = oxygen, yellow = sulfur). (C) Protein residues that are predicted to interact with VPC-18005 at the ERG-ETS domain. The red dotted lines indicate hydrogen bonds, and the green lines represent non-polar packing interactions. (D) Amide chemical shift perturbations resulting from the addition of a 10-fold molar excess of VPC-18005 to the ERG-ETS
domain (derived from Figure 2-5). Colored bars denote significant changes (magenta ≥ mean ± standard deviation, cyan ≥ mean). The secondary structure of the EGR-ETS domain is shown at the bottom. (E) Amino acid residues exhibiting significant chemical shift perturbations were mapped to their corresponding locations on the ERG-ETS domain (same color code as in D). (F) $^1$H-NMR monitored titration of VPC-18005 (sharp signals) with increasing concentrations (red through purple) of the ERG-ETS domain (broad signals). Signals from $^1$H nuclei directly bonded to the indicated chemical moieties shift and broaden upon binding the protein.
**Figure 2-5 VPC-18005 directly binds to the ERG-ETS domain.**

(A) Overlaid 15N-HSQC spectra of ERG-ETS domain (100 μM) in the absence (red) and presence of increasing protein:compound molar ratios of DMSO-solubilized VPC-18005 (orange 1:1, green 1:2, cyan 1:4, and purple 1:10). (B) Expanded regions of the overlaid spectra. (C) Fitting of the VPC-18005-induced chemical shift perturbations of the amide 1HN-15N signals of residues 319, 323, 334, 371, and 379 (shown) to a simple 1:1 binding isotherm yielded an average $K_D \sim 3$ mM. (D) Overlaid $^{15}$N-HSQC spectra of ERG-ETS domain (100 μM) in the absence (red) and presence of increasing concentration of DMSO (orange 0.2%, green 0.4%, cyan 0.8%, and purple 2%). (E) Expanded regions of the overlaid spectra, showing no perturbations due to the control titration with DMSO.
2.2.3 VPC-18005 disrupts binding of the ERG-ETS domain to DNA

As there was no obvious effect of VPC-18005 on general cytotoxicity (Figure 2-1H), we assessed whether the impact of VPC-18005 treatment on ETS reporter activity was due to decreased ERG protein stability. Pre-treatment with the protein synthesis inhibitor cycloheximide did not induce ERG degradation after treatment with VPC-18005 at up to 50 μM for 4 h (Figure 2-6A). At extended time points of 24 and 48 h, there was still no observable ERG protein degradation (Figure 2-3D).

Since ERG protein levels were stable in cells treated with VPC-18005 and NMR data supported its direct binding to the ERG-ETS domain, we next assessed whether VPC-18005 could disrupt ERG-DNA binding. Electrophoretic mobility shift assays (EMSA) were performed using purified ERG-ETS domain and a DNA oligonucleotide containing the consensus GGAA recognition motif. The recombinant ERG-ETS domain binds this cognate DNA with a $K_D \approx 1$ nM (Figure 2-7A). VPC-18005, but not DMSO control, exhibited dose-dependent disruption of recombinant ERG-ETS/DNA complex formation (Figure 2-6B) with a $K_I$ value of $\sim 250$ μM (Figure 2-7B). Although indicative of relatively weak binding, this agrees with the $K_D$ value determined for the interaction of the ERG-ETS domain and VPC-18005 using $^{15}$N-HSQC spectroscopy (Figure 2-7C). In contrast, YK-4-279 did not disrupt binding between the ERG-ETS domain and DNA (Figure 2-6B). These results were further confirmed using VCaP nuclear lysate where VPC-18005, but not YK-4-279, disrupted ERG-DNA complex formation (Figure 2-7C–E). Collectively, these results indicate that VPC-18005 can disrupt binding of the ERG protein to the DNA containing ETS-response elements.

A previous study has shown that ERG induces SOX9 gene expression through an AR-regulated enhancer in VCaP. SOX9, a member of the SOX (SRY-related HMG box) family, is a
transcription factor that is required for prostate organogenesis, and its dysregulation has been implicated in cancer pathogenesis\textsuperscript{209}. SOX9 overexpression in an LNCaP xenograft mouse model resulted in increased tumor growth and invasion\textsuperscript{210}, and SOX9 depletion in VCaP was shown to inhibit \textit{in vitro} and \textit{in vivo} invasion\textsuperscript{208}. SOX9 is basally expressed in VCaP cells and elevated following Metribolone (R1881) treatment (Figure 2-8A). Basal and R1881-stimulated SOX9 mRNA and protein expressions were markedly decreased following VPC-18005 treatment. Reduction of ERG and SOX9 expression was also confirmed following siRNA knockdown of ERG in VCaP cells compared to non-specific (NS) siRNA control (Figure 2-8B).
Figure 2-6 VPC-18005 disrupts binding of the ERG-ETS domain to DNA.

(A) Western blot analysis of ERG expression (upper panel) relative to vinculin (lower panel) in lysates from VCaP cells treated for 1 h with cycloheximide and then cultured for 0, 2 or 4 h with VPC-18005 at the indicated concentrations. (B) EMSA shows binding of 4 nM ERG-ETS domain to 1 nM fluorescently-labelled dsDNA alone and in the presence of increasing concentrations of DMSO (top panel, 0.008–17%), VPC-18005 (middle panel, 2 μM–8 mM), and YK-4-279 (lower panel 4 μM–8 mM).
Figure 2-7 VPC-18005 disrupts binding of the ERG-ETS domain to DNA.

(A) EMSA analysis of the binding of from 0.6 pM to 1 mM of purified ERG-ETS domain to a fixed concentration of 1 nM fluorescently labeled dsDNA was performed as detailed in manuscript Materials and Methods. Fitting of densitometric analysis of free (lower band) and bound (upper band) DNA probe data to a 1:1 binding isotherm yielded a KD value of $\sim 1$ nM. (B) Fitting of the data from Figure 3B (VPC-18005; middle panel) to simple competition isotherm yielded a $K_I$ value of $\sim 250$ µM for the interaction of VPC-18005 with the ERG-ETS domain. (C) EMSA analysis of the binding of from 1.1 pg/µl to 1.76 µg/µl of VCaP nuclear lysate to a fixed concentration of 1 nM fluorescently labeled dsDNA. EMSA analysis of the binding of 55 ng/µl ERG-ETS domain to 1 nM fluorescently labeled dsDNA in the presence of 4 µM to 8 mM of (D) VPC-18005, and (E) YK-4-279.
Figure 2-8 VPC-18005 inhibits SOX9 gene expression.

(A) SOX9 and ERG mRNA levels in VCaP cells treated with and without 1 nM R1881 and 25 μM VPC-18005. ***p < 0.001; *p < 0.05, Kruskal–Wallis test with Dunn’s multiple comparison post-hoc test. (B) SOX9 and ERG mRNA levels in VCaP cells transfected with ERG siRNA for 48 h. *p < 0.0001, unpaired t-test; * p = 0.0007, Mann-Whitney-U test. Expression of ERG and SOX9 was normalized to GAPDH. Data points represent experiment performed in triplicate. Error bars indicate standard error of mean for n = 9 values.
2.2.4 VPC-18005 inhibits migration and invasion of ERG-overexpressing cells in vitro

ERG promotes EMT, which enables cells to acquire migratory and invasive characteristics. We have previously shown that PNT1B cells acquired these invasive characteristics when ERG was stably overexpressed. Therefore, we aimed to determine if VPC-18005 was able to affect migration and invasion of these cells. PNT1B-MOCK and -ERG cells were plated into the upper chamber of a double chamber real-time cell analysis system and treated with VPC-18005 after 24 h. As expected, in the absence of VPC-18005, PNT1B-ERG exhibited an increased rate of migration toward the serum-containing bottom chamber compared to the PNT1B-MOCK control (Figure 2-9A and 9B). After 24 h exposure and in comparison to a DMSO control, VPC-18005 (5 μM) significantly reduced the rate of migration of the PNT1B-ERG cells relative to vehicle-treated cells, and the resulting migration rate was indistinguishable from that observed for vehicle treated PNT1B-MOCK cells (Figure 2-9C). In contrast, but consistent with the cytotoxicity results described earlier, treatment with YK-4-279 resulted in cytotoxicity in both cell lines. RWPE prostate cells, engineered to overexpress ERG, were also tested in this assay. No effect on cell viability was observed following treatment with increasing concentrations of VPC-18005 (Figure 2-10A), and VPC-18005 had a moderate effect on RWPE-ERG cell migration compared to MOCK control (Figure 2-10B and C). To further explore this inhibitory effect of VPC-18005, PNT1B-ERG spheroids, pretreated for 24 h with vehicle control or VPC-18005, were submerged in matrix in the presence or absence of treatments, and monitored for 6 days (Figure 2-9D). Analysis of images captured every 2 days revealed that the rate of invasion between day 2 and 6 was significantly reduced in both VPC-18005 (t-test; p = 0.02) and YK-4-279 (p = 0.005) treated cells compared to
vehicle control. These results indicated that VPC-18005 inhibited migration and invasion of ERG-overexpressing cells, without inducing cytotoxicity.

2.2.5 **VPC-18005 inhibits metastasis of ERG-overexpressing cells in vivo**

To determine whether VPC-18005 could affect cell migratory behavior in an animal model, we utilized zebrafish xenotransplantation as a tool to investigate cell extravasation\(^\text{211}\). We first investigated whether PNT1B-MOCK and PNT1B-ERG could disseminate through the zebrafish body (Figure 2-9E). In two separate experiments, fluorescently tagged cells were injected into the yolk sac (10 fish / treatment) and after 5 days PNT1B-ERG could be seen throughout the body of the fish. In contrast, PNT1B-MOCK cells were not detected outside of the yolk sac. The embryos also remained viable when cultured in the presence of up to 50 μM VPC-18005 for 72 h. In contrast, YK-4-279-treated embryos exhibited toxicity at concentrations > 10 μM (Figure 2-9F). Yolk sac-inoculated PNT1B-ERG and VCaP cells were found to become disseminated toward the head and tail of 65 to 70% of embryos, respectively. When cultured in the presence of VPC-18005 at 1 and 10 μM, this percentage of fish with PNT1B-ERG or VCaP dissemination was reduced to 20–30% of inoculated animals (Figure 2-9G). Culturing embryos in YK-4-279 at 1 and 10 μM resulted in yolk sac dissemination in 40–60% of inoculated animals (Figure 2-9G). These assays provide first evidence that small molecules such as VPC-18005 can antagonize the metastatic potential of ERG-expressing prostate cells.
Figure 2-9 VPC-18005 inhibits migration and invasion of prostate cell lines in vitro and in vivo

(A) PNT1B-Mock cells and (B) PNT1B-ERG cells were seeded in the upper chamber of a real-time cell analysis system (xCelligence) and treated with 5 μM VPC-18005 (red line), YK-4-279 (blue line) or 0.01% DMSO (control; black line) at 24 h. The normalized cell index is a measure of the migration of the cells through the pores of the upper chamber and is used as the migration index. Dotted lines represent standard deviations (n = 3). The horizontal dotted red line indicates the level of migration the PNT1B-MOCK cells reached at 48 h in comparison to -ERG cells. (C) Rates of migration were determined by the slopes of the curves between 24–48 h for VPC-18005 (red) (p = 0.031, unpaired t-test) and YK-4-279 (blue) (p < 0.001, unpaired t-test) relative to DMSO control (black). (D) Quantitative analysis of PNT1B-ERG spheroid invasion into the surrounding matrix in the presence or absence of VPC-18005 (red line), YK-4-279 (blue line), or 0.01% DMSO (black line) over the period of 6 days. The rate of invasion between day 2 and 6 was significantly reduced in those cells treated with VPC-18005 (p = 0.02, unpaired t-test) and YK-4-279 (p = 0.005, unpaired t-test) compared to vehicle control. Error bars indicate standard error of the mean (n = 3). (E) Pre-stained PNT1B-Mock and PNT1B-ERG cells were microinjected into the yolk sac (green arrows) of the zebrafish, and the metastatic capability of the cells (white arrows) were detected using confocal microscope at day 2 and day 5. (F) Evaluation of compound toxicity to zebrafish embryos. Zebrafish embryos were treated with increasing concentration of VPC-18005 and YK-4-279 in their water. After 4 days, surviving embryos were counted. (G) Following 5 days of daily treatment, VPC-18005 reduced occurrence of metastasis in zebrafish grafted with PNT1B-ERG and VCaP cells. DMSO versus 1 μM (p = 0.03/0.03, chi square) and 10 μM (p = 0.002/<0.001, chi square) VPC-18005 (PNT1B/VCaP). YK-4-279 was significant only at 10 μM (p = 0.02/0.04, chi square).
Figure 2-10  VPC-18005 suppresses migration, but not growth of RWPE-1-ERG cells.

(A) Cell viability (MTS) of ERG-expressing cells (RWPE-ERG) (closed circle) and non-ERG expressing cells (RWPE-MOCK (open square)) after treatment with 0.2 to 25 μM VPC-18005 (red) or published inhibitor YK-4-279 (blue) for 72 h. Impact on viability is presented as the mean ± SEM of 3 technical replicates and expressed as a percentage of absorbance at 490 nm relative to DMSO control. (B) RWPE-1-Mock or (C) RWPE-1-ERG cells were seeded in the upper chamber of a real-time cell analysis system (xCelligence) and treated with vehicle (0.01% DMSO, black), or 5 μM VPC-18005 (red) or YK-4-279 (blue) for 24 h. The normalized cell index is a measure of the migration of the cells through the pores of the upper chamber and is used as the migration index. Dotted lines represent 1 standard deviation (SD) of the mean migration rate (* = p <0.05)
2.3 Discussion

The ETS family of transcription factors are important targets for drug development because of their strong implications in numerous cancers\(^\text{14}\). However, targeting these transcription factors with small molecules is a challenging task due to their lack of “druggable” active sites. ERG is an important therapeutic target in PCa. We confirmed ERG overexpression in PCa by comparing tumor-specific upregulated genes from three published datasets\(^\text{15,16,101}\) based on a 2-fold differential expression threshold (Figure 2-11). Whereas there were a number of genes dysregulated in each pair-wise dataset comparison (Table 2-2 and Table 2-3), the only upregulated gene common in all three datasets was ERG. This highlights ERG as a potential major influencer of PCa. There are currently several reports describing efforts to target ERG with various agents, but none have resulted in approved therapeutics. These include the use of siRNA\(^\text{104,193}\), shRNA\(^\text{183}\), peptidomimetics\(^\text{195}\) and a small molecule, DB1255 that interacts not with the ERG protein but rather the ETS recognition site on the DNA (GGAA)\(^\text{194}\). In addition, ERG has been targeted indirectly through inhibition of ERG binding proteins including PARP1\(^\text{187}\) and USP9X\(^\text{196}\), as well as via ERG-regulated genes, such as YAP1\(^\text{197}\). As noted earlier, among the various efforts to antagonize ERG, only one small molecule, YK-4-279, has been reported to directly target the ERG protein\(^\text{200}\), but with limitations detailed above.

In contrast to these attempts, we utilized an established rational drug discovery approach\(^\text{58}\) to directly target the DNA-binding interface of the ERG-ETS domain. NMR spectroscopy experiments demonstrated that VPC-18005 binds directly to the ETS domain of the ERG protein. The NMR data were consistent with the in silico modelling of the binding mode of VPC-18005 with ERG. In particular, the perturbation of Tyr371, a key residue required for the ERG-DNA interaction\(^\text{205}\), by VPC-18005 observed in \(^{15}\)N-HSQC spectra provides a possible antagonizing
mechanism. Superimposing VPC-18005 over the DNA at the ERG pocket further revealed the predicted mutually exclusive nature of their binding interfaces (Figure 2-12). Not only does VPC-18005 partially occupy the same interface as the DNA, but the negatively charged carboxyl group is also mapped directly on top of the negatively charged phosphate group of the DNA backbone.

It is also noteworthy that VPC-18005 binds the isolated ERG-ETS domain in vitro with mM affinity, and this is substantially weaker than the nM affinity interaction of ERG with its cognate DNA sequences. However, in the context of the cellular milieu, transcription factor binding to DNA is malleable and susceptible to chemical perturbations\textsuperscript{212,213}. Thus, VPC-18005 is biologically active when present in cell-based assays at μM concentrations. The differences between these in vitro versus in vivo results could arise for numerous reasons spanning from potentially elevated intracellular concentrations of VPC-18005 to highly sensitive effects of this compound on the network of cooperative intermolecular interactions required for transcription. Overall, VPC-18005 is the first reported small molecule inhibitor (SMI) that directly antagonizes the DNA-binding interface on the ERG protein, and it adds to the few successful examples where SMIs have been shown to disrupt protein-DNA interactions of transcription factors\textsuperscript{212}.

The use of a minimal promoter-Renilla luciferase control suggested that inhibition from VPC-18005 was specific to the ERG-responsive reporter and not a non-specific effect on general transcription (Figure 2-14B). However, due to the sequence conservation at the ETS domain, it is expected that anti-ERG compounds such as VPC-18005 will have the potential to bind and inhibit other ETS factors. Indeed, preliminary NMR spectroscopic experiments revealed that VPC-18005 also interacts with the ETS domains of PU.1 and ETV4 (Figure 2-15). Nevertheless, as demonstrated here, VPC-18005 is non-toxic at active concentrations. It should also be noted that many of these ETS factors are oncogenic and have been implicated in a wide spectrum of cancers\textsuperscript{14}. 
Although future development of new VPC-18005 derivatives can be prioritized based on their selective binding towards ERG, those showing promiscuous or increased specificities for alternative ETS factors should be considered as lead therapeutics towards cancers linked to those factors.

VPC-18005 inhibits ERG transcriptional activity in a dose dependent manner. Additionally, VPC-18005 can inhibit the expression of an ERG-regulated gene, SOX9, which has been previously shown to stimulate PCa invasion. VPC-18005 did not affect cell viability, but did influence cell motility, leading to reduced migration/invasion of cells in vitro and in a zebrafish xenograft model. As cancer cell death is a measure of toxicity and cancer cell immobility is a measure of metastasis prevention, our study supports non-toxic anti-metastatic applications of VPC-18005 and its derivatives. Future clinical studies of such anti-ERG drugs can be modelled based on previous clinical trials for anti-metastatic drugs to target patients with metastatic disease of low burden. Key indicators of anti-metastatic drug efficacy in patients include inhibition of further metastasis/invasion of tissues, decreased skeletal related conditions, decreased pain/narcotic use, increased survival (decreased end organ destruction), and decrease of circulating tumor cells.

In summary, these results demonstrate proof-of-principal that small molecule targeting of the ERG-ETS domain can suppress transcriptional activity and reverse transformed characteristics of PCa’s aberrantly expressing ERG. The current lead compound, VPC-18005, inhibited ERG with low micromolar concentrations at in vitro and in vivo experiments. In murine pharmacodynamics and toxicology studies, VPC-18005 is soluble, stable, and orally bioavailable, and does not exhibit general toxicity at single doses of up to 500 mg/kg, and after a 4 week BID at 150 mg/kg trial (Figure 2-16). We anticipate that future medicinal chemistry (medchem) efforts will improve its
activity into a sub-micromolar or nanomolar range. Indeed, our initial medchem development has identified additional derivatives through chemical similarities and modifications of the VPC-18005 scaffold (Table 2-4). Of these candidates VPC-18065 and 18098, with terminal moieties that are more hydrophobic, demonstrated slightly better IC\textsubscript{50} values (2 μM and 1 μM respectively in luciferase assays) compared to VPC-18005 (Figure 2-13). The removal of the carboxyl group in VPC-18100 resulted in the loss of inhibition in the luciferase reporter assays, as we expected given that the carboxylate is predicted to form a salt-bridge with the nearby Lys357. Although the modifications tested to date have not yet resulted in significant sub-micromolar activity, these derivatives do provide a working structure-activity relationship that will guide future medchem efforts. DNA binding domains of transcription factors are often conserved and exhibit low rates of mutations as a structural compromise is likely to translate into a loss of function \textsuperscript{216}. Targeting such DNA-interacting regions may increase the value of the corresponding drugs due to less mutation-driven resistance, but also makes direct assessment of binding specificity by mutagenesis challenging.

PCa, one of the most common malignancies in men, is treated by surgery and radiation at the early stage, but eventually progresses to advanced forms that are managed primarily by the androgen deprivation therapy (ADT). The effectiveness of ADT is only temporary due to resistance mechanisms related to aberrant androgen production and mutations in the androgen receptor\textsuperscript{217}. Recent studies not only established ERG as a critical drug target in PCa\textsuperscript{103}, but also reported on ERG feed-forward regulation. This supports the notion that despite initial dependence on the androgen receptor, ERG expression can eventually become self-driven and resistant against ADT \textsuperscript{187}. Thus, anti-ERG drug prototypes such as VPC-18005 developed through rational drug design as reported here can specifically target the malignant transformation and metastasis driven
by the ERG and are not susceptible to current PCa treatment limitations such as drug resistance against anti-androgens and side effects from ADT. With the availability of non-invasive urine tests for ERG detection 179, future anti-ERG drugs can be specifically prescribed to the 50% of PCa patients who are ERG-positive and pave the way for precision medicine.
ERG is overexpressed in prostate cancer

A) A Venn diagram that shows the number of upregulated genes from each of the three gene expression datasets: Vancouver Prostate Centre (VPC), Memorial Sloan-Kettering Cancer Center (MSKCC), and The Cancer Genome Atlas (TCGA), based on a bioinformatic protocol (see Materials and Methods). ERG is the only overexpressed gene common to the three datasets. (B) The fold changes of ERG gene expression in PCa tumor samples, compared to normal samples, range from 2.66 to 3.29.
Figure 2-12 Mutually exclusive binding of VPC-18005 and DNA with the ERG-ETS domain.
The position of the carboxyl group in VPC-18005 (orange = carbon, blue = nitrogen, red = oxygen, yellow = sulfur) is predicted to coincide with that of a phosphate group on the DNA backbone (cyan ribbons and sticks). Thus, binding of VPC-18005 or DNA with the ERG-ETS domain is expected to be mutually exclusive. Whereas the lower portion of the VPC-18005 chemical structure occupies the same general area as the DNA, the upper portion, consisted of the isopropyl moiety and aromatic ring, extends further into the pocket (shown in grey net). The ERG-ETS protein structure is rotated −90 degree on the vertical axis, compared to those shown in Figure 2-4.
Figure 2-13  Preliminary SAR studies using derivatives of VPC-18005. Modifications of the isopropyl moiety (VPC-18005) into tert-butyl (VPC-18065) and cyclobutyl (VPC-18098) improved the IC50 values in the luciferase reporter assays in PNT1B-ERG cells. Removal of the carboxyl moiety (VPC-18100) resulted in the loss of activity. Progressive differences between the derivatives are highlighted in green.
Figure 2-14
(A) Comparison of empty vector vs. ETS responsive vector and the effect of DMSO control and VPC-18005 accordingly. Data points represent the mean of quadruplicate values. Error bars indicate standard error of mean for \( n = 4 \) values. (B) Raw renilla (pRL-tk) luciferase readings from a dose response experiment where PNT1B-ERG cells were treated with 0.1 – 100 \( \mu \text{M} \) of VPC-18005. (C) General scheme of chemical synthesis for VPC-18005.
Figure 2-15 NMR analysis for titration of VPC-18005 with ETV4 or PU1 ETS domain.

Amide chemical shift perturbations resulting from the addition of a 10-fold molar excess of VPC-18005 to the ETV4 or PU.1 domain. Colored bars denote significant changes (magenta ≥ mean + standard deviation, cyan ≥ mean). The secondary structure of the EGR-ETS domain is shown at the bottom.
Figure 2-16 Preliminary data for in vivo study of VPC-18005

(A) In a preliminary study to evaluate in vivo toxicity and establish working oral (PO) doses, Nu/nu mice (n = 4) received a single PO dose of VPC-18005 (100 mg/kg or 500 mg/kg) dissolved in cyclodextrin. Blood was collected directly after injection and at the times indicated. Serum concentration of VPC-18005 was determined for each time point. (B) In a 4 week “repeated dose” study, body weight was measured daily for mice (n = 4) that received a daily dose of VPC-18005 (150 mg/kg BID) dissolved 50% ethanol (PO) or 50% ethanol (Vehicle). Change in weight of each animal was normalized to starting pretreatment weight. (C) Cytotoxicity serum markers of vehicle and VPC-18005 treated animals described in B) were determined and presented as average (Avg. Conc.) and standard deviation (SD) with corresponding paired t-test p values. LC-MS analysis of serum indicators of major organ toxicity (albumin (ALB), alkaline phosphatase (ALP) and alanine aminotransferase (ALT) as indicators of liver dysfunction, amylase (AMY) to test for acute pancreatitis, total bilirubin (TBIL) to assess jaundice, and blood urea nitrogen (BUN) to assess kidney function from endpoint samples indicated no significant impact of VPC-18005 treatment on their levels.
2.4 Materials and Methods

2.4.1 In silico modeling and virtual screening

The published ERG-ETS domain X-ray crystal structure (PDB: 4IRG) was subjected to the Site Finder algorithm, implemented in the Molecular Operating Environment (MOE), which used virtual atomic probes to search the protein surface for suitable small molecule binding pockets. The crystal structure of an ERG/DNA complex (PDB: 4IRI) was used to define the ERG-DNA interface. The top-ranked pocket was identified and used for the subsequent virtual screening. Before molecular docking, the ERG-ETS domain structural model was prepared by using the Protein Preparation Wizard module of the Maestro v9.3 program from the Schrodinger 2012 software suite. The docking grid was centered at the pocket composed of the following amino acids: Pro306, Gly307, Gln310, Ile311, Gln312, Leu313, Trp314, Trp351, Lys355, Met360, Lys364, Leu365, Ala368, Tyr371, Tyr372, Lys375, Ile377, Ile395, Ala398, Leu399 (residue numbering based on ERG isoform 5, UniProt ID: P11308-4). A total of 19,607,722 (~ 20 million) small molecule structures were downloaded from the ZINC database version 12. Among the 20 million set, a total of 2,990,102 (~ 3 million) molecules that possess the following lead-like and drug-like properties were extracted for molecular docking: molecular weight between 250 and 400 Da, logP ≤ 5, hydrogen-bond donors ≤ 5, hydrogen-bond acceptors ≤ 10, number of rotatable bonds ≤ 10, and number of rings ≤ 4. Each molecule was given its expected protonation state at pH 7 and energy-minimized under the MMFF94x (solvation: Born) force field using MOE. Each molecule was docked into the previously defined docking grid on the ERG-ETS domain protein model, using the Glide program (Small-Molecule Drug Discovery Suite, version 5.8, Schrödinger, LLC, New York, NY, 2012). Standard Precision with all other parameters set to default. The top 1% (~30,000 molecules), as ranked by the docking scores calculated based on interaction forces.
including hydrogen bonds and hydrophobic interactions, were selected to advance into the next stage of virtual screening. Within this set, a predicted pK\text{a} was calculated for each molecule using a custom MOE SVL script, and ligand efficiency was calculated using Glide. In addition, this set of 30,000 molecules was re-docked into the same pocket, using the eHiTs docking program \textsuperscript{219}. A root-mean-square deviation (RMSD) was calculated between the docking poses from Glide and eHiTs for each molecule. A consensus scoring (voting) method was used each compound received one vote from each of the following criteria met: 1) top 20% pK\text{a} values, 2) top 20% ligand efficiency values, and 3) top 20% eHiTs docking scores and 4) RMSD ≤3 Å. The top 3,000 molecules, as ranked by the number of votes, were selected for the final stage of selection. During this step, the chemical structure of each molecule within the predicted ERG-ETS binding pocket was manually examined using the 3D visual environment in MOE. Preference was given to compounds with favorable binding poses and interactions with the surrounding amino acid residues. Molecules were removed from the selection if they contain problematic or promiscuous moieties. In addition to manual examination, the FAFDrugs program \textsuperscript{220} was used to assist identification of such problematic groups. A total of 48 compounds were selected for testing. MOE and MarvinSketch were used to visualize and represent the protein models and chemical structures. Chemical similarity searches based on the Tanimoto coefficient was performed on the hit compound VPC-18005 (prepared as detailed in Materials and Methods) in the ZINC database, with additional medchem derivatives designed using MOE.

### 2.4.2 Cell Culture

VCaP (CRL-2876) and PC3 (CRL-1435) human prostate carcinoma cells were obtained from the American Type Culture Collection (ATCC). VCaP cells harbor an endogenous \textit{TMPRSS2-ERG} gene fusion, whereas PC3 cells do not, but do express the ETS family member
ETV4. The immortalized prostatic epithelial cell line, PNT1B\textsuperscript{206} was purchased from ATCC. PNT1B-Mock, PNT1B-ERG, RWPE-Mock, and RWPE-ERG cells are lineage-matched control and ERG-expressing prostatic epithelial lines generated in house\textsuperscript{106}. PC3 cells were maintained in RPMI 1640 medium (Life Technologies) supplemented with 5\% (v/v) fetal bovine serum (FBS). VCaP cells were maintained in low bicarbonate DMEM (ATCC) supplemented with 10\% FBS. PNT1B-Mock and –ERG cells were maintained in DMEM (Life Technologies) supplemented with 10\% FBS and under blasticidin selection. Cells were grown in a humidified, 5\% CO\textsubscript{2} incubator at 37\°C.

2.4.3 Western blot

Cells were lysed on ice with RIPA buffer containing a protease inhibitor cocktail (Pierce). Primary antibodies: ERG (1:1,000, EPR3864(2), Abcam), α-Tubulin (1:20,000, Millipore), Vinculin (1:1,000, Abcam). Immunoreactivity was detected with the use of the goat anti-rabbit or rabbit anti-mouse horseradish peroxidase (HPR)–conjugated secondary antibody (1:10,000) (Santa Cruz), and visualization was achieved by chemiluminescence (Pierce). To inhibit protein synthesis, 10 μM cycloheximide was added for 1 h and then replaced with treatment medium for indicated time frame.

2.4.4 Dual reporter luciferase assay

All of the compounds selected from the virtual screening were tested in a luciferase-based ERG-responsive reporter assay, using two ERG-overexpressing cell lines, VCaP and PNT1B-ERG, previously developed in house\textsuperscript{106}. Cells (3000) in 150 μL per well of a 96 well plate were seeded and after a 24 h incubation were transfected with 50 ng of an Endoglin E3 promoter-derived ETS-responsive firefly luciferase reporter (–507/–280 of (E3) promoter\textsuperscript{207} inserted into luciferase reporter vector (Signosis), ARR\textsubscript{3}tk-luc\textsuperscript{58}, and 5 ng of the Renilla luciferase reporter (pRL-tk,
Promega) using TransIT 20/20 transfection reagent (Mirus, USA). After 16 h incubation, treatment media was added for further 48 h. Firefly and Renilla luciferase activities were measured using a TECAN M200Pro plate reader. Comparison of empty vector versus ETS responsive reporter demonstrates activation only in the presence of the ETS responsive sequence (Figure 2-14A). Data were normalized first to Renilla luciferase and then to the DMSO-media control on each plate, unless otherwise stated. Initial hit compounds were identified as those with an average normalized luciferase reading (firefly luciferase/Renilla luciferase readings) that is 60% or less of the average normalized luciferase reading of the DMSO-media control (i.e. 40% or more reduction of luciferase activity) at 10 μM. Representative example of raw Renilla luciferase readings presented in Figure 2-14B. The luciferase assays were repeated for each lead compound under multiple concentrations (0.1 to 100 μM) to establish a dose-dependent response and an IC<sub>50</sub> value. AR reporter assay was performed as previously described<sup>58</sup>.

2.4.5  **Proliferation/ cell viability assay**

3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt (MTS). Cells were seeded at a density of 3000 cells per well (except VCaP at 20,000/well) in 100 μL of appropriate media in 96 well culture dishes. Twenty-four hours later, 100 μL of medium containing vehicle control or compounds. Each treatment was prepared in triplicate. After a 72 h treatment, cellular viability was assessed using CellTiter 96<sup>®</sup> Aqueous One Solution Cell Proliferation Assay reagent (Promega) according to the manufacturer’s instructions. Values were normalized to the DMSO control.

2.4.6  **NMR spectroscopy**

ERG-ETS domain expression and purification. A pET28a plasmid encoding residues 307–400 of the ERG-ETS domain was expressed in *E. coli* BL21 (DE3). Cultures of 1 L were grown
at 37°C in M9 media supplied with 3 gm/L ¹³C₆-glucose and/or 1 gm/L ¹⁵NH₄Cl. Cells were allowed to grow to O.D.₆₀₀ = 0.6 and protein expression was induced by adding 1 mM IPTG. After an induction time of 4 h, cells were harvested by centrifugation and stored at –80°C for at least 1 round of freeze/thaw. Cells were resuspended in 40 mL of lysis buffer for every 1 L of culture. Cells were lysed by passing through 5 rounds of homogenization and 10 mins of sonication. The cell lysate centrifuged at 15k rpm for 1 hr, and the supernatant subjected to nickel column purification. The column was washed using 25 mM imidazole (50 mM phosphate, 1 M NaCl, pH 7.4) and proteins were eluted with 1 M imidazole. Fractions containing the ETS domain were confirmed by SDS-PAGE and pooled. The His₆-tag was cleaved by thrombin and the tag-free sample was concentrated to 2 mL and subjected to S75 size exclusion chromatography. Fractions were checked by SDS-PAGE and those containing the pure sample were pooled and concentrated. The protein ample was dialyzed to NMR buffer (20 mM sodium phosphate, 150 mM NaCl, 2 mM DTT, 0.1 mM EDTA, pH 6.5) for all NMR experiments. NMR spectral assignments. NMR data were recorded at 25 or 28°C on cryoprobe-equipped 850 MHz Bruker Avance III spectrometer. Data were processed and analyzed using NMRpipe and Sparky. Signals from backbone and sidechain ¹H, ¹³C, and ¹⁵N nuclei were assigned by standard multidimensional heteronuclear correlation experiments. NMR-monitored titrations. Interactions of compounds with the ERG-ETS domain were monitored via sensitivity-enhanced ¹⁵N-HSQC spectra. Experiments involved titrating unlabeled DMSO-solubilized compound or control DMSO into ¹⁵N-labeled ERG-ETS domain. Chemical shift perturbations were calculated from the combined amide ¹H and ¹⁵N shift changes as Δδ = [(0.2 × Δδₙ)² + (Δδₕ)²]¹/². Reciprocal titrations were carried out using ¹H-NMR to monitor the effects of progressively adding unlabeled protein to a sample of VPC-18005 (180
μM) in 20 mM phosphate, 150 mM NaCl, 2 mM DTT, 0.1 mM EDTA, pH 6.5. The signal from water was suppressed by pre-saturation.

2.4.7 Electrophoretic mobility shift assay (EMSA)

Purified ERG-ETS domain (see NMR spectroscopy) was stored in buffer (20 mM sodium phosphate, 200 mM NaCl, 2 mM DTT, 0.1 mM EDTA, pH 6.5). To prepare the probe for the gel shift assay, equal amounts (200 nM) of Alexa-488 fluorophore-labeled DNAs (5′-CGGCAAGCCGGAAGTGAGTG-3′ and its complement) were mixed, heated to 95°C for 30 minutes, and then slowly cooled to 25°C in several hours. An initial gel shift assay was performed by titrating constant 1 nM labeled dsDNA with ERG at concentrations spanning 0.3 pM to 0.5 μM. Glycerol (3%) and 0.2 mg/mL bovine serum albumin (BSA) were included in the reaction mixture. After incubated at room temperature for 1 hr, samples were loaded on to 10% polyacrylamide native gel, and electrophoresed at 10°C. The gel was scanned with Typhoon 9200 Imager equipped with blue laser to excite at 490 nm and fluorescence was measured at 520 nm. The scanned image was analyzed with Image J. Non-linear least squares fitting (GraphPad Prism) of the titration data to a 1:1 binding isotherm yielded the equilibrium dissociation constant (K_D value ~ 1 nM) for the ERG-ETS domain interaction with DNA. The binding isotherm equation is f_{b,i} = [ERG]/([ERG]_i + K_D) where [ERG]_i is the total concentration of the ERG-ETS domain (a valid approximation as K_D > 1 nM total dsDNA) at each titration point (i), and the fraction bound, f_{b,i} was calculated as the intensity of the bound DNA band at that point relative to the intensity with saturating 0.5 μM protein. The result of this initial study was used to set the molar ratio of ERG-ETS domain:DNA in subsequent competition assays with VPC-18005. For these assays, 4 nM of the ERG-ETS domain was mixed with 1 nM of fluorophore-labeled dsDNA, titrated with VPC-18005 (diluted from a DMSO stock) and analyzed by the same EMSA protocol. The data were fit to the equation.
for competitive binding, \( f_{b,i} = \frac{[\text{ERG}]}{([\text{ERG}] + K_D[1 + [\text{VPC-18005}]/K_I])} \), where \( K_I \) is the inhibitor dissociation constant and the fraction bound, \( f_{b,i} \), was calculated the intensity of the bound DNA band at each titration point relative to that without added VPC-18005. A control experiment was carried out by titrating with equivalent quantities of DMSO.

For experiments involving ERG from VCaP cells, VCaP nuclear protein was extracted using CellLytic NuCLEAR Extraction Kit (Sigma). An initial gel shift assay was performed by titrating constant 1 nM labeled dsDNA nuclear extract at concentrations spanning 1.1 pg/μl to 1.76 μg/μl (Figure 2-7C). For subsequent assays, 55 ng/μl of the nuclear extract was mixed with 1 nM of fluorophore-labeled dsDNA, titrated with VPC-18005 (diluted from a DMSO stock) and analyzed by the same EMSA protocol.

**2.4.8 Analyses of gene expression**

Total RNA was extracted from VCaP cells with the use of RNeasy Plus kit (Qiagen). Reverse transcription was performed with the use of the iScript First-Strand cDNA Synthesis Kit (Bio-Rad Laboratories) with 100 ng total RNA used as template. Real time reverse-transcription (RT) polymerase chain reaction (PCR) primers for ERG synthesized by IDT (forward, 5′-CGCAGATTATCGT GCCAGCAGAT-3′; reverse, 5′-CCATATTCTTTCACC GCCCACTCC-3′) and SOX9 (Quantitect primer assay, Qiagen). Real-time quantitative RT-PCR was performed in triplicate for each sample with the use of the ABI ViiA7 QPCR thermocycler. In each reaction, 1 μL cDNA, 1 μL forward and reverse primers (or 1 μL of Quantitect primers), and 6 μL Sybr Green Master Mix (Applied Biosystems) were added with water to make a final volume of 12 μl. All primers were used at a concentration of 5 μmol/l. PCR cycling conditions were 95°C for 10 minutes followed by 40 cycles of 95°C for 15 seconds, 60°C for 1 min. Data was normalized to reference genes: GAPDH (forward, 5′-CCATATTCTTTCACC GCCCACTCC-3′; reverse, 5′-
GGCATGGACTGTGGTCATGAG -3′) The $2^{-\Delta\Delta CT}$ method was used to compare samples. PCR product specificity was validated with the use of a melt curve.

### 2.4.9 Real time cell analysis (xCELLigence)

Cell migration was monitored using CIM-16 migration plates via the xCELLigence platform (ACEA). FBS-supplemented media (160 μL) was added to the lower chamber of the plate and incubated at RT for 30 min. The upper chamber was then mounted and 30 μL of serum free media (SFM) was added to each well and left to equilibrate in the incubator for 1 h at 37°C. After the incubation, a background reading was taken for each well. PNT1B-ERG or –MOCK cells, cultured for 24 h in SFM, were seeded into the wells of the upper chamber at 30,000 cells per well and after 24 h 100 μL of desired treatment was added (vehicle control, VPC-18005, and YK-4-279). Real time readings of cell index values were recorded initially every 5 min until the end of the experiment (48 hr).

### 2.4.10 Spheroid invasion assay

3D Spheroid BME Cell Invasion Assay (Trevigen) was performed as per manufacturer’s instructions. Briefly, 5,000 PNT1B-ERG cells and 5 μL of ECM were prepared in growth media to a total volume of 50 μL and seeded in 3D culture qualified 96 well spheroid formation plate and incubated at 37°C for 72 hr. Spheroids were pre-treated with VPC-18005 or DMSO for 24 h after which 50 μL gel invasion matrix was added. Spheroids were then incubated at 37°C for 3 to 7 days and photographed using Zeiss AxioObserver Z1 microscope in each well on the day of invasion mix addition and every two days following. Spheroids were retreated with 50 μL of vehicle control or compound after 72 hr.
2.4.11 Zebrafish

Research was carried in accordance with protocols compliant to the Canadian Council on Animal Care and with the approval of the Animal Care Committee at the University of British Columbia. The wildtype zebrafish strain was maintained in aquaria according to standard protocols. Embryos were generated by natural pair-wise mating’s and raised at 28.5°C on a 14 h light/10 h dark cycle in a 100 mm² petri dish containing aquarium water. Phenylthiourea (0.2 mM PTU, Sigma) was added to the embryos at 10 h post-fertilization (hpf) to prevent pigment formation.

Yolk sac dissemination assay. PCa cell lines were fluorescently labelled the day before microinjection with 1.5 μM of CellTracker CM-Dil dye (Life Technologies) as per manufacturer’s instructions. Wild-type embryos were dechorionated at 2 dpf. Following anaesthetization with tricane, approximately 50–70 cancer cells were microinjected into the yolk sac. Embryos were then transferred to 100 mm² plates that contained aquaria water with added PTU and VPC-18005, YK-4-279 or DMSO control. Embryos were visually assessed for presence of xenograph. Those embryos that did not contain cells were removed from the experiment. Embryos were kept at 35°C for the duration of the experiment. Approximately, 50 fish were injected per cell line and metastasis was determined on Day 4 and 5 by observation using the Zeiss Axio Observer microscope (5X objective) controlled with Zen 2012 software. Fixed (dead) cells were used as a control to ensure that the dissemination observed was not due to yolk sac absorption.

2.4.12 Statistics

Data are presented as mean ± standard error of the mean (SEM) unless indicated otherwise. The Kruskal–Wallis test with Dunn’s Multiple Comparison post hoc test, chi squared test, two-way ANOVA followed by Fisher’s LSD post hoc test, and t-test were used for analyses as indicated.
in the respective figure legends. $p < 0.05$ was considered significant. Statistical analyses were performed with the use of GraphPad Instat or GraphPad Prism 6 (GraphPad Software, Inc.).

2.4.13 Compound solubility and stability

Stock solutions of compounds at 50 mM in dimethyl sulfoxide (DMSO) were diluted 1,000× into methanol (MeOH), RPMI + 5% charcoal stripped serum (CSS) (media), and phosphate buffered saline (PBS) and vortex mixed for 1 hr, 800 rpm at room temperature (RT). The resulting solutions were centrifuged at 20,000 g for 5 min (RT) and saturated supernatants were transferred to fresh Eppendorf tubes. Saturated PBS samples were further diluted with an equal volume of PBS. Aliquots of these solutions were analyzed, and the remainder stored at RT in the dark. Aliquots taken at later time points were vortex mixed for 1 h prior to sampling. MeOH and diluted PBS samples required no further processing; media samples were extracted with two volumes acetonitrile (ACN) and centrifuged at 20,000 g for 5 min. These MeOH, and diluted media and PBS samples, were analyzed using an Acquity UPLC coupled in series with an eLambda PDA and a Quattro Premier (Waters). A 100 mm BEH C18, 1.7 µ column (Waters) was used for separations with a 10–95% acetonitrile (ACN) gradient from 0.2–7 min followed by a 1 min 95% ACN flush and 2 min re-equilibration for a 10 min run length (0.1% formic acid present throughout). Wavelengths from 210–800 nm at 1.2 nm resolution and 2 points/sec were collected with the PDA. The sampler was maintained at RT and all MS data was collected in ES+ scan or single ion recording (SIR) mode at unit resolution with the following instrument parameters: capillary, 3.0 kV; extractor and RF lens, 3 V and 0.1 V; cone, 40 V; source and desolvation temperatures, 120°C and 350°C; desolvation and cone (N2) flow, 900 L/hr and 50 L/hr. The m/z for SIR functions were selected from MeOH scan datasets.
Quanlynx (Waters) was used for analysis of data, using extracted wavelength chromatograms selected for best signal to noise for PDA data and SIR for MS data. All compounds dissolved well in MeOH and these were used for calibration purposes with slopes forced through the origin. OD data was used in most cases with MS data mainly for PBS samples; SIR data was calibrated by applying the SIR/OD ratio from corresponding media samples where less saturation of MS data is expected. This rudimentary method is useful to 50 µM, performs well for solubility and relative stability at higher concentrations and gives reasonable estimates when the use of MS endpoints is needed.

2.4.14 Cell cycle analysis

Cells were detached by treatment with Accutase (Gibco), then underwent APC BrdU Flow Kit protocol (BD Pharmingen). Cells were analyzed on a FACSCanto™ II (BD Biosciences). Data was analyzed using FlowJo software (TreeStar, USA). Biological replicates were analyzed statistically by Two-Way ANOVA.

2.4.15 Proliferation/cell viability assay (Incucyte generated growth curves)

VCaP cells (20,000 cells/well) were plated in a 96 well plate. After 24 h, plates were treated with vehicle control, VPC-18005 or YK-4-279 at the indicated concentrations. Growth curves were constructed by imaging plates using the Incucyte system (Essen Instruments), where the growth curves were built from real-time confluence measurements acquired during round-the-clock kinetic imaging for 7 days.

2.4.16 Bioinformatics and statistical analyses on gene expression datasets from PCa patients

The gene expression datasets included 26 PCa and 5 normal patient samples from Vancouver Prostate Centre (VPC) 16, 150 PCa and 29 normal patient samples from Memorial
Sloan-Kettering Cancer Center (MSKCC)\textsuperscript{15}, and 498 PCa and 52 normal patient samples from The Cancer Genome Atlas (TCGA)\textsuperscript{101}. A list of upregulated genes were identified from each dataset by the following steps: 1) log2 transformation; 2) two sample t-test between tumor and normal samples; 3) multiple testing correction on p-values; 4) selection of genes with corrected (adjusted) p-values < 0.05; and 5) among those with significant p-values, selection of genes with fold-change $\geq 2$ (tumor vs. normal).

2.4.17 Chemical synthesis of VPC-18005 (Figure 2-14C)

All reagents and solvents were purchased from commercial suppliers and used without further purification unless otherwise stated. The reactions were monitored by thin layer chromatography (TLC) on pre-coated silica gel F254 plates (Sigma-Aldrich) with a UV indicator using ethylacetate/hexane (1:2 v/v). Yields were of purified product were not optimized. The purities of the newly synthesized compounds were determined by LC-MS analysis using an Agilent 1100 LC system. The compound solution was injected into the ionization source operating positive and negative modes with a mobile phase acetonitrile/water/formic acid (50:50:0.1% v/v) at 1.0 mL/min. The instrument was externally calibrated for the mass range m/z 100 to 650. The $^1$H-NMR spectra were measured on a Varian GEMINI 2000 NMR spectrometer system with working frequency of 400 MHz. Chemical shifts $\delta$ are given in ppm, and the following abbreviations are used: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m), and broad singlet (br s). VPC-18005: 2-(((Z)-2-(((Z)-4-isopropylbenzylidene) hydrazono)-4-oxothiazolidin-5-yl)acetic acid (4d). To a stirred solution of 4-isopropylbenzaldehyde (1d) (563 mg, 3.8 mmol) in PhMe (2 mL) and DMF (2 mL) were added thiosemicarbazide (2) (290 mg, 3.2 mmol) and p-TsOH acid (5 mg, 0.03 mmol). The reaction mixture was heated in stirred microwave vial for 10 min at 90oC. After formation of thiosemicarbazone derivate and testing by TLC, maleic anhydride
(3) (343 mg, 3.5 mmol) was added, and the reaction mixture was heated for 40 min at 110°C in the microwave. Recrystallized from AcOH yielded (4d, VPC-18005) (300 mg, 29% yield) as a white solid with 99% purity by LC/MS. 1H-NMR (DMSO-d6, 400 MHz): 1.20–1.22 (6H, d), 2.69–2.76 (1H, m), 2.90–2.94 (2H, m), 4.25–4.28 (1H, d), 7.31–7.33 (2H, d), 7.66–7.68 (2H, d), and 8.34 (1H, s). MS (ESI) m/z (M + H)+ calculated for C15H17N3O3S: 319.4, found: 320.2. The final product is racemic and has several possible isomeric forms that have not been experimentally defined.
Table 2-1 Analysis of the serum concentration curves produced in Supp. Figure 2.16A

<table>
<thead>
<tr>
<th>Dosing</th>
<th>mg/kg</th>
<th>hr</th>
<th>µM</th>
<th>interval hr</th>
<th>µM*hr</th>
<th>L/hr/kg</th>
<th>L/kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>PO</td>
<td>100</td>
<td>1</td>
<td>7.36</td>
<td>(6-24)</td>
<td>2.38</td>
<td>85.4</td>
<td>3.67</td>
</tr>
<tr>
<td>n=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1-6)</td>
<td></td>
<td>11.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dosing</th>
<th>mg/kg</th>
<th>hr</th>
<th>µM</th>
<th>interval hr</th>
<th>µM*hr</th>
<th>L/hr/kg</th>
<th>L/kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>PO</td>
<td>500</td>
<td>9</td>
<td>27.1</td>
<td>(9-72)</td>
<td>5.73</td>
<td>385.3</td>
<td>4.07</td>
</tr>
<tr>
<td>n=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1-6)</td>
<td></td>
<td>-5.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2-2 Overexpressed genes common in the VPC and TCGA gene expression sets.

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>NCBI Gene ID</th>
<th>P-value (VPC)</th>
<th>Adjusted P-value (VPC)</th>
<th>Fold change (VPC)</th>
<th>P-value (TCGA)</th>
<th>Adjusted value (TCGA)</th>
<th>P-value (TCGA)</th>
<th>Fold change (TCGA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACRV1</td>
<td>56</td>
<td>2.52E-04</td>
<td>2.64E-02</td>
<td>3.46</td>
<td>3.94E-05</td>
<td>1.03E-04</td>
<td>2.49</td>
<td></td>
</tr>
<tr>
<td>ANGPTL3</td>
<td>27329</td>
<td>6.15E-04</td>
<td>4.33E-02</td>
<td>3.55</td>
<td>2.40E-12</td>
<td>2.19E-11</td>
<td>3.61</td>
<td></td>
</tr>
<tr>
<td>ASF1B</td>
<td>55723</td>
<td>3.04E-05</td>
<td>8.37E-03</td>
<td>2.89</td>
<td>2.10E-10</td>
<td>1.32E-09</td>
<td>2.21</td>
<td></td>
</tr>
<tr>
<td>B4GALNT4</td>
<td>338707</td>
<td>3.85E-07</td>
<td>8.52E-04</td>
<td>2.87</td>
<td>2.81E-21</td>
<td>2.32E-19</td>
<td>3.39</td>
<td></td>
</tr>
<tr>
<td>BARX1</td>
<td>56033</td>
<td>2.42E-04</td>
<td>2.58E-02</td>
<td>6.22</td>
<td>3.80E-13</td>
<td>4.15E-12</td>
<td>2.68</td>
<td></td>
</tr>
<tr>
<td>BUB1</td>
<td>699</td>
<td>1.89E-05</td>
<td>6.41E-03</td>
<td>2.91</td>
<td>9.15E-11</td>
<td>6.12E-10</td>
<td>2.74</td>
<td></td>
</tr>
<tr>
<td>BUB1B</td>
<td>701</td>
<td>3.31E-04</td>
<td>3.11E-02</td>
<td>3.32</td>
<td>7.82E-15</td>
<td>1.24E-13</td>
<td>3.54</td>
<td></td>
</tr>
<tr>
<td>C17orf55</td>
<td>284185</td>
<td>7.88E-04</td>
<td>4.95E-02</td>
<td>2.80</td>
<td>3.65E-07</td>
<td>1.31E-06</td>
<td>2.16</td>
<td></td>
</tr>
<tr>
<td>CCNA2</td>
<td>890</td>
<td>2.59E-06</td>
<td>2.88E-03</td>
<td>2.38</td>
<td>1.10E-13</td>
<td>1.34E-12</td>
<td>2.63</td>
<td></td>
</tr>
<tr>
<td>CDC20</td>
<td>991</td>
<td>6.08E-05</td>
<td>1.29E-02</td>
<td>2.88</td>
<td>7.24E-15</td>
<td>1.16E-13</td>
<td>3.57</td>
<td></td>
</tr>
<tr>
<td>CDC25C</td>
<td>995</td>
<td>2.91E-06</td>
<td>3.05E-03</td>
<td>3.79</td>
<td>3.12E-15</td>
<td>5.49E-14</td>
<td>4.93</td>
<td></td>
</tr>
<tr>
<td>CDC45</td>
<td>8318</td>
<td>7.08E-04</td>
<td>4.68E-02</td>
<td>2.65</td>
<td>1.66E-10</td>
<td>1.06E-09</td>
<td>3.66</td>
<td></td>
</tr>
<tr>
<td>CDC46</td>
<td>83461</td>
<td>1.98E-04</td>
<td>2.32E-02</td>
<td>3.37</td>
<td>2.13E-14</td>
<td>3.06E-13</td>
<td>2.34</td>
<td></td>
</tr>
<tr>
<td>CDK9</td>
<td>1033</td>
<td>2.19E-05</td>
<td>6.91E-03</td>
<td>2.64</td>
<td>2.26E-11</td>
<td>1.70E-10</td>
<td>3.39</td>
<td></td>
</tr>
<tr>
<td>CELSR3</td>
<td>1951</td>
<td>5.26E-04</td>
<td>3.97E-02</td>
<td>2.67</td>
<td>2.46E-13</td>
<td>2.79E-12</td>
<td>2.74</td>
<td></td>
</tr>
<tr>
<td>COMP</td>
<td>1311</td>
<td>8.56E-05</td>
<td>1.41E-02</td>
<td>6.10</td>
<td>3.20E-11</td>
<td>2.34E-10</td>
<td>6.66</td>
<td></td>
</tr>
<tr>
<td>CPLX1</td>
<td>10815</td>
<td>4.23E-04</td>
<td>3.44E-02</td>
<td>3.52</td>
<td>8.90E-22</td>
<td>8.23E-20</td>
<td>2.54</td>
<td></td>
</tr>
<tr>
<td>CPNE7</td>
<td>27132</td>
<td>1.63E-05</td>
<td>5.80E-03</td>
<td>5.27</td>
<td>2.14E-09</td>
<td>1.11E-08</td>
<td>4.29</td>
<td></td>
</tr>
<tr>
<td>DLGAP5</td>
<td>9787</td>
<td>7.49E-04</td>
<td>4.82E-02</td>
<td>6.64</td>
<td>1.19E-12</td>
<td>1.15E-11</td>
<td>5.15</td>
<td></td>
</tr>
<tr>
<td>EPHA8</td>
<td>2046</td>
<td>1.48E-05</td>
<td>5.67E-03</td>
<td>4.28</td>
<td>4.29E-16</td>
<td>9.32E-15</td>
<td>4.48</td>
<td></td>
</tr>
<tr>
<td>ERG</td>
<td>2078</td>
<td>2.47E-04</td>
<td>2.62E-02</td>
<td>3.29</td>
<td>6.60E-11</td>
<td>4.53E-10</td>
<td>2.90</td>
<td></td>
</tr>
<tr>
<td>ESPL1</td>
<td>9700</td>
<td>2.20E-06</td>
<td>2.88E-03</td>
<td>3.47</td>
<td>3.99E-11</td>
<td>2.87E-10</td>
<td>2.43</td>
<td></td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>NCBI Gene ID</td>
<td>P-value (VPC)</td>
<td>Adjusted P-value (VPC)</td>
<td>Fold change (VPC)</td>
<td>P-value (TCGA)</td>
<td>Adjusted value (TCGA)</td>
<td>P-value (TCGA)</td>
<td>Fold change (TCGA)</td>
</tr>
<tr>
<td>FAM57B</td>
<td>83723</td>
<td>3.84E-05</td>
<td>9.36E-03</td>
<td>7.06</td>
<td>5.44E-10</td>
<td>3.17E-09</td>
<td>2.41</td>
<td></td>
</tr>
<tr>
<td>FAM64A</td>
<td>54478</td>
<td>6.11E-04</td>
<td>4.33E-02</td>
<td>3.65</td>
<td>6.45E-13</td>
<td>6.65E-12</td>
<td>3.99</td>
<td></td>
</tr>
<tr>
<td>GTSE1</td>
<td>51512</td>
<td>5.97E-04</td>
<td>4.25E-02</td>
<td>3.64</td>
<td>5.34E-14</td>
<td>7.01E-13</td>
<td>4.28</td>
<td></td>
</tr>
<tr>
<td>HJURP</td>
<td>55355</td>
<td>2.53E-06</td>
<td>2.88E-03</td>
<td>4.09</td>
<td>8.10E-16</td>
<td>1.65E-14</td>
<td>6.22</td>
<td></td>
</tr>
<tr>
<td>IQGAP3</td>
<td>128239</td>
<td>7.06E-04</td>
<td>4.68E-02</td>
<td>3.48</td>
<td>3.54E-15</td>
<td>6.15E-14</td>
<td>4.20</td>
<td></td>
</tr>
<tr>
<td>KIF4A</td>
<td>24137</td>
<td>9.83E-06</td>
<td>5.15E-03</td>
<td>4.86</td>
<td>2.23E-14</td>
<td>3.20E-13</td>
<td>5.69</td>
<td></td>
</tr>
<tr>
<td>KIFC1</td>
<td>3833</td>
<td>1.59E-04</td>
<td>2.02E-02</td>
<td>2.80</td>
<td>9.58E-13</td>
<td>9.50E-12</td>
<td>2.77</td>
<td></td>
</tr>
<tr>
<td>KLLH35</td>
<td>283212</td>
<td>1.57E-04</td>
<td>2.01E-02</td>
<td>3.84</td>
<td>2.41E-19</td>
<td>1.20E-17</td>
<td>2.92</td>
<td></td>
</tr>
<tr>
<td>LRRCC36</td>
<td>55282</td>
<td>9.70E-06</td>
<td>5.15E-03</td>
<td>2.25</td>
<td>6.28E-09</td>
<td>3.01E-08</td>
<td>2.40</td>
<td></td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene ID</td>
<td>Gene Symbol</td>
<td>Gene ID</td>
<td>Gene Symbol</td>
<td>Gene ID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGAT5B</td>
<td>146664</td>
<td>MLF11P</td>
<td>79682</td>
<td>MYBL2</td>
<td>4605</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEIL3</td>
<td>55247</td>
<td>NCK2</td>
<td>4751</td>
<td>NETO2</td>
<td>81831</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NKX6-1</td>
<td>4825</td>
<td>NKX2-2</td>
<td>4821</td>
<td>NRE2E1</td>
<td>7101</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OIP5</td>
<td>11339</td>
<td>ONECUT2</td>
<td>9480</td>
<td>PBK</td>
<td>55872</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPP1R14B</td>
<td>26472</td>
<td>PLK1</td>
<td>5347</td>
<td>PRCC</td>
<td>9055</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>NCBI Gene ID</td>
<td>P-value (VPC)</td>
<td>Adjusted P-value (VPC)</td>
<td>Fold change (VPC)</td>
<td>P-value (TCGA)</td>
<td>Adjusted P-value (TCGA)</td>
<td>Fold change (TCGA)</td>
<td></td>
</tr>
<tr>
<td>PRC1</td>
<td>9055</td>
<td>3.02E-06</td>
<td>3.05E-03</td>
<td>2.51</td>
<td>6.91E-14</td>
<td>8.86E-13</td>
<td>2.01</td>
<td></td>
</tr>
<tr>
<td>PTPRT</td>
<td>11122</td>
<td>3.75E-04</td>
<td>3.26E-02</td>
<td>6.51</td>
<td>2.43E-06</td>
<td>7.67E-06</td>
<td>3.14</td>
<td></td>
</tr>
<tr>
<td>RAB19</td>
<td>401409</td>
<td>4.54E-04</td>
<td>3.61E-02</td>
<td>4.02</td>
<td>9.23E-09</td>
<td>4.29E-08</td>
<td>2.12</td>
<td></td>
</tr>
<tr>
<td>RDM1</td>
<td>201299</td>
<td>2.62E-05</td>
<td>7.58E-03</td>
<td>6.57</td>
<td>4.14E-10</td>
<td>2.47E-09</td>
<td>2.35</td>
<td></td>
</tr>
<tr>
<td>RNFT2</td>
<td>84900</td>
<td>3.55E-05</td>
<td>8.97E-03</td>
<td>2.16</td>
<td>2.74E-10</td>
<td>1.69E-09</td>
<td>2.32</td>
<td></td>
</tr>
<tr>
<td>RRM2</td>
<td>6241</td>
<td>3.38E-04</td>
<td>3.14E-02</td>
<td>4.06</td>
<td>1.56E-12</td>
<td>1.48E-11</td>
<td>4.42</td>
<td></td>
</tr>
<tr>
<td>SHCBP1</td>
<td>79801</td>
<td>2.19E-04</td>
<td>2.48E-02</td>
<td>3.38</td>
<td>2.30E-10</td>
<td>1.43E-09</td>
<td>2.81</td>
<td></td>
</tr>
<tr>
<td>SPAG5</td>
<td>10615</td>
<td>2.59E-04</td>
<td>2.66E-02</td>
<td>2.02</td>
<td>1.96E-14</td>
<td>2.85E-13</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>SPC24</td>
<td>147841</td>
<td>3.43E-04</td>
<td>3.14E-02</td>
<td>2.31</td>
<td>2.42E-13</td>
<td>2.75E-12</td>
<td>3.05</td>
<td></td>
</tr>
<tr>
<td>SRCIN1</td>
<td>80725</td>
<td>3.88E-04</td>
<td>3.32E-02</td>
<td>2.26</td>
<td>1.47E-08</td>
<td>6.62E-08</td>
<td>2.06</td>
<td></td>
</tr>
<tr>
<td>TK1</td>
<td>7083</td>
<td>2.28E-06</td>
<td>2.88E-03</td>
<td>2.45</td>
<td>1.76E-12</td>
<td>1.65E-11</td>
<td>2.24</td>
<td></td>
</tr>
<tr>
<td>TOP2A</td>
<td>7153</td>
<td>5.07E-05</td>
<td>1.13E-02</td>
<td>2.78</td>
<td>3.34E-10</td>
<td>2.03E-09</td>
<td>3.05</td>
<td></td>
</tr>
<tr>
<td>TPX2</td>
<td>22974</td>
<td>6.81E-05</td>
<td>1.33E-02</td>
<td>3.27</td>
<td>1.79E-13</td>
<td>2.09E-12</td>
<td>3.49</td>
<td></td>
</tr>
<tr>
<td>TRIP13</td>
<td>9319</td>
<td>1.45E-04</td>
<td>1.93E-02</td>
<td>2.84</td>
<td>3.81E-12</td>
<td>3.33E-11</td>
<td>2.65</td>
<td></td>
</tr>
<tr>
<td>TROAP</td>
<td>10024</td>
<td>7.58E-05</td>
<td>1.36E-02</td>
<td>4.77</td>
<td>5.09E-16</td>
<td>1.08E-14</td>
<td>4.45</td>
<td></td>
</tr>
<tr>
<td>TTTL6</td>
<td>284076</td>
<td>1.08E-04</td>
<td>1.61E-02</td>
<td>2.58</td>
<td>4.24E-13</td>
<td>4.59E-12</td>
<td>2.23</td>
<td></td>
</tr>
<tr>
<td>UBE2C</td>
<td>11065</td>
<td>6.92E-04</td>
<td>4.60E-02</td>
<td>5.45</td>
<td>4.86E-13</td>
<td>5.16E-12</td>
<td>4.86</td>
<td></td>
</tr>
<tr>
<td>UGT2B4</td>
<td>7363</td>
<td>4.24E-04</td>
<td>3.44E-02</td>
<td>7.18</td>
<td>2.32E-14</td>
<td>3.30E-13</td>
<td>8.55</td>
<td></td>
</tr>
<tr>
<td>WDR62</td>
<td>284403</td>
<td>1.52E-04</td>
<td>1.98E-02</td>
<td>2.71</td>
<td>6.52E-09</td>
<td>3.13E-08</td>
<td>2.17</td>
<td></td>
</tr>
<tr>
<td>ZIC2</td>
<td>7546</td>
<td>6.39E-07</td>
<td>1.11E-03</td>
<td>25.22</td>
<td>1.21E-22</td>
<td>1.32E-20</td>
<td>24.17</td>
<td></td>
</tr>
<tr>
<td>ZIC5</td>
<td>85416</td>
<td>2.19E-04</td>
<td>2.48E-02</td>
<td>7.87</td>
<td>8.79E-26</td>
<td>2.26E-23</td>
<td>10.42</td>
<td></td>
</tr>
<tr>
<td>ZP3</td>
<td>7784</td>
<td>1.42E-04</td>
<td>1.92E-02</td>
<td>2.18</td>
<td>2.74E-12</td>
<td>2.47E-11</td>
<td>2.33</td>
<td></td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>NCBI Gene ID</td>
<td>$P$-value (MSKCC)</td>
<td>Adjusted $P$-value (MSKCC)</td>
<td>Fold change (MSKCC)</td>
<td>$P$-value (TCGA)</td>
<td>Adjusted $P$-value (TCGA)</td>
<td>Fold change (TCGA)</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>------------------</td>
<td>---------------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>ABCC4</td>
<td>10257</td>
<td>8.29E-12</td>
<td>6.78E-10</td>
<td>2.06</td>
<td>3.56E-11</td>
<td>2.58E-10</td>
<td>3.62</td>
<td></td>
</tr>
<tr>
<td>ACSM1</td>
<td>116285</td>
<td>6.75E-20</td>
<td>9.77E-17</td>
<td>2.72</td>
<td>3.91E-14</td>
<td>5.30E-13</td>
<td>7.13</td>
<td></td>
</tr>
<tr>
<td>AMACR</td>
<td>23600</td>
<td>1.57E-22</td>
<td>7.58E-19</td>
<td>4.03</td>
<td>3.76E-27</td>
<td>1.31E-24</td>
<td>8.17</td>
<td></td>
</tr>
<tr>
<td>ARHGEF38</td>
<td>54848</td>
<td>6.68E-12</td>
<td>5.70E-10</td>
<td>2.39</td>
<td>5.11E-11</td>
<td>3.59E-10</td>
<td>4.10</td>
<td></td>
</tr>
<tr>
<td>CRISP3</td>
<td>10321</td>
<td>2.36E-11</td>
<td>1.75E-09</td>
<td>2.30</td>
<td>2.49E-02</td>
<td>4.05E-02</td>
<td>2.72</td>
<td></td>
</tr>
<tr>
<td>DNAH5</td>
<td>1767</td>
<td>4.22E-14</td>
<td>7.77E-12</td>
<td>2.48</td>
<td>6.80E-15</td>
<td>1.10E-13</td>
<td>5.66</td>
<td></td>
</tr>
<tr>
<td>EPCAM</td>
<td>4072</td>
<td>6.18E-12</td>
<td>5.31E-10</td>
<td>2.02</td>
<td>3.32E-14</td>
<td>4.57E-13</td>
<td>2.27</td>
<td></td>
</tr>
<tr>
<td>ERG</td>
<td>2078</td>
<td>7.77E-18</td>
<td>3.75E-15</td>
<td>2.66</td>
<td>6.60E-11</td>
<td>4.53E-10</td>
<td>2.90</td>
<td></td>
</tr>
<tr>
<td>GCNT1</td>
<td>2650</td>
<td>6.01E-13</td>
<td>7.43E-11</td>
<td>2.02</td>
<td>3.31E-14</td>
<td>4.57E-13</td>
<td>3.18</td>
<td></td>
</tr>
<tr>
<td>GLYATL1</td>
<td>92292</td>
<td>9.86E-13</td>
<td>1.13E-10</td>
<td>2.29</td>
<td>4.00E-09</td>
<td>1.99E-08</td>
<td>5.88</td>
<td></td>
</tr>
<tr>
<td>OR51E1</td>
<td>143503</td>
<td>6.83E-06</td>
<td>8.76E-05</td>
<td>2.01</td>
<td>5.53E-04</td>
<td>1.20E-03</td>
<td>3.01</td>
<td></td>
</tr>
<tr>
<td>PCA3</td>
<td>50652</td>
<td>2.11E-16</td>
<td>7.01E-14</td>
<td>4.84</td>
<td>6.46E-15</td>
<td>1.05E-13</td>
<td>28.21</td>
<td></td>
</tr>
<tr>
<td>SIM2</td>
<td>6493</td>
<td>1.11E-10</td>
<td>6.44E-09</td>
<td>3.02</td>
<td>5.32E-16</td>
<td>1.12E-14</td>
<td>7.94</td>
<td></td>
</tr>
<tr>
<td>TARP</td>
<td>445347</td>
<td>9.56E-06</td>
<td>1.16E-04</td>
<td>2.23</td>
<td>4.92E-10</td>
<td>2.89E-09</td>
<td>7.06</td>
<td></td>
</tr>
<tr>
<td>TDRD1</td>
<td>56165</td>
<td>8.69E-30</td>
<td>3.77E-25</td>
<td>3.05</td>
<td>1.72E-18</td>
<td>6.92E-17</td>
<td>17.90</td>
<td></td>
</tr>
<tr>
<td>TMEFF2</td>
<td>23671</td>
<td>4.74E-05</td>
<td>4.45E-04</td>
<td>2.06</td>
<td>6.16E-04</td>
<td>1.33E-03</td>
<td>2.24</td>
<td></td>
</tr>
<tr>
<td>TRIB1</td>
<td>10221</td>
<td>3.35E-11</td>
<td>2.35E-09</td>
<td>2.31</td>
<td>2.99E-12</td>
<td>2.67E-11</td>
<td>2.67</td>
<td></td>
</tr>
</tbody>
</table>
Table 2-4  A list of active VPC-18005 derivatives.

<table>
<thead>
<tr>
<th>VPC-ID</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>IC_{50} in Luciferase assay PNT1B-ERG cells</th>
<th>Solubility in media (n.d. = not determined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18005</td>
<td>CH₃</td>
<td>H</td>
<td>H</td>
<td>3 µM</td>
<td>&gt; 50 µM</td>
</tr>
<tr>
<td>18065</td>
<td>CH₃</td>
<td>H</td>
<td>H</td>
<td>2 µM</td>
<td>&gt; 50 µM</td>
</tr>
<tr>
<td>18098</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>1 µM</td>
<td>&gt; 50 µM</td>
</tr>
<tr>
<td>18104</td>
<td>CH₃</td>
<td>H</td>
<td>OH</td>
<td>3 µM</td>
<td>n.d.</td>
</tr>
<tr>
<td>18106</td>
<td>CH₃</td>
<td>H</td>
<td>H</td>
<td>4 µM</td>
<td>n.d.</td>
</tr>
<tr>
<td>18113</td>
<td>CH₃</td>
<td>Cl</td>
<td>H</td>
<td>8 µM</td>
<td>n.d.</td>
</tr>
<tr>
<td>18114</td>
<td>CH₃</td>
<td>H</td>
<td>F</td>
<td>8 µM</td>
<td>n.d.</td>
</tr>
<tr>
<td>18118</td>
<td>CH₃</td>
<td>H</td>
<td>CH₃</td>
<td>2 µM</td>
<td>n.d.</td>
</tr>
<tr>
<td>18119</td>
<td>CH₃</td>
<td>CH₃</td>
<td>H</td>
<td>4 µM</td>
<td>n.d.</td>
</tr>
<tr>
<td>18120</td>
<td>CH₃</td>
<td>H</td>
<td>Cl</td>
<td>1 µM</td>
<td>n.d.</td>
</tr>
</tbody>
</table>
Chapter 3: Utilization of the small molecule screening pipeline to identify inhibitors of ERG in the Prestwick chemical library

3.1 Introduction

The drug discovery pipeline established in Chapter 2 of this thesis provides the basis to screen additional chemical libraries. A preliminary study has been performed to screen the Prestwick Chemical Library which consists of 1280 small molecules that includes FDA-approved drugs. It is common in drug discovery to screen libraries of known drug molecules with established safety and bioavailability profiles to expedite lead development through Selective Optimization of Side Activities (SOSA), as previously demonstrated in other drug discovery programs for human diseases including cancers.

3.2 Results and Discussion

The ERG-responsive luciferase assay was used to screen the 1280 small molecules from the Prestwick Chemical Library. From this screen we identified parthenolide (PTL) as the most potent inhibitor of transcriptional activity in two ERG expressing PCa cell lines (VCaP and PNT1B-ERG) (Figure 3-1A&B). Although initial screening identified several other molecules that had increased inhibition compared to PTL, further assessment of IC\textsubscript{50} determined that these compounds did not give a dose response and were toxic. PTL inhibited the ERG-driven transcription activity in VCaP cells at IC\textsubscript{50} of 0.9µM and in PNT1B-ERG at IC\textsubscript{50} of 11µM.

Protein NMR, as the next step in our pipeline, was used to determine direct binding of a compound of interest to the DNA-binding ETS domain of ERG. The N-HSQC spectrum of N-labelled protein (100 µM) was assessed in the presence of increasing concentrations of DMSO-solubilized PTL, as well as with a DMSO control. A dose dependent chemical shift was observed in PTL treated samples compared to the DMSO in the spectra in a number of amide $^1$H\textsubscript{N}–$^{15}$N groups.
A chemical shift perturbation plot (Figure 3-2A) with PTL at 1:10 molar ratio (i.e. 1 mM) showed that the perturbed residues with shifts greater than the mean (0.026 ppm) were mostly located along helix α1, loop β1-β2, helix α3 and strand β3. These perturbed residues cluster around the two binding sites (pockets) of ERG as predicted by molecular docking, supportive of the mode of PTL binding to the ERG-ETS protein domain (Figure 3-2B). At this binding site, PTL can form hydrogen bonds with Leu313 and Lys357, and hydrophobic interactions with nearby residues including Gln312, Ala368, Tyr371, and Tyr372. Molecular docking also revealed a weaker binding site as illustrated in Figure 3.2. Both sites are located at the ERG-DNA interface where PTL binding is predicted to cause steric clashes with the DNA-backbone and interfere with the DNA interaction. PTL can form hydrogen bonds with Leu313 and Lys357, and hydrophobic interactions with nearby residues including Gln312, Trp351, Ala368, Tyr371, and Tyr372.

Electrophoretic mobility shift assays (EMSA) were performed using purified ERG-ETS domain and a DNA oligonucleotide containing the consensus GGAA recognition motif. Previously experiments have demonstrated that the recombinant ERG-ETS domain binds the cognate DNA with a KD~ 1 nM, not affected by the DMSO control. As shown in Figure 3-3, PTL disrupted the ERG/DNA complex dose-dependently.

Taken together, this preliminary experimental data is consistent PTL binding to the ERG-ETS protein domain at the interface required for DNA interaction and inhibiting ERG transcriptional activity. PTL is a sesquiterpene lactone that is derived from shoots of the feverfew (Tanacetum parthenium) plant and has been used as an herbal medicine to treat migraine and rheumatoid arthritis for centuries. Previous studies have reported anti-inflammatory and anticancer effects of PTL and its derivative dimethylamino-PTL (DMAPT), and showed that they can inhibit the NF-κB signaling pathway, induce apoptosis of cancer cells, and reduce tumor growth.
of lung, bladder and PCas. PTL has been shown to increase apoptosis in cancer cells through inhibition of pathways including NF-κB and PI3K. One of the mechanisms of action of PTL is through the inhibition of the P65 subunit of the NF-κB complex. Interestingly, TMPRSS2-ERG gene fusion isoforms differentially increase NF-κB mediated transcription through phosphorylation of NF-κB p65 on Ser536. Thus, combined with our study in which we showed that PTL can directly inhibit ERG, this suggests multiple modes of inhibition of PTL on the NF-κB pathway. Furthermore, it was recently shown that PTL can sensitize PCa cells to radiotherapy, taken together with the knowledge that ERG inhibition can reduce resistance to taxane therapies, PTL may synergize with current treatments to benefit patients with advanced PCa. Here we report a potential new mode of action by PTL in PCa and demonstrate that PTL can directly bind to the ERG-ETS domain, disrupt ERG-DNA interaction, inhibit ERG-driven transcriptional activity, and inhibit cell proliferation and invasion of ERG-expressing PCa cells.

As a known drug with established safety and bioavailability data, future medicinal chemistry effort on optimizing PTL can be expedited to further enhance its pharmacologic property and potency against ERG. Overall, the preliminary results of this study support an interesting avenue to investigate new anti-ERG mechanism by PTL in PCa that has not been previously reported. With an ever-increasing knowledge on molecular subtyping of PCa and association between ERG expression with aggressive PCa subtypes, there is an urgent demand for therapeutics targeting ERG-positive PCa. The screening strategy described in this thesis provide a thorough way to investigate potential candidates.
Figure 3-1 Discovery of parthenolide as an inhibitor of the ERG-ETS domain and its transcriptional activity. (A) Schematic representation of ERG-inhibitor screening from the Prestwick Chemical Library. Chemical structure of parthenolide is shown in its isomeric form as provided by the vendor, Cayman Chemical (catalog #70080). Molecular weight = 248 g/mol at pH 7.4. (B) Parthenolide inhibited ERG transcriptional activity in luciferase assays in PNT1B-ERG and VCaP cells. (C) Overlaid 15N-HSQC spectra of ERG-ETS domain (100 µM) in the absence (red) and presence of increasing protein: compound molar ratios of DMSO-solubilized parthenolide (yellow 1:1, green 1:2, cyan 1:4, purple 1:10, dark blue 1:20).
Figure 3-2 Binding mode of parthenolide to the ERG-ETS domain.

(A) Amide chemical shift perturbations resulting from the addition of a 10-fold molar excess of parthenolide to the ERG-ETS domain (derived from Figure 3.1C). Coloured bars denote significant changes (magenta ≥ mean + standard deviation, cyan ≥ mean). The secondary structure of the EGR-ETS domain is shown at the bottom. (B) Molecular docking of parthenolide to the ERG-ETS protein domain surface (PDB: 4IRG). Among the top 9 binding poses, 7 out of 9 docked to the same site. The top scoring pose is shown. Amino acid residues exhibiting significant chemical shift perturbations from panel A were mapped to their corresponding locations on the ERG-ETS domain. The DNA backbone (green) superimposed from the ERG-DNA complex (PDB: 4IRI) is shown to illustrate that the parthenolide binding can sterically block ERG-DNA interactions. (C) A ribbon-representation of the same binding model as in B. (D) Protein residues that are predicted to interact with parthenolide at this binding site. The red dotted lines indicate hydrogen bonds, and the green lines represent non-polar packing interactions.
Figure 3-3 Parthenolide disrupts binding of the ERG-ETS domain to DNA.
EMSA shows binding of 4 nM ERG-ETS domain to 1 nM fluorescently-labelled dsDNA alone and in the presence of increasing concentrations of DMSO (top panel, 0.008 - 17%) and parthenolide (middle panel, 4 μM – 8 mM).
Figure 3.4 Alternative binding mode of parthenolide to the ERG-ETS domain.

(A) Molecular docking of parthenolide to the ERG-ETS protein domain surface (PDB: 4IRG). Among the top 9 binding poses, 7 out of 9 docked to the main site (Figure 3-2B) and 2 out of 9 docked to this alternative site as shown. Amino acid residues exhibiting significant chemical shift perturbations from Figure 3-2 were mapped to their corresponding locations on the ERG-ETS domain. The DNA backbone (green) superimposed from the ERG-DNA complex (PDB: 4IRI) is shown to illustrate that this alternative parthenolide binding can also sterically block ERG-DNA interactions. (B) A ribbon-representation of the same binding model as in A. (C) Protein residues that are predicted to interact with parthenolide at this binding site. The red dotted lines indicate hydrogen bonds, and the green lines represent non-polar packing interactions.
Table 3-1  A summary of top 15 compounds in the Prestwick library that resulted in approximately 60 percent or lower activity after compound treatment in ERG$^+$ cells and at 10 and 1 µM concentrations.

<table>
<thead>
<tr>
<th>Sample (Prestwick-#)</th>
<th>Chemical Name</th>
<th>% Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average (10µM)</td>
</tr>
<tr>
<td>925</td>
<td>Thonzonium bromide</td>
<td>5.2</td>
</tr>
<tr>
<td>1362</td>
<td>Vorinostat</td>
<td>6.1</td>
</tr>
<tr>
<td>777</td>
<td>Alexidine dihydrochloride</td>
<td>10.4</td>
</tr>
<tr>
<td>487</td>
<td>Daunorubicin hydrochloride</td>
<td>16.7</td>
</tr>
<tr>
<td>550</td>
<td>Parthenolide</td>
<td>18.9</td>
</tr>
<tr>
<td>787</td>
<td>Merbromin</td>
<td>19.5</td>
</tr>
<tr>
<td>862</td>
<td>Raloxifene hydrochloride</td>
<td>21.7</td>
</tr>
<tr>
<td>126</td>
<td>Mefloquine hydrochloride</td>
<td>29.2</td>
</tr>
<tr>
<td>270</td>
<td>Fendiline hydrochloride</td>
<td>29.8</td>
</tr>
<tr>
<td>78</td>
<td>Thioridazine hydrochloride</td>
<td>37.2</td>
</tr>
<tr>
<td>1068</td>
<td>Thiethylperazine dimalate</td>
<td>38.1</td>
</tr>
<tr>
<td>708</td>
<td>Benzethonium chloride</td>
<td>38.5</td>
</tr>
<tr>
<td>149</td>
<td>Thioproperazine dimesylate</td>
<td>39.3</td>
</tr>
<tr>
<td>790</td>
<td>Cycloheximide</td>
<td>40.9</td>
</tr>
<tr>
<td>1014</td>
<td>Sertraline</td>
<td>41.2</td>
</tr>
</tbody>
</table>
3.3 Materials and Methods

3.3.1 Cell culture

VCaP (CRL-2876) human prostate carcinoma cells, harboring an endogenous TMPRSS2-ERG gene fusion, were obtained from the American Type Culture Collection (ATCC), and maintained in low bicarbonate DMEM (ATCC) supplemented with 10 % FBS VCaP cells. The immortalized prostatic epithelial cell line, PNT1B was purchased from ATCC. PNT1B- Mock and PNT1B-ERG are previously derived lineage-matched control and ERG-expressing prostatic epithelial lines maintained in DMEM (Life Technologies) supplemented with 10 % FBS and under blasticidin selection. Cells were grown in a humidified, 5 % CO2 incubator at 37°C.

3.3.2 Dual reporter luciferase assay

Compounds from the Prestwick library were tested in a luciferase-based ERG-responsive reporter assay as described previously (20). VCaP and PNT1B- ERG cells (3000) were plated in 150 μL per well of a 96 well plate. After seeding for 24 h, cells were transfected with 50 ng of an Endoglin E3 promoter-derived ETS-responsive firefly luciferase reporter (~507/~280 of (E3) promoter inserted into luciferase reporter vector (Signosis), ARR3tk-luc, and 5 ng of the Renilla luciferase reporter (pRL-tk, Promega) using TransIT 20/20 transfection reagent (Mirus, USA). After 16 h incubation, treatment media was added for an additional 48 h. Data were normalized first to Renilla luciferase and then to the DMSO-media control on each plate, unless otherwise stated. Initial hit compounds were identified as those with an average normalized luciferase reading (firefly luciferase/Renilla luciferase readings) that was 60% or less (Table 3-1) of the average normalized luciferase reading of the DMSO-media control (i.e. 40% or more reduction of luciferase activity) at 10 μM. The luciferase assays were repeated for each lead
compound under multiple concentrations (0.1 to 100 μM) to establish a dose-dependent response and an IC$_{50}$ value.

3.3.3 NMR spectroscopy: ERG-ETS domain expression and purification

A pET28a plasmid encoding residues 307-400 of the ERG-ETS domain was expressed in *E. coli BL21 (λDE3)*. Cultures of 1 L were grown at 37°C in M9 media supplied with 3 gm/L $^{13}$C$_6$-glucose and/or 1 gm/L $^{15}$NH$_4$Cl. Cells were allowed to grow to O.D.$\text{$_{600}$}$ = 0.6 and protein expression was induced by adding 1 mM IPTG. After an induction time of 4 h, cells were harvested by centrifugation and stored at -80 °C for at least 1 round of freeze/thaw. Cells were resuspended in 40 mL of lysis buffer for every 1 L of culture. Cells were lysed by passing through 5 rounds of homogenization and 10 min of sonication. The cell lysate centrifuged at 15k rpm for 1 hr, and the supernatant subjected to nickel column purification. The column was washed using 25 mM imidazole (50 mM phosphate, 1 M NaCl, pH 7.4) and proteins were eluted with 1 M imidazole. Fractions containing the ETS domain were confirmed by SDS-PAGE and pooled. The His$_6$-tag was cleaved by thrombin and the tag-free sample was concentrated to 2 mL and subjected to S75 size exclusion chromatography. Fractions were checked by SDS-PAGE and those containing the pure sample were pooled and concentrated. The protein ample was dialyzed to NMR buffer (20 mM sodium phosphate, 150 mM NaCl, 2 mM DTT, 0.1 mM EDTA, pH 6.5) for all NMR experiments. *NMR spectral assignments.* NMR data were recorded at 25 or 28 °C on cryoprobe-equipped 850 MHz Bruker Avance III spectrometer. Data were processed and analyzed using NMRpipe$^{221}$ and Sparky$^{222}$. Signals from backbone and sidechain $^1$H, $^{13}$C, and $^{15}$N nuclei were assigned by standard multidimensional heteronuclear correlation experiments.

*NMR-monitored titrations.* Interactions of compounds with the ERG-ETS domain were monitored via sensitivity-enhanced $^{15}$N-HSQC spectra. Experiments involved titrating unlabeled
DMSO-solubilized compound or control DMSO into $^{15}$N-labeled ERG-ETS domain. Chemical shift perturbations were calculated from the combined amide $^1$H$^N$ and $^{15}$N shift changes as $\Delta \delta = [(0.2 \times \Delta \delta_N)^2 + (\Delta \delta_H)^2]^{1/2}$. Reciprocal titrations were carried out using $^1$H-NMR to monitor the effects of progressively adding unlabeled protein to a sample of PTL (180 μM) in 20 mM phosphate, 150 mM NaCl, 2 mM DTT, 0.1 mM EDTA, pH 6.5. The signal from water was suppressed by pre-saturation.

3.3.4 Electrophoretic mobility shift assay (EMSA)

Purified ERG-ETS domain (see NMR spectroscopy) was stored in buffer (20 mM sodium phosphate, 200 mM NaCl, 2 mM DTT, 0.1 mM EDTA, pH 6.5). To prepare the probe for the gel shift assay, equal amounts (200 nM) of Alexa-488 fluorophore-labeled DNAs (5’ CGGCCAAGCCGGAAGTGAGTG-3’ and its complement) were mixed, heated to 95 °C for 30 minutes, and then slowly cooled to 25 °C in several hours. An initial gel shift assay was performed by titrating constant 1 nM labeled dsDNA with ERG at concentrations spanning 0.3 pM to 0.5 μM. Glycerol (3 %) and 0.2 mg/mL bovine serum albumin (BSA) were included in the reaction mixture. After incubated at room temperature for 1 hr, samples were load on to 10 % polyacrylamide native gel, and electrophoresed at 10 °C. The gel was scanned with Typhoon 9200 Imager equipped with blue laser to excite at 490 nm and fluorescence was measured at 520 nm. The scanned image was analyzed with Image J. Non- linear least squares fitting (GraphPad Prism) of the titration data to a 1:1 binding isotherm yielded the equilibrium dissociation constant ($K_D$ value ~ 1 nM) for the ERG-ETS domain interaction with DNA. The binding isotherm equation is $f_{b,i} = [\text{ERG}]_i/([\text{ERG}]_i + K_D)$ where $[\text{ERG}]_i$ is the total concentration of the ERG-ETS domain (a valid approximation as $K_D > 1$ nM total dsDNA) at each titration point (i), and the fraction bound, $f_{b,i}$ was calculated as the intensity of the bound DNA band at that point relative to the intensity with saturating 0.5 μM
protein. The result of this initial study was used to set the molar ratio of ERG-ETS domain: DNA in subsequent competition assays with PTL. For these assays, 4 nM of the ERG-ETS domain was mixed with 1 nM of fluorophore-labeled dsDNA, titrated with PTL (diluted from a DMSO stock) and analyzed by the same EMSA protocol. The data were fit to the equation for competitive binding, \( f_{b,i} = \frac{[\text{ERG}]}{[\text{ERG}] + K_D \left(1 + \frac{[\text{PTL}]}{K_I}\right)} \), where \( K_I \) is the inhibitor dissociation constant and the fraction bound, \( f_{b,i} \), was calculated the intensity of the bound DNA band at each titration point relative to that without added PTL. A control experiment was carried out by titrating with equivalent quantities of DMSO.

3.3.5 *In silico* modeling

Molecular docking. The published X-ray crystal structure of ERG-ETS domain (PDB: 4IRG)\(^{204}\) was prepared by Protein Preparation Wizard module of the Maestro program from the Schrodinger software suite (Maestro Version 10.7.014, Release 2016-3). The docking grid, centered at the ERG-ETS protein structure, was enlarged to include the entire domain. This grid configuration enabled ‘blind docking’ where no specific pocket residues were specified, and the docking program would search for the best sites for small molecule binding over the entire protein surface. The 2D chemical structure of PTL was extracted from the vendor Cayman Chemical (Catalog number: 70080), and its 3D structure was protonated at pH 7 and energy-minimized under the MMFF94x (solvation: Born) force field using the Molecular Operating Environment (MOE)\(^{218}\). The PTL chemical structure was docked against the prepared ERG-ETS docking grid by using the Glide program (Standard Precision, Schrodinger software suite, Release 2016-3). Top 9 docking poses were generated. The crystal structure of an ERG/DNA complex (PDB: 4IRI)\(^{204}\) was used to define the ERG-DNA interface. MOE and Marvin Sketch were used to visualize and represent the protein and chemical structures.
Chapter 4: Discussion

Through the development of in vitro screening assays and biochemical characterization, this thesis was able to identify the chemical compound, VPC-18005, as an inhibitor of ERG action and the first small molecule to demonstrate direct binding to ERG. VPC-18005 joins a small group of known ETS small molecule inhibitors, including YK-4-279 (FLI1, ETV1, and ERG)\textsuperscript{199,201}, BRD32048 (ETV1)\textsuperscript{237}, and ERGi-USU2 (ERG)\textsuperscript{238}. (Note: ERGi-USU2 was published after the initial preparation of this thesis and inhibits ERG protein expression through its direct binding to RIOK2 which induces ribosomal stress\textsuperscript{238}. ETS factors play a role in the control of many processes that are important in tumorigenesis, such as cell cycle, apoptosis, cell migration, and angiogenesis. Aberrant expression of ETS factors is associated with several cancers, including prostate cancer, Ewing sarcoma, and leukemia, thus, small molecule inhibitors targeting ETS factors have the potential to be effective cancer treatments. Equally, it also makes sense that targeting members of the ETS family could result in toxicity in tissues that normally express ETS transcription factors, such as those tissues involved in hematopoiesis. Preliminary in vivo data presented in this thesis does not show a significant impact of VPC-18005 on hematopoiesis. In support of this finding, YK-4-279 treatment in EWS-FLI1+ leukemic mice resulted in a correction of abnormal hematopoiesis and improved overall survival without any apparent side effects\textsuperscript{239}.

Due to advancements in computer-aided drug discovery approaches, the notation of ‘druggability’, previously defined by conventional targets such as enzymes and GPCRs, has been slowly expanded to include surface sites at protein-DNA and particularly protein-protein interfaces\textsuperscript{240,241}. While there are continual advances in modeling it is likely that it will never be able to simulate a complete biological system. For example, in the model used in this thesis, the target
protein is in a fixed position with limited flexibility. In their native state target molecules are highly flexible, thus, identifying interacting molecules using a rigid structure may lead to higher proportion of inaccurate hits. Furthermore, not all proteins have structural information. Indeed, at the commencement of this study the ERG-ETS domain X-ray crystal structure was not available. Homology modeling and protein threading methods are capable of building models based on sequence conversation across protein families. However, these are based on a lot of assumptions. In this study, a homology model was initially created for the ERG-ETS domain using the structural template from another ETS transcription factor, FEV, a close homolog of ERG with 85% sequence identity at the ETS domain. Fortuitously, the ERG structure was published as this model was being created. While our predicted structure closely resembled the ERG-ETS published model, the study moved forward with the published ERG-ETS model.

Sequence conservation at the ETS domain means that there is similarity across the ETS transcriptional family. Thus, it is expected that anti-ERG compounds such as VPC-18005 will have the potential to bind and inhibit other ETS factors. As discussed in Chapter 2, preliminary NMR spectroscopic experiments revealed that VPC-18005 also interacts with the ETS domains of PU.1 and ETV4. Since the initial preparation of this thesis, additional work (Cox unpublished) has been done whereby HEK cells were transfected with individual ETS factors (ETV5, ETV4, ETV1 and Fli1) and treated with a VPC-18005 derivatives. Cross reactivity was observed but this represents a tool by which we can screen derivatives for ETS specificity more efficiently. Ultimately, this information can help to enable further development of small molecules targeting ERG and other ETS factors. Hsing et al. demonstrates that while there is sequence conservation across the ETS domains, the 3D structures of the DNA-binding pockets are different with unique polar and
hydrophobic regions \(^{134}\). Additionally, as that there are known X-ray crystal structures for select ETS factors, homology modelling and protein threading methods are capable of building models for the other ETS factors based on sequence conversation across protein families. This information combined with additional in silico methods can be used to introduce new substituents systematically and replace selected moieties in the small molecule with those that have similar/different shape, chemistry and electrostatics to enable the identification of molecules with characteristics that have increased selectivity toward individual ETS factors. Conversely, interacting with more than one ETS factor may not be an entirely negative finding. YK-4-279 has been shown to interact with both EWS-FLI1 and ERG and its clinical derivative is currently in a phase 1 trial of EWS patients \(^{202}\).

This study used luciferase assay to screen the small molecules identified by the in silico screen. The advantage of luciferase assays is that they are high throughput, sensitive, reproducible and have relatively low cost. However, the results of this thesis do highlight differences between VPC-18005’s EMSA Kd/Ki and luciferase IC50 values. Given a Kd between ERG and DNA in a nanomolar range is also reported by Regan et al. \(^{204}\), it is expected that first-generation anti-ERG small molecules, such as VPC-18005, would require high concentrations to compete with ERG-DNA binding in the EMSA experiments. As for the lower IC50 values of VPC-18005 in cellular assays, the differences could arise for numerous reasons, spanning from potentially elevated intracellular concentrations of VPC-18005 to highly sensitive effects of this compound on the network of cooperative intermolecular interactions required for transcription. The high sensitivity of cells towards perturbation by small molecule inhibitors due to the local arrangement of enhancer elements on DNA has been previously reported \(^{244}\). Alternatively, it can be acknowledged that there are limitations to the firefly luciferase assay used in this study. Firstly, ERG is constitutively
expressed and not inducible, so luciferase reporter expression will increase with time following transfection. Thus, ERG antagonizing agents will always need to be chronically active. Secondly, firefly luciferase-based luminescence readouts are known to be susceptible to small molecule inhibition of firefly luciferase\textsuperscript{245,246}. Since the initial preparation of this thesis, unpublished work from the laboratory tested the nanoLuc platform (Promega), which due to its greater structural rigidity is less likely to bind nonspecifically to small molecules\textsuperscript{247}. This method has successfully confirmed VPC-18005 inhibition with an IC50 more comparable to the EMSA result and identified more potent derivatives with complementary data that demonstrates suppression of SOX9 as well as other target genes (Cox unpublished).

\section*{4.1 Future Directions}

\subsection*{4.1.1 Characterize the impact of molecular probes on ERG-Mediated Metastasis}

Ultimately, to progress VPC-18005 or a more potent candidate lead molecular probe to the clinic additional \textit{in vivo} studies will need to be performed to demonstrate that these candidate molecules are capable of suppress ERG-mediated oncogenic transformation. This thesis used the zebrafish model as an initial attempt to test VPC-18005 in an \textit{in vivo} system. Zebrafish are increasingly being tested in preclinical models because they are suitable for large scale screening, are inexpensive, (soluble) small molecule treatment can be easily added to the water, and the process of metastasis can be followed in real-time\textsuperscript{248}. The initial assessment of mammalian tolerability using mice for pharmacokinetic/toxicology studies identified VPC-18005 to be soluble, stable, and orally bioavailable. These findings are a significant step forward toward the development of ERG-targeted therapeutics and provides a solid basis upon which to refine the molecule to increase potency. While the lack of overt murine toxicity for VPC-18005 is encouraging; there are challenges in the identification and development of an appropriate
spontaneous metastasis model driven by ERG overexpression. Future studies will apply the following two models to assess the effect on ERG driven metastasis.

4.1.1.1 Murine metastasis assays

Tail vein/cardiac injection assays will be used to measure capacity of ERG chemical probes to impact seeding of lung, liver, bone lesions. To establish the metastatic homing of ERG expressing cells, VCaP-Luc cells have already been engineered using lentiviral UBC-luciferase vector with blasticidin resistance that grow as a subcutaneous xenograft in nu/nu mice. VCaP-Luc cells inoculated via tail vein to monitor lung metastasis formation, and intracardiac inoculation will be used to monitor for growth of luciferase-expressing cells in distal sites using the Xenogen live imaging system. Mice will receive daily VPC-18005 PO treatment and be monitored daily for their weight and general health. At experimental end point, mice will be harvested, and liver, lung, heart, brain, bone, and kidneys will be examined individually for the presence/rate of metastatic VCaP-Luc cells by histopathology. This approach will determine whether VPC-18005 or lead chemical probes can directly impact metastasis.

4.1.1.2 Xenograft growth and metastasis

To test whether ERG chemical probes can truly suppress tumor growth and metastasis, PNT1B-ERG cells will be grafted under the renal capsule of intact SCID mice (12/group). Tumor take and growth rate will be monitored using ultrasound imaging. One week after inoculation, mice will be treated with 18005 (dose determined from the previous xenograft models) or vehicle control once daily PO for 4 weeks and the resulting primary tumor volume, degree of invasion and metastatic burden will be assessed based on histopathology of tumor-bearing and contralateral kidney, lung and liver. As an additional measure of tumor cell dissemination, RNA isolated from
blood collected at endpoint will be analyzed for TMPRSS2-ERG transcript levels normalized to human GAPDH \(^{249}\).

Preliminary results demonstrate that PNT1B-ERG cells, but not PNT1B-MOCK, are invasive \(^{106}\) and that VPC-18005 inhibits dissemination of PNT1B-ERG and VCaP cells in zebrafish embryos. Other ERG-PrEC lines may need to be included and ERG-null PCa cells will need to be tested as negative controls. Furthermore, the impact of ERG chemical probes on metastasis in these models could be compared to that of cells engineered to express ERG-directed shRNA as definitive ERG loss-of-function controls. Lastly, Tg(fli1a:EGFP)y1 transgenic zebrafish embryos may be utilized. Having eGFP expressed in their vasculature off an ETS promoter would allow us to study the effect of our ERG chemical probes on angiogenesis in the fish and localization of fluorescently labelled PCa cells. While the precise effects of ERG on PCa cell survival/proliferation remains to be elucidated \(^{102}\), several studies have shown ERG knockdown reduced tumor volume in PCa xenograft mouse models \(^{104,183,196}\). Thus, our lead ERG inhibitors have the potential to reduce not only metastasis, but also tumor growth, both of which will be monitored closely. In addition, previous studies have demonstrated that over expression of ERG in normal prostate epithelial cells induced transcriptional changes of genes involved in cell-cell adhesion and EMT \(^{106}\). Therefore, for both tumor models, ERG-driven phenotypic features such as EMT markers will be assessed by IHC/RNA as a measure of the reversal of ERG transcriptional activity and this is expected to correlate with xenograft growth and metastatic changes.
4.1.2 Functional Characterization of ERG chemical probes

4.1.2.1 Transcriptome analysis of VPC-18005-treated ERG PCa models

The evidence presented in this thesis that VPC-18005 can suppress ERG transcriptional activity without overt cytotoxicity in ERG-PNT1B and VCaP cells provides two models in which to assess how selective targeting of ERG impacts the transcriptome of these cells. Agilent whole genome expression microarrays will be used to compare how VPC-18005 and ERG siRNA impact the transcriptome of VCaP and ERG-PrECs. In addition, these treatments will be compared to that of VCaP cells treated by androgen deprivation to reveal how androgen pathway signaling is altered when ERG activity is muted in this system. Analysis will focus on potential for oncogenic phenotype reversion using differential expression of EMT markers previously identified \(^{10,106}\). Comparison with transcriptional changes observed in the ERG-PrECs, will identify potential markers of recurrent phenotypic changes in ERG-transformed PCa’s that may be amenable for prognostics or therapeutic intervention.

4.1.2.2 Validation of ERG transcriptional targets and identify novel ERG partners

4.1.2.2.1 ERG Protein Interactions

The regulation of downstream genes by ERG is dependent on its ability to form protein interactions. ERG has been shown to directly interact with AR at the ETS domain and this interaction is independent of DNA \(^{10}\), and to form a complex with Fos and Jun that regulates MMP1 and MMP3 gene expression \(^{250}\). Similarly, ERG also interacts with PARP1 and DNA-PKcs, which affects transcription of downstream genes that include PLA1, FKBP5 and TMPRSS2. These interactions between ERG and DNA-PKcs and PARP1 are independent of DNA binding. In the presence of DNA, ERG-DNA-PKcs forms a larger complex with Ku70 and Ku80 \(^{187}\). These interactions are with the C-terminus of ERG where the ETS domain is located and may be impacted
by VPC-18005-based chemical probes. Since previous studies investigating ERG binding partners have been performed in VCaP cells, this model will be used to test whether treatment with VPC-18005 or its derivatives disrupts formation of ERG protein-protein interactions with AR, Fos, Jun, PARP, DNA-PKcs, Ku70 and Ku80, by reciprocal co-immunoprecipitation (CoIP) with ERG. As protein interactions may be influenced by cell type, factors that associate with ERG will also be assessed in the ERG-PrEC models.

4.1.2.2.2 ERG-regulated gene expression

By targeting the ETS domain of ERG with novel chemical probes, the goal is to disrupt the binding of ERG to DNA, and to subsequently affect gene transcription. While the transcriptome analysis above will help identify global impact of the chemical probes on cellular transcriptome, there is also a need to assess direct binding of ERG to DNA to demonstrate the predicted targeting mechanism. Whether VPC-18005 or its derivatives disrupt ERG-DNA interactions will be assessed using EMSA and chromatin immunoprecipitation (ChIP) studies. The ETS family of transcription factors recognize the core consensus sequence 5’-GGA(A/T)-3’. ERG binds to the CXCR4 promoter via specific oligonucleotide sequences that contain the GGAA core116. These experiments will be replicated by testing whether incubating ERG with VPC-18005 or derivatives disrupts binding to specific CXCR4 promoter sequences. Subsequent studies will evaluate other previously reported direct ERG gene targets identified by previously published ChiP and ChIP-Seq data such as: PLAU, PLAT, MMP3, MMP9152, EZH210 and cMYC104. Specificity will be assessed by including genes that do not bind ERG, such as KIAA00610. As we have described, we have not ruled out cross reactivity with other ETS factors (i.e. ETV1, ETV4, Fli1). As medicinal chemistry improvements are used to fine tune structure of the chemical probes based on the subtlety of the ETS binding pockets, this EMSA system can be used to assess selectivity. More
selective chemical probes can highlight differences between the binding of the different factors. In addition, site directed oligonucleotide mutations in EMSA assays will also be used to assess impacts on binding.

It should be noted that the impact of ERG chemical probes on different ERG-expressing cell models may vary. While there is no evidence of cytotoxicity in the three lines tested to date, impact on others remains to be determined. The lack of selective toxicity to ERG-expressing lines may be due to the assays being performed in optimal culture conditions. Viability assays could be performed under stress conditions, such as hypoxia or nutrient deprivation, to determine whether the combination of environmental stress reveals ERG-related viability dependence in the VCaP and ERG-PrEC lines. Evidence of senescence induction can be further assessed by measurement of proinflammatory cytokines and other factors which are known as the senescence associated secretory phenotype (SASP). Determining if ERG antagonism induces stress responses will be important considerations for developing ERG-targeted therapeutics since SASP and autophagy can be tumor promoting states. Microarrays are limited by the fact dependence on bait probe design. As an alternative, deep sequencing (RNA-Seq) methods can cover the entire transcriptome which allows for more specific analysis of novel transcripts, splice junctions and noncoding RNAs. Thus, RNA-Seq using an optimized chemical probe may be considered to provide a more extensive analysis of impact on ERG-PrECs transcriptomes. As protein interactions may be influenced by cell type novel binding partners will be assessed in our ERG-PrECs using Silac\textsuperscript{251} as an unbiased protein correlation profiling means of identifying proteins that interact with ERG.

4.2 Conclusion

There is currently a need for novel therapeutics to target metastatic CRPC. Combination therapies which include agents that target both epithelial and mesenchymal phenotypes may
ultimately produce a more enhanced effect. Thus, targeting ERG represents an important step forward toward achieving this goal. The development of ERG inhibitors may offer a novel option for precision medicine as they could be specifically prescribed to 50% of PCa patients carrying the ERG mutation which can be detected noninvasively by urine tests. The chemical probe, VPC-18005, and future chemical probes that can be derived from this work provide a foundation for the clinical development of ERG targeted agents as well as the targeting of other ‘undruggable’ oncogenic ETS factors. It is our ultimate goal that ERG inhibitors will be advanced into clinical trials (Figure 4-1) to be used in combination with current therapies to benefit patients with the most advanced forms of PCa.

**Figure 4-1 The future of ERG inhibitors in the clinic**
ERG inhibitors can be prescribed as preventive measures (during active surveillance) to patients of high risk that have positive TMPRSS2-ERG fusion urine test or in patients that have progressive disease with ERG\(^+\) phenotype.
Bibliography


Zong, Y. et al. ETS family transcription factors collaborate with alternative signaling pathways to induce carcinoma from adult murine prostate cells. Proc Natl Acad Sci U S A 106, 12465-12470, doi:0905931106 [pii]


Oncternal Therapeutics, I. *TK216 in Patients With Relapsed or Refractory Ewing Sarcoma*, (2016).


Group, C. C. *Molecular Operating Environment (MOE)*.


Goddard TD, K. D. *SPARKY 3.*, 1999).


