AD_____

Award Number: W81XWH-11-1-0494

TITLE: The Snail-Induced Sulfonation Pathway in Breast Cancer Metastasis

PRINCIPAL INVESTIGATOR: Dr. Frank J. Rauscher III, Ph. D.

CONTRACTING ORGANIZATION: Wistar Institute of Anatomy & Biology
Philadelphia, PA 19104

REPORT DATE: September 2014

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

REPORT DATE	2. REPORT TYPE	3. DATES COVERED
September 2014	Final	1 Jul 2011 – 30 Jun 2014

4. TITLE AND SUBTITLE

The Snail-Induced Sulfonation Pathway in Breast Cancer Metastasis

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-11-1-0494

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
Dr. Frank J. Rauscher III

E-Mail: rauscher@wistar.org

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Wistar Institute of Anatomy & Biology
Philadelphia, PA 19104

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This support from DOD has provided funding for a 3-year project that has resulted in fundamental new insights into how the transcription factor Snail can control gene regulation processes that regulate cancer metastases. The new discoveries include: 1. Binding of Snail to PRKG1, 14-3-3 and RING1a/B proteins; 2. Derivation of a model for Snail-DNA interaction and its verification using mutagenesis; 3. Phosphorylation of Snail by the kinase PRKG1 and recruitment of PO4-Snail to target genes; 4. Demonstration that LIMD2 is in the EMT pathway and also contributes to cancer metastases by regulating EMT, likely thru Snail pathways. Many of these have been published or are in preparation for publication. The demonstration that the PAPSS1 and PAPSS2 genes are direct targets for Snail during EMT continues to be elusive, mostly for technical reasons. Overall, this support has resulted in the discovery of key novel principles in EMT mediated gene regulation and has provided new exploitable targets for metastasis prevention and therapeutics which will be followed up in further studies which are ongoing.

15. SUBJECT TERMS Key words or phrases identifying major concepts in the report.
Metastasis, snail, transcription factor, DNA binding, repression, PRKG1, LIMD2, EMT

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	50	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

INTRODUCTION:

Background:

Snail and Slug are master regulators of the epithelial-mesenchymal transition (EMT) differentiation pathway involved in neural crest development in embryogenesis, which is also critically involved in the invasiveness and lethality of cancer metastasis. An absolute pre-requisite for undergoing EMT is the down-regulation of E-cadherin, claudins and other epithelial cell adhesion molecules, which facilitate cell motility. The Snail subfamily (Snai1, Snai2, Snai3) are activated during early EMT and bind directly to target promoters containing the canonical E-box recognition sequence: CACGTG. The Snail protein contains four COOH-terminal tandem C2-H2 ZF motifs, possessing sequence-specific DNA binding activity and anchoring a protein complex to chromatin at specific target genes. One of the most commonly derived sequences recognized by the Snail transcription factor family is the CAGGTG E-box sequence, usually found in multiple copies in promoters regulated by Snail. Induction of EMT by Snail during tumor progression is virtually a universal property of late stage carcinoma progression and metastasis as the cells become motile, invasive and highly metastatic. Hence, targeting the functions of Snail superfamily members with small molecule therapeutics is highly relevant. In order to understand and target the Snail family of factors, we must define the complete set of target genes that are bound by Snail in the pro-metastatic cell undergoing EMT. Moreover, we must define the post-translational modifications that occur on Snail and the Proteins to which it binds during EMT. We hypothesized that the PAPSS1 and PAPSS2 are direct target genes for Snail. As described below, we have made the most progress on defining Snail binding proteins, modifications and effector pathways in EMT.

Hypothesis/Objective:

The original hypothesis was that the Snail protein induces enzymes which result in changes in cellular sulfonation and that this is a required part of EMT during metastasis in breast cancer.

Specific Aims:

The original Specific Aims were: 1) test the hypothesis that the PAPSS2 gene is a direct target of Snail and that it is required for maintaining the mesenchymal phenotype and invasive property of the breast cancer cells during EMT and 2) test the hypothesis that Snail-stimulated sulfonation of proteoglycans (PGs) are crucial for the induction of EMT and metastasis.

BODY:

Overview: In the course of the 3-year project, we considerably broadened our work to include other aspects of Snail function including defining DNA binding specificity, target gene interaction, post-translational modifications, and protein-protein interactions. This broad approach has been very successful and has led to many novel discoveries. This broadening of the project was approved by the DOD and was described in our annual reports, along with the experimental and technical difficulties in addressing the original hypothesis. Both the successes and the problems are detailed below.

A method for in silico identification of SNAIL/SLUG DNA binding potentials to the E-box sequence using molecular dynamics and evolutionary conserved amino acids.

Binding of Snail to DNA is a dynamic process allowing for spatial- and sequence-specificity of the E box. Many methods for determination of DNA-protein structures do not allow for identification of dynamics of the search process, but provide only a single snapshot of the most stable binding. In order to better understand the dynamics of DNA binding as a protein encounters its cognate site, we

have created a computer-based DNA scanning array macro that sequentially inserts a high affinity DNA consensus binding site at all possible locations in a predicted protein-DNA interface. We showed, using short molecular dynamic simulations at each location in the interface, that energy minimized states and decreased movement of evolutionary conserved amino acids can be readily observed and used to predict the consensus binding site. The macro was applied to SNAIL class C2H2 zinc finger family proteins. The analysis suggests that (1) SNAIL binds to the E-box in multiple states during its encounter with its cognate site; (2) several different amino acids contribute to the E-box binding in each state; (3) the linear array of zinc fingers contributes differentially to overall folding and base-pair recognition; and (4) each finger may be specialized for stability and sequence specificity. Moreover, the macromolecular movement observed using this dynamic approach may allow the NH2-terminal finger to bind without sequence specificity yet result in higher binding energy. This macro and overall approach could be applicable to many evolutionary conserved transcription factor families and should help to better elucidate the varied mechanisms used for DNA sequence-specific binding. Moreover, this analysis should allow us to better distinguish real, direct targets of Snail and Slug during EMT than previous resolution has allowed.

Prokop JW, Liu Y, Milsted A, Peng H, Rauscher FJ 3rd. A method for in silico identification of SNAIL/SLUG DNA binding potentials to the E-box sequence using molecular dynamics and evolutionary conserved amino acids. J Mol Model. 2013 Sep;19(9):3463-9. doi: 10.1007/s00894-013-1876-y. Epub 2013 May 25. PMCID: PMC3745821.

High resolution DNA recognition by the Snail zinc finger protein: Testing of a Molecular Dynamics based model defines the atomic level interactions required for high affinity binding, E-box specificity and reveals potential strategies for small molecule control of EMT transcriptional programs

To verify the molecular models for Snail-DNA binding we performed a saturation site-directed mutagenesis study. These experiments yielded the following conclusions: Multiple AA-Base contacts occur on both strands of DNA; All 4 ZF might not contribute equally to DNA affinity & specificity; High affinity docking of the model does not require large change in DNA structure: B-form DNA is maintained during Snail DNA binding. Moreover, canonical AA-base contacts seen in other ZF structures, but new modes of content are evident. AA-Base contacts were verified and revealed specifically thru mutagenesis. TheW193 in finger 2, and the S221 in finger 3 and the R247 in the finger 4 determine the sequence specificity for Snail DNA-binding, while R191 in finger 2 and N222 in finger 3 stabilize the complex via phosphate backbone contacts. Snail has high specificity for DNA recognition , which requires the second central G to be presented in the E-box, while the central Gs are not required for Twist bHLH DNA-binding; E-box methylation has less influence for Snail DNA-binding as compared to bHLH DNA-binding. These results provide the first comparative analysis of Twist and Snail binding the canonical E-box.

In preparation for submission.

Snail Recruits Ring1B to Mediate Transcriptional Repression and Cell Migration in Pancreatic Cancer Cells.

Transcriptional repressor Snail is a master regulator of epithelial-mesenchymal transition (EMT), yet the epigenetic mechanism governing Snail to induce EMT is not well understood. We discovered elevated levels of the ubiquitin E3 ligase Ring1B and Snail, along with elevated monoubiquitination of H2A at K119 (H2AK119Ub1), are highly correlated with poor survival. Mechanistic investigations identified Ring1B as a Snail-interacting protein and showed that the carboxyl zinc fingers of Snail recruit Ring1B and its paralog Ring1A to repress its target promoters. Simultaneous depletion of Ring1A and Ring1B in cancer cells decreased Snail binding to the target chromatin, abolished H2AK119Ub1 modification, and thereby compromised Snail-mediated transcriptional repression and cell migration. We found that Ring1B and the SNAG-associated chromatin modifier EZH2 formed

distinct protein complexes with Snail and that EZH2 was required for Snail-Ring1A/B recruitment to the target promoter. Collectively, our results unravel an epigenetic mechanism underlying transcriptional repression by Snail, suggest Ring1A/B as a candidate therapeutic target, and identify H2AK119Ub1 as a potential biomarker for diagnosis and prognosis.

Chen J, Xu H, Zou X, Wang J, Zhu Y, Chen H, Shen B, Deng X, Zhou A, Chin YE, Rauscher FJ 3rd, Peng C, Hou Z. Snail Recruits Ring1B to Mediate Transcriptional Repression and Cell Migration in Pancreatic Cancer Cells. Cancer Res. 2014 Aug 15;74(16):4353-63. doi: 10.1158/0008-5472.CAN-14-0181. Epub 2014 Jun 5. PMID: 24903147.

LIMD2 is a small LIM-only protein overexpressed in metastatic lesions which regulates cell motility and tumor progression by directly binding to and activating the integrin-linked kinase.

Proteins that communicate signals from the cytoskeleton to the nucleus are prime targets for effectors of metastasis as they often transduce signals regulating adhesion, motility, and invasiveness. LIM domain proteins shuttle between the cytoplasm and the nucleus and bind to partners in both compartments, often coupling changes in gene expression to extracellular cues. As part of our DOD funded work on Snail, we looked at potential effectors of EMT. LIMD2 is a mechanistically undefined LIM-only protein originally found to be overexpressed in metastatic lesions but absent in the matched primary tumor. LIMD2 levels in fresh and archival tumors positively correlate with cell motility, metastatic potential, and grade including bladder, melanoma, breast and thyroid tumors. LIMD2 directly contributes to these cellular phenotypes as shown by overexpression, knockdown and reconstitution experiments in cell culture models. The solution structure of LIMD2 determined using NMR revealed a classic LIM-domain structure that was highly related to LIM1 of PINCH1, a core component of the Integrin Linked Kinase-Parvin-Pinch (IPP) complex. Structural and biochemical analyses revealed that LIMD2 bound directly to the kinase domain of ILK near the active site and strongly activated ILK kinase activity. Cells that were null for ILK failed to respond to the induction of invasion by LIMD2. This strongly suggests that LIMD2 potentiates its biological effects through direct interactions with ILK, a signal transduction pathway strongly linked to cell motility and invasion. In summary, we discovered that LIMD2 is a new component of the signal transduction cascade that links integrin-mediated signaling to cell motility/metastatic behavior and may be a promising target for controlling tumor spread. Moreover, recent work suggests that the LIMD2-ILK axis is directly involved in the Snail EMT pathway.

Peng H, Talebzadeh-Farrooji M, Osborne MJ, Prokop JW, McDonald PC, Karar J, Hou Z, He M, Kebebew E, Orntoft T, Herlyn M, Caton AJ, Fredericks W, Malkowicz B, Paterno CS, Carolin AS, Speicher DW, Skordalakes E, Huang Q, Dedhar S, Borden KL, Rauscher FJ 3rd. LIMD2 is a small LIM-only protein overexpressed in metastatic lesions that regulates cell motility and tumor progression by directly binding to and activating the integrin-linked kinase. Cancer Res. 2014 Mar 1;74(5):1390-1403. PMID: 24590809.

A Novel Phospho-Switch in the Linker Region of the Snail Zinc Finger Protein which regulates 14-3-3 association, DNA binding and Epithelial-Mesenchymal Differentiation.

Regulation of transcription factor function in eukaryotes often occurs at the level of DNA binding site recognition, via dimerization, allosteric regulation of the DNA binding domain, or post-translational modification. The C2H2 family of zinc finger (ZF) proteins comprises the largest family of transcription factors (TFs) existing in the human proteome; their ZFs bind to DNA as monomers, however, little is known about how DNA binding is regulated. We discovered a highly conserved phospho-switch in the ZF domains of Snail and Slug, a subfamily of C2H2 ZFs that function in both normal development and tumor progression to metastasis by regulating the Epithelial-Mesenchymal Transition (EMT) program. Using highly specific phospho-Snail antibodies, we showed that the phosphorylation at Ser11 and Thr177 sites occurs robustly and the modified Snail proteins are bound to Snail target genes during EMT. Kinase PRKG1 was identified to modify Snail at T177, which is

located in the canonical HC linker region between fingers 1 and 2. Phosphorylation at T177 reduces Snail-DNA interaction while recruiting 14-3-3 proteins. Recognition of pT177-Snail is 14-3-3 isoform-specific and is modulated by the specific DNA target site for Snail. A molecular dynamics based model of the pT177-Snail /14-3-3 /DNA complex shows dramatic 14-3-3-mediated unfolding of finger 1, largely supporting the experimental observations. In conclusion, we discovered a novel phosphorylation in the DNA binding region of Snail, which plays a critical role in regulating Snail-induced EMT through dictating the states of Snail in protein recruitment and protein-DNA binding.

Submitted and under review: J. Mol. Biol.

An evolutionarily conserved DNA architecture determines target specificity of the Twist-family bHLH transcription factors

Basic helix-loop-helix (bHLH) transcription factors and the Snail family of zinc finger proteins recognize the canonical E-box (CANNTG) to regulate gene transcription; however, given the prevalence of E-boxes in a genome, it has been puzzling how individual bHLH and Snail proteins selectively recognize E-box sequences on their targets. Twist is a bHLH transcription factor that promotes Epithelial-Mesenchymal Transition (EMT) during development and tumor metastasis. High resolution mapping of Twist occupancy in human and Drosophila genomes reveals that Twist, but not other bHLH proteins, recognizes a unique double E-box motif with two E-boxes spaced preferentially by five nucleotides. Using molecular modeling and binding kinetic analyses, we found that the strict spatial configuration in the double E-box motif aligns two Twist-E47 dimers on the same face of DNA, thus providing a high affinity site for a highly stable intramolecular tetramer. Biochemical analyses showed that the WR domain of Twist dimerizes to mediate tetramer formation, which is functionally required for Twist-induced EMT. These results uncover a novel mechanism for a bHLH transcription factor to recognize a unique spatial configuration of E-boxes to achieve target specificity. Moreover, the identification of this double E-box allows us to easily distinguish a Snail target gene from a Twist target gene. Whether these sets of genes overlap in regulation during EMT is of great interest. Moreover, the WR-WR domain interaction uncovered here sets an example of target gene specificity of a bHLH protein being controlled allosterically by a domain outside of the bHLH region and represents a promising target for small molecule therapeutics directed at EMT transcription factors.

Under Revision: Genes & Dev.

As described above, our analyses of the biochemical properties of Snail has been extremely successful. However, our analysis of the PAPSS1 and 2 genes as Snail targets has been fraught with difficulties over the entire project period. This has been noted in each yearly progress report. Saying it simply: Nothing we tried worked. These 2 genes seem to be highly required for cell viability and any change to their status via knockdown or over expression leads to cell death. Specifically, we could not overcome the following problems:

However, we have had unexpected experimental approach difficulties which were completely unforeseen during design of the original project. Specifically, a cornerstone experimental procedure for testing our hypothesis that PAPSS2 controls the mesenchymal phenotype, is the ability to stably eradicate expression of PAPSS2 gene expression in various cell lines and in tumors. The state-of-the-art technique for doing this is miRNA mediated gene knockdown. However, this simply does not work in our hands for this particular gene. We have tried multiple independent experimental approaches to miRNA KO including transient transfection, (either naked RNA or synthetic oligos), transient vector (lentiviral-mediated) expression of miRNA , and finally selection of independent clonal lines of stable expressing miRNA vectors. The results show that contrary to any gene which we have ever worked with, none of these techniques are able to silence PAPSS2 (or PAPSS1). The cells either immediately shut off expression of the transgene miRNA vector or the cells die. Thus, either PAPSS2 is an extremely important gene for normal cell homeostasis, or there are very significant off-

target effects in our constructs. However, the latter explanation is very curious as we have used multiple independent targeting sequences for PAPSS2, which should control for these effects. In order to perform the comparative genomic and proteomic analyses of metastatic cell that do or do not have PAAPSS2 expression as originally proposed, we must be able to do this. This must await development of better technologies to tackle properly.

Summary: Despite the difficulties, we have made fundamental new discoveries regarding the role of Snail and other E-box binding transcription factors in EMT.

KEY RESEARCH ACCOMPLISHMENTS: 1 Jul 2011 – 30 Jun 2014

This project seeks to establish a key role for Snail in the metastatic progression of breast cancer. To this end, we have accomplished the following tasks over the entire research period:

1. Identification of the mode of DNA binding for Snail.
2. Identification of PRKG1 as the key kinase that modifies Snail.
3. Demonstration that RING1a/a, 14-3-3 and PRKG1 directly bind to Snail.
4. Demonstration MCF-10A cells undergo dramatic morphologic EMT in the presence of WT Snail expression, with the downregulation of the Cell adhesion molecule E-cadherin and the upregulation of the mesenchymal markers vimentin and fibronectin.
5. Identification of a double E-box binding site for Twist, which is separate and distinct from that of Snail.
6. Demonstration that cooperative binding to the double E-box is mediated by the WR domain.
7. Discovery that LIMD2 induces the EMT phenotype and may be an effector of the Snail induced morphogenetic pathway.

REPORTABLE OUTCOMES:

- Published and in preparation, in revision, in review manuscripts.

Published manuscripts:

A method for in silico identification of SNAIL/SLUG DNA binding potentials to the E-box sequence using molecular dynamics and evolutionary conserved amino acids.
Prokop JW, Liu Y, Milsted A, Peng H, Rauscher FJ 3rd.
J Mol Model. 2013 Sep;19(9):3463-9. doi: 10.1007/s00894-013-1876-y. Epub 2013 May 25. PMCID: PMC3745821.

LIMD2 is a small LIM-only protein overexpressed in metastatic lesions that regulates cell motility and tumor progression by directly binding to and activating the integrin-linked kinase.
Peng H, Talebzadeh-Farrooji M, Osborne MJ, Prokop JW, McDonald PC, Karar J, Hou Z, He M, Kebebew E, Orntoft T, Herlyn M, Caton AJ, Fredericks W, Malkowicz B, Paterno CS, Carolin AS, Speicher DW, Skordalakes E, Huang Q, Dedhar S, Borden KL, Rauscher FJ 3rd.
Cancer Res. 2014 Mar 1;74(5):1390-403. doi: 10.1158/0008-5472.CAN-13-1275. PMID: 24590809.

Snail Recruits Ring1B to Mediate Transcriptional Repression and Cell Migration in Pancreatic Cancer Cells.
Chen J, Xu H, Zou X, Wang J, Zhu Y, Chen H, Shen B, Deng X, Zhou A, Chin YE, Rauscher FJ 3rd, Peng C, Hou Z.
Cancer Res. 2014 Aug 15;74(16):4353-63. doi: 10.1158/0008-5472.CAN-14-0181. Epub 2014 Jun 5. PMID: 24903147.

CONCLUSION:

In summary, we have been highly motivated to describe completely the pathways controlled by Snail through activation of gene expression to regulate EMT and metastasis in breast cancer cells. Our preliminary studies have shown that targeting of this transcription factor in the highly aggressive, mesenchymal-like, metastatic human breast cancer cells should lead to profound reversal of the invasive phenotype.

So What?

These results together suggest that a small molecule therapeutics strategy targeting Snail or the kinases that post-translationally modify these proteins and other biomolecules are critical in the regulation of EMT and metastasis and that it would be a fruitful strategy to move forward with. This work provides the critical "proof-of-principle" for undertaking these future studies.

REFERENCES: none

SUPPORTING DATA:

List of Personnel – Paid from Research Effort:

Name	Role
Rauscher, Frank	Principal Investigator
Ayyanathan, Kasarajan	Visiting Scholar
Ho, Jenny	Part-time Summer Lab Assistant
Liu, Yuanjie	Postdoctoral Fellow
Peng, Hongzhuang	Senior Staff Scientist
Reddy, Nithin	Part-time Summer Lab Assistant

APPENDICES:

Appendix 1
Prokop JW, Liu Y, Milsted A, Peng H, Rauscher FJ 3rd. A method for in silico identification of SNAIL/SLUG DNA binding potentials to the E-box sequence using molecular dynamics and evolutionary conserved amino acids. J Mol Model. 2013 Sep;19(9):3463-9. doi: 10.1007/s00894-013-1876-y. Epub 2013 May 25. PMCID: PMC3745821.

Appendix 2
Peng H, Talebzadeh-Farrooji M, Osborne MJ, Prokop JW, McDonald PC, Karar J, Hou Z, He M, Kebebew E, Orntoft T, Herlyn M, Caton AJ, Fredericks W, Malkowicz B, Paterno CS, Carolin AS, Speicher DW, Skordalakes E, Huang Q, Dedhar S, Borden KL, Rauscher FJ 3rd. LIMD2 is a small LIM-only protein overexpressed in metastatic lesions that regulates cell motility and tumor progression by directly binding to and activating the integrin-linked kinase. Cancer Res. 2014 Mar 1;74(5):1390-1403. PMID: 24590809.

Appendix 3
Chen J, Xu H, Zou X, Wang J, Zhu Y, Chen H, Shen B, Deng X, Zhou A, Chin YE, Rauscher FJ 3rd, Peng C, Hou Z. Snail Recruits Ring1B to Mediate Transcriptional Repression and Cell Migration in Pancreatic Cancer Cells. Cancer Res. 2014 Aug 15;74(16):4353-63. doi: 10.1158/0008-5472.CAN-14-0181. Epub 2014 Jun 5. PMID: 24903147.

Published in final edited form as:
J Mol Model. 2013 September ; 19(9): 3463–3469. doi:10.1007/s00894-013-1876-y.

A method for *in silico* identification of SNAIL/SLUG DNA binding potentials to the E-box sequence using molecular dynamics and evolutionary conserved amino acids

JW Prokop[a,*], **Y Liu**[b], **A Milsted**[a], **H Peng**[b], and **FJ Rauscher III**[b]
[a]Program of Integrated Bioscience, The University of Akron, Akron OH 44325
[b]The Wistar Institute, Philadelphia PA 19104

Abstract

Binding of transcription factors to DNA is a dynamic process allowing for spatial- and sequence-specificity. Many methods for determination of DNA-protein structures do not allow for identification of dynamics of the search process, but only a single snap shot of the most stable binding. In order to better understand dynamics of DNA binding, as a protein encounters its cognate site, we have created a computer based DNA scanning array macro which sequentially inserts high affinity DNA consensus binding site at all possible locations in a predicted protein-DNA interface. We show that using short molecular dynamic simulations at each location in the interface, energy minimized states and decreased movement of evolutionary conserved amino acids can be readily observed and used to predict the consensus binding site. This macro is applied to SNAIL class C2H2 zinc finger family proteins. The analysis suggests that 1) SNAIL binds to the E-box in multiple states during encounter with its cognate site; 2) several different amino acids contribute to the E-box binding in each state; 3) the linear array of zinc fingers contributes differentially to overall folding and base-pair recognition, and; 4) each finger may be specialized for stability and sequence specificity. Moreover, the macromolecular movement observed using this dynamic approach may allow the NH2-terminal finger to bind without sequence specificity yet result in higher binding energy. This macro and overall approach could be applicable to many evolutionary conserved transcription factor families and should help elucidate better the varied mechanisms used for DNA sequence specific binding.

Keywords

SNAIL; zinc finger recognition; E-box binding; transcription factor binding; protein-DNA dynamics

Introduction

How eukaryotic transcription factors find their cognate DNA binding sites in the promoters/enhancers of target genes is critical to understanding how these factors regulated gene expression and hence cellular phenotype. This function is defined in the context of chromatin and large excess of non-specific DNA binding sites in the nucleus. Further complicating our understand is the transcription factors' expression patterns [1], disorder of the protein structure [2], and the complicated interactions each protein has [3,4] in a network which may influence DNA binding specificity. With many proteins we have the power to predict protein structure and function [5,6]; however, less work has been done to use

[*]**Corresponding author:** Jeremy W Prokop (jwp7@zips.uakron.edu), telephone (330)573-3176, fax (330)972-8445.

computer algorithms to predict how transcription factors finds and regulate a high affinity recognition site. A short DNA sequence for transcription factor binding has a high probability of being ubiquitous in the genome (with hundreds of thousands of sites), yet only a fraction of these sites are occupied by the transcription factor. Even fewer of these occupied sites yield regulatory roles [7]. This suggests that binding is a dynamic process, and that recruitment and chromatin state (spatial components) play a large role. Transcription factors have the ability to "bind and search" in a scanning mechanism in the nucleus/DNA to identify DNA sequence and the proteins exist in multiple states of DNA interaction [8–11]. These all point to a very dynamic role of proteins in how they interact and identify their consensus sequence which is undervalued in solid state or solution structures of DNA-protein complexes.

In this study we have employed *in silico* biology to develop a new DNA scanning array program of a consensus DNA sequence through a model of a protein-DNA structure to identify dynamic binding properties. We have elected to use the zinc finger transcription factor SNAIL and its homologous family to address the potential mechanism involved in DNA binding to the E-box sequence CAGGTG [12]. SNAIL contains four canonical zinc fingers, each independently folded into a helix and an antiparallel beta sheet when complexed to zinc. SNAIL contains three C2H2 domains, which have two cysteines (C) and two histidines (H) that directly coordinate zinc. A fourth finger is an atypical C3H domain. While the cognate E-Box site bound by SNAIL has been known for more than 15 years, remarkably very little is understood about how SNAIL interacts with DNA, what specific amino acids give it sequence specificity, and what the structure of the complex is. Herein we propose the first molecular model for SNAIL-E-box binding using this molecular dynamics based approach. This modeling approach will allow for many future studies at the bench top to be performed on the SNAIL-E-box complex, allowing for a better understanding of how SNAIL is involved in cancer and cancer metastasis. Although SNAIL was chosen for this study, the approach can be applied to many other transcription factor families with the hopes of elucidating mechanisms and networks of transcriptional activation and repression.

Methods

Modeling of the SNAIL protein with DNA

The model of SNAIL, containing 4 Zn fingers bound to DNA, was created using PDB structures 1tf3 (TFIIIA protein, for zinc fingers 1 and 2 of SNAIL) and 2i13 (an artificial Zn finger, for Zn fingers 3 and 4 of SNAIL) based on the most homologous sequence of known zinc fingers in the pdb to each one of the four zinc fingers of SNAIL. To create the homology model of SNAIL, fingers 1 and 2 of 1tf3 were parsed and amino acids changed to those present in fingers 1 and 2 of SNAIL based on CLUSTAL [13] sequence alignment. These fingers (from 1tf3) were then aligned to the 2nd and 3rd fingers of the structure 2il3 using Mustang [14]. Zn fingers 1, 2, 3 and 6 of 2i13 were then deleted. The remaining Zn fingers (4 and 5) were mutated *in silico* to the amino acids matching fingers 3 and 4 of SNAIL as determined by alignments. A bond was created joining the 2nd finger's carboxyl terminus (originally from structure 1tf3) to the amino terminus of finger 3 (from 2i13). This resulted in a structure contain 4 fingers, each with its own Zn ion, and a 20 mer DNA sequence (the original recognition sequence from 2i13). The linker domain between fingers 1 and 2 was corrected for SNAIL, which contains a slightly shortened linker domain compared to most Zn fingers. The final SNAIL sequence (Figure 1), including the deletion of two amino acids from structure 1tf3, was energy minimized to relieve bond stress. All atoms were then freed and energy minimizations performed multiple times using AMBER03 force field [15] with 0.997 g/mL of water. ConSurf analysis [16] for SNAIL, SNAIL like, Slug and Scratch proteins was done using ClustalW [13] alignment with default settings and Bayesian method with the T92 model for phylogenetic analysis.

Creation and molecular dynamics simulations of the DNA scanning array

To the modeled protein-DNA structure above the DNA sequence was changed four times to one of four DNA bases (A, T, C, or G) on the first DNA strand of the structure, with the complement base placed on the opposite strand using the "*6bp Scanning Array*" macro (Supplemental file pages 11–68) in YASARA [17], so that all the DNA of one strand is the same. This creates a background with all possible DNA bases at each site (four backgrounds), which serves as a relative starting point for stability when scanning the consensus binding sequence. The macro then inserts the 6 base pair consensus sequence (CAGGTG, used in this study) at multiple sites on each DNA strand with all four DNA base backgrounds, generating four YASARA scene files for each location. Changes to only the first lines of the macro are required for changing the sequence of the consensus site, to allow for the use with other proteins. At each site the "*md and analysis for scanning array*" macro (supplemental file pages 69–121) was run. Details of the conditions for all molecular dynamics simulations can be found in the macro scripts. The macro runs a 500 picosecond md simulation, which then calculates multiple root mean squared deviations (RMSD). The score for each amino acid from the Consurf analysis ($Cons_R$, on a scale of 1–9 with 9 being highly conserved and 1 having no conservation, Figure S1) was divided by the averaged heavy atom RMSD (movement of residue, M_R) of each amino acid in the four DNA base pair backgrounds at each location of the consensus (Figure 2 Equation 1) yielding X_{RZ} (amino acid R's conserved movement, at the Z position of the consensus sequence). This quantity was then averaged for all the amino acids in SNAIL (Figure 2 Equation 2) and normalized to the value from the control DNA sequences without a consensus (Figure 2 Equation 3) yielding the $ConSurf_{relz}$ (The Consurf total relative to the Z position of the consensus sequence). For the DNA base pair movement, each of the bases' RMSD from consensus sequence (M_{DZ}) was divided by the same location on the control DNA sequence (M_{DO}) to give the relative movement (M_{DNAZ}) (Figure 2 Equation 4). This was then averaged for the six bases of the consensus sequence (Figure 2 Equation 5) yielding DNA_{relz}. The compiled movement (R_z) was calculated by dividing the $ConSurf_{relz}$ by the DNA_{relz} (Figure 2 Equation 6) and the positions ±1 (Figure 2 Equation 7) or ±2 (Figure 2 Equation 8) of the R_z were added to address dynamic binding and transitions. The models were analyzed (z-score) using YASARA2 force field and model quality (normality of dihedrals and packing).

Results

The overall approach was to create a program which could help illuminate the dynamic mechanism(s) by which a transcription factor binds a well-established DNA consensus sequence using a computational approach which could be advanced to high throughput methodology. This method should be able to confirm or identify the energy stable binding state, while also addressing the dynamic energy landscape in which that binding is directed. On the energy minimized structure of SNAIL, Zn finger 1 was found around DNA base pairs C13–C16, finger 2 at C9–C12, finger 3 at C6–C8, and finger 4 at C3–C5. The combination of the two macros used to study the Zn finger SNAIL binding to the E-box sequence CAGGTG, allowed for placement of the consensus sequence at each location throughout the DNA strand and studied the energetics of the DNA-protein interaction at each location, with the consensus sequence (C of CAGGTG) starting at C1 (C1–C6), C2 (C2–C7),…., C15 (C15–C20) or on the D strand of DNA. Potential energies of the modeled SNAIL protein from the multiple consensus locations showed few sites (D32 and D33) to have lower values, suggesting most of the SNAIL structures are stabilized and similar in all consensus locations (Figure S2A). Binding energy of SNAIL to DNA were not altered much for the multiple locations, with the highest value at the location C3 (Figure 3). The relative conserved movement (R_z) was highest at the C4 location (1.12581), with a high value also

found at C13 (1.114211, Figure 4A). A value of 1 represents the R_z calculated for the DNA which did not contain a consensus sequence. The value of Z ± 1 (which allows a single metric for multiple binding locations to identify energy profiles with multiple sites in close proximity with high binding energies) enabled identification of two areas of probable binding, with the consensus sequence starting around C3–C5 and again around C12 to C13 (Figure 4B). The Z ± 2 values are similar to the Z ± 1 values (Figure S2B–C), and therefore do not allow further identification of possible sites. Future experiments should allow for optimizing the Z ± 1 value for quick identification of probable binding sites.

To address the specifics of why these two locations (C4 and C13) are highly favored, individual components of our calculations were studied. The DNA movement (DNA_{relz}) of the consensus sequence, which shows the stability of the consensus base pairs in simulations, was lowest when the consensus started at both C4 and the C13 locations (Figure S3), suggesting increased stability due to the protein environment at that location. Next, the X_{RZ} scores were used to address the amino acids contacting the consensus sequence at each of the two highly favored sites. The score was averaged for all of the consensus locations in the array for each amino acid (Figure 5A) to understand the intrinsic variation of the assay. A difference from each location of the array to this average allows for identification of potential amino acid leading to specificity of binding and stability of the DNA base pair movement. For the consensus found with the E-box sequence at the C4 location, the largest variation was found at amino acids 246 and 247 with a difference greater than 2 (Figure 5B). Two times the standard error of the average for these amino acids was only 0.322957 and 0.295983, suggesting a strong significance of the difference seen at the C4 location from all the other locations. Both of these amino acids are found in the fourth zinc finger with a highly conserved Ser at 246 and an Arg at 247. The hydroxyl group on the Ser likely hydrogen bonds with the phosphate backbone, putting the Arg side chain in the perfect location to hydrogen bond with the first or second G of the consensus sequence (CA**GG**TG, Figure 6). The average structures for each of the backgrounds with the C4 location confirm the proper distance is maintained in all simulations (Figure S4A). Mutation of these two amino acids significantly altered the ability to shift the E-box sequence (Figure S4B). Tracking the score of Ser 246 and Arg 247 over the entire consensus locations shows that they are higher at locations flanking the E-box consensus at the C4 location (C1–C5) and no other consensus locations (Figure S4C).

Besides Ser 246 and Arg 247 of finger 4, several amino acids are found in the normal positions for DNA specificity of a zinc finger. These amino acids include Met 171 of finger 1, Arg 191/Trp 193 of finger 2, Ser 221 of finger 3, and Met 248/Ser 249 of finger 4. Met 171, Arg 191, and Met 248 have no increased X_{RZ} score for any of the consensus locations (Figure S5), thus suggesting they do not contribute to specificity of binding. Trp 193 has an increased X_{RZ} score around the E-box sequence at the C9 location with additional elevated levels for C8, C11 and C12. These locations put the E-box between fingers 2 and 3. Ser 221 had an elevated X_{RZ} score at E-box location C2 and D35 while Ser 249 showed elevated values at C14. As the E-box locations C12–C13 yielded the second highest Z ± 1 score, the difference between each consensus X_{RZ} score and the control X_{RZ} score was observed. None of the amino acids around finger one or two (amino acids 156–202) contributed to the elevation of the X_{RZ} score (Figure S6). However, using heavy atom RMSD values for these consensus locations showed that the linker region between fingers 1 and fingers 2 (amino acids 176–180) were lower and therefore more stable in DNA interaction with consensus sites placed starting at C12, C13, and C14 (Figure S7, this linker region is not conserved in the SNAIL like and Scratch proteins and therefore received no scoring in the X_{RZ} scores based on ConSurf results). In summary it appears that critical contact residues for sequence specificity are Ser 246 and Arg 247, with additional contributions from the first linker

domain and amino acids Trp 193, Ser 221, and Ser 249. This dynamic approach taken here suggests a multiple state binding for SNAIL.

Discussion

Here we introduce a program which should be useful in localizing DNA-protein contacts in a cognate transcription factor-DNA complex. Many structure determination methods are based on static, non-biological conditions, yet biology functions in dynamic processes in aqueous environments. Thus, new approaches are needed to address how proteins bind to DNA. Recent evidence has supported a role of zinc fingers binding and shuffling on DNA at a nanosecond timescale in an asymmetrical role[11]. Many solid state structures lack this dynamic component, thus there is a clear need for tools to help determine these dynamic processes. *In silico* experiments provide a low cost preliminary data generation relative to expensive NMR and protein purification, allowing scientists to screen a larger dataset of proteins for their potential role in these dynamic processes. By using either known or modeled structures of proteins interacting with DNA, it is possible to elucidate dynamic binding mechanisms.

SNAIL is involved in the epithelial-mesenchymal transition (EMT) involved in breast [18] and prostate [19] cancers. This involvement in cancer regulation is through an E-box sequence with a consensus of six base pairs (CAGGTG). The canonical model shows that each zinc finger domain recognizes three DNA base pairs [20]. SNAIL contains four, highly conserved fingers, suggesting that only some of the fingers are involved in recognition of the E-box sequence. Until now it has been unclear which fingers contact which bases in the E-box consensus. We employed an *in silico* approach, creating a consensus sequence array using molecular dynamics simulations and evolutionary conservation to determine a likely mechanism of binding. Our initial model before the scanning array yielded a z-score of −0.997 and with the consensus location starting at C4 of −0.891. Z-score values between 0 and −2 are considered fair, and the calculations of the z-scores for the original structures are −1.261 for 1tf3 and −0.433 for 2i13. Although subtle variations in the structure may exist than those of our model for SNAIL, it provided an opportunity to test the macro created in this paper while providing significant hypothesis generation that can be tested in benchtop experiments for the SNAIL system. Our final models of SNAIL bound to DNA provide similar results (potential and binding energies) to other Zn fingers complexed to DNA and thus provide a strong starting point for threading the E-box consensus sequence through the DNA of our structure. Further comparisons of SNAIL structure to other Zn finger proteins binding different consensus sequences and other non-Zn finger proteins binding to the E-box sequence will be detailed in subsequent publications. In addition, the role of amino acid 246 and 247 have been confirmed through use of bench top methods.

Although details of this method are still being actively modified for use with other proteins, the length of md simulation (500ps) in this manuscript appears to allow for stability of the protein in simulation, as can be seen in the representative plots for both energy and carbon alpha RMSD of the C4 location (Figure S8). This time was initially determined as RMSD values stabilized around 100 picoseconds of SNAIL interaction with random DNA, allowing for full 400 picoseconds of stabilized simulations. This amount of time we suggest optimizes for detection of stability of critical amino acids, while reducing the computational requirements on performing hundreds of molecular dynamics simulations for longer periods of time. Each protein-DNA complex will need to be adjusted and validated for the length of simulations to allow for stability of the complex.

Results from this novel *in silico* approach suggests that binding likely occurs through the fourth finger with conserved contacts created by amino acids Ser 246 and Arg 247.

Additional contacts on fingers two and three maintain a shuffling protein around this consensus sequence. As this small shuffle of SNAIL on the DNA E-box sequence takes place, finger one may stabilize and bind tighter with minimal to no sequence specificity, through electrostatic interaction of the polar basic amino acids of SNAIL with the phosphate backbone of the DNA. As the linker domain between fingers one and two is smaller and atypical from most other zinc finger proteins, it may require more time to stabilize with the phosphate backbone. This mechanism would allow for the SNAIL proteins to scan the DNA and when the DNA consensus identified, SNAIL to reduce its movement to a shutter around the site and have a tighter binding through stabilization of finger one (Figure 7). Interactions of this linker domain with 14-3-3 [21] may alter the stability and binding affinity of SNAIL to E-box sequence. Similar mechanisms may also exist to that previously shown for Egr-1 [11] allowing finger one of SNAIL to bind to another strand of DNA, translocating the other fingers to another DNA strand.

These results, in this study, provide the first molecular model for SNAIL class zinc finger protein binding to the E-box, a DNA sequence of high biological and disease relevance. This model may serve as a tool in understanding cancer progression and metastasis, allowing for drug design to specific amino acids. The data generated using this *in silico* DNA scanning array can be used to suggest amino acids to mutate for benchtop analysis. At minimal cost (excluding time to run simulations and building computers) to perform, this allows for a screening of potential mechanisms that are conserved in a family of transcription factors. When combined with mutagenesis, DNA gel shift assays, NMR and other molecular/biochemical techniques, it provides a strong addition to our understanding of protein-DNA interactions in a dynamic, rather than static, process.

Conclusions

Use of this *in silico* DNA scanning array has elucidated a potential mechanism for SNAIL-E-box sequence specificity. A multiple state binding mechanism appears likely, in which the first of four zinc fingers on SNAIL may form tight complexes when fingers 2–4 complex to the E-box. The use of these methods can be applied to all transcription factor families with hopes of further explaining DNA sequence specificity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

JWP is funded by the Ohio Board of Regents, American Heart Predoctoral fellowship, and The University of Akron. Work in the Rauscher laboratory is supported by NIH grants CA129833, CA010815, CA163761, DOD-BCRP W81XWH-11-1-0494, The Samuel Waxman Cancer Research Foundation, SUSAN B. KOMEN FOR THE CURE and The Noreen ONeill Foundation for Melanoma Research. YL is supported by NCI Training Grant: T32 CA09171.

References

1. Gilad Y, Oshlack A, Smyth GK, Speed TP, White KP. Expression profiling in primates reveals a rapid evolution of human transcription factors. Nature. 2006; 440:242–245. [PubMed: 16525476]
2. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic Disorder in Transcription Factors. Biochemistry. 2006; 45:6873–6888. [PubMed: 16734424]
3. Spirin V, Mirny LA. Protein complexes and functional modules in molecular networks. Proc Natl Acad Sci USA. 2003; 100:12123–12128. [PubMed: 14517352]

4. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004; 5:101–113. [PubMed: 14735121]
5. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protocols. 2010; 5:725–738.
6. Baker D, Sali A. Protein Structure Prediction and Structural Genomics. Science. 2012; 294:93–96. 2001. [PubMed: 11588250]
7. Fisher WW, Li JJ, Hammonds AS, Brown JB, Pfeiffer BD, Weiszmann R, et al. NA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in Drosophila. Proc Natl Acad Sci. 109:21330–21335. [PubMed: 23236164]
8. Misteli T. Protein dynamics: implications for nuclear architecture and gene expression. Science. 2001; 291:843–847. [PubMed: 11225636]
9. Catez F, Lim JH, Hock R, Postnikov YV, Bustin M. HMGN dynamics and chromatin function. Biochem Cell Biol. 2003; 81:113–122. [PubMed: 12897844]
10. Shav-Tal Y, Darzacq X, Singer RH. Gene expression within a dynamic nuclear landscape. The EMBO J. 2006; 25:3469–3479.
11. Zandarashvili L, Vuzman D, Esadze A, Takayama Y, Sahu D, Levy Y, Iwahara J. Asymmetrical roles of zinc fingers in dynamic DNA-scanning process by the inducible transcription factor Egr-1. Proc Natl Acad Sci USA. 2012; 109:E1724–E1732. [PubMed: 22675124]
12. Nieto MA. The snail superfamily of zinc-finger transcription factors. Nat Rev Mol Cell Biol. 2002; 3:155–166. [PubMed: 11994736]
13. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948. [PubMed: 17846036]
14. Konagurthu AS, Whisstock JC, Stuckey PJ, Lesk AM. MUSTANG: A multiple structural alignment algorithm. Proteins. 2006; 64:559–574. [PubMed: 16736488]
15. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem. 2003; 24:1999–2012. [PubMed: 14531054]
16. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. 2010; 38:W529–W533. [PubMed: 20478830]
17. Krieger E, Darden T, Nabuurs SB, Finkelstein A, Vriend G. Making optimal use of empirical energy functions: force-field parameterization in crystal space. Proteins. 2004; 57:678–683. [PubMed: 15390263]
18. Côme C, Magnino F, Bibeau F, De Santa Barbara P, Becker KF, Theollet C, Savagner P. Snail and slug play distinct roles during breast carcinoma progression. Clin Cancer Res. 2006; 12:5395–5402. [PubMed: 17000672]
19. Smith BN, Odero-Marah VA. The role of Snail in prostate cancer. Cell Adh Migr. 2012; 6:433–441. [PubMed: 23076049]
20. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science. 1991; 252:809–817. [PubMed: 2028256]
21. Hou Z, Peng H, White DE, Wang P, Lieberman PM, Halazonetis T, Rauscher FJ 3rd. 14-3-3 Binding Sites in the Snail Protein Are Essential for Snail-Mediated Transcriptional Repression and Epithelial-Mesenchymal Differentiation. Cancer Res. 2010; 70:4385–4393. [PubMed: 20501852]

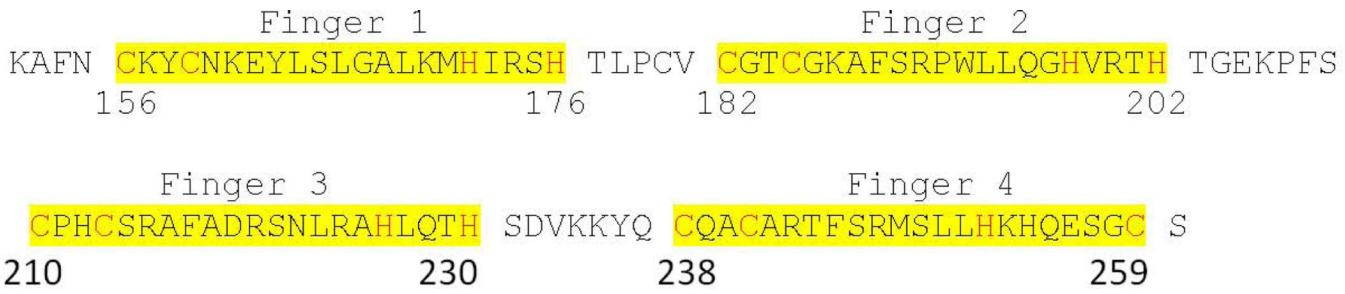

Figure 1. Sequence of SNAIL zinc fingers

Sequence used of SNAIL to make model with the four zinc fingers highlighted in yellow and the amino acids interacting with zinc shown in red. Number is based on human SNAIL.

Amino Acid Movement

1). $X_{RZ} = \frac{Cons_R}{M_R}$

2). $\overline{X}_{RZ} = \frac{1}{n}\sum_{i=1}^{n} X_{RZ}$

3). $ConSurf_{relz} = \frac{\overline{X}_{RZ}}{\overline{X}_{RO}}$

DNA Movement

4). $M_{DNAZ} = \frac{M_{DZ}}{M_{DO}}$

5). $DNA_{relz} = \overline{M}_{DNAZ} = \frac{1}{n}\sum_{i=1}^{n} M_{DNAZ}$

Compiled Movement

6). $R_z = \frac{ConSurf_{relz}}{DNA_{relz}}$

7). $Z\pm1 = R_{(Z+1)}+R_{(Z)}+R_{(Z-1)}$

8). $Z\pm2 = R_{(Z+2)}+R_{(Z+1)}+R_{(Z)}+R_{(Z-1)}+R_{(Z-2)}$

Figure 2. Math equations used in the DNA scanning array analysis

$Cons_R$ = conservation score of each residue (based on ConSurf results), M_R = heavy atom RMSD of each residue, X_{RZ} = conserved movement for amino acid R with consensus at the Z location. X_{RO} = conserved movement for amino acid R with no consensus sequence, $ConSurf_{relz}$ = relative conserved movement of the protein with the consensus at the Z location, M_{DZ} = movement of one of the bases of the consensus site of the DNA, M_{DO} = movement of the base M_{DZ} in the background with no consensus sequence, M_{DNAZ} = relative movement of the consensus DNA sequence. DNA_{relz} = average movement relative to the background of the consensus sequence at each position (Z), R_z = score for conserved movement of the protein and the DNA at position Z of the consensus sequence, Z±1(2) = additive scores for positions plus or minus one site (or 2) around the Z site of the consensus location.

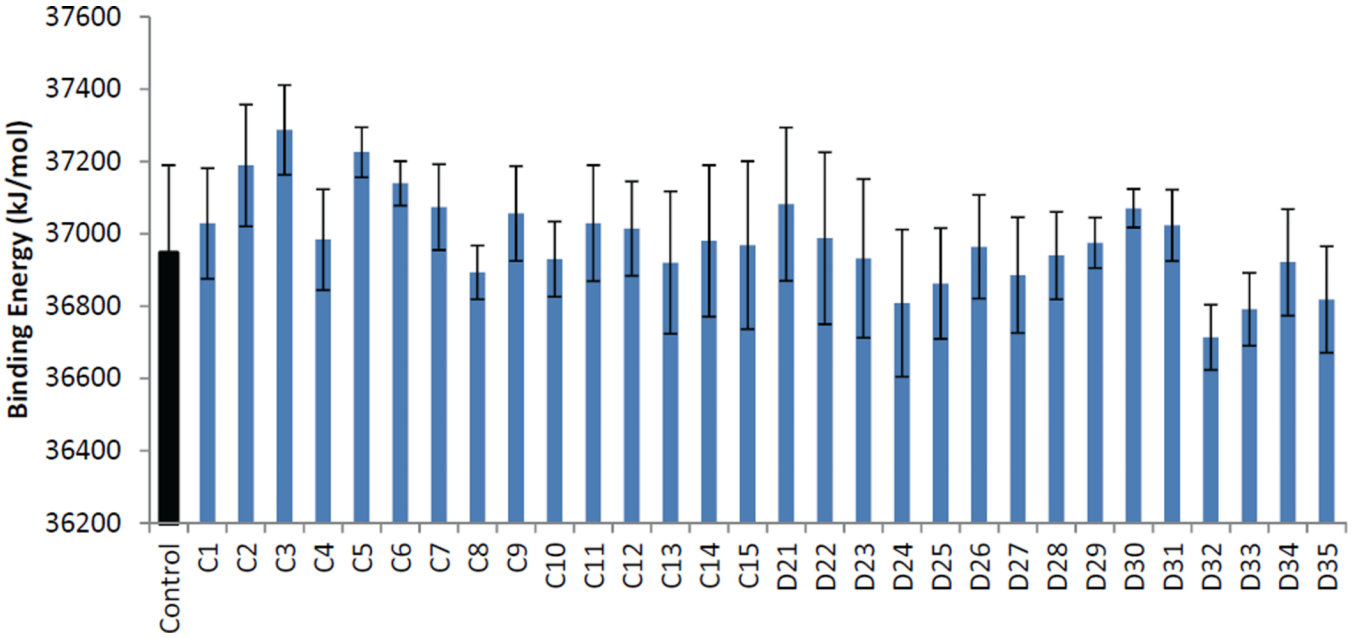

Figure 3. Binding energies of SNAIL at each location of the consensus DNA or the control

Binding energy of SNAIL to DNA from the energy minimized structure for each background at various locations of the two strands (C1–C15 and D21–D35) as determined in YASARA with AMBER03 force field.

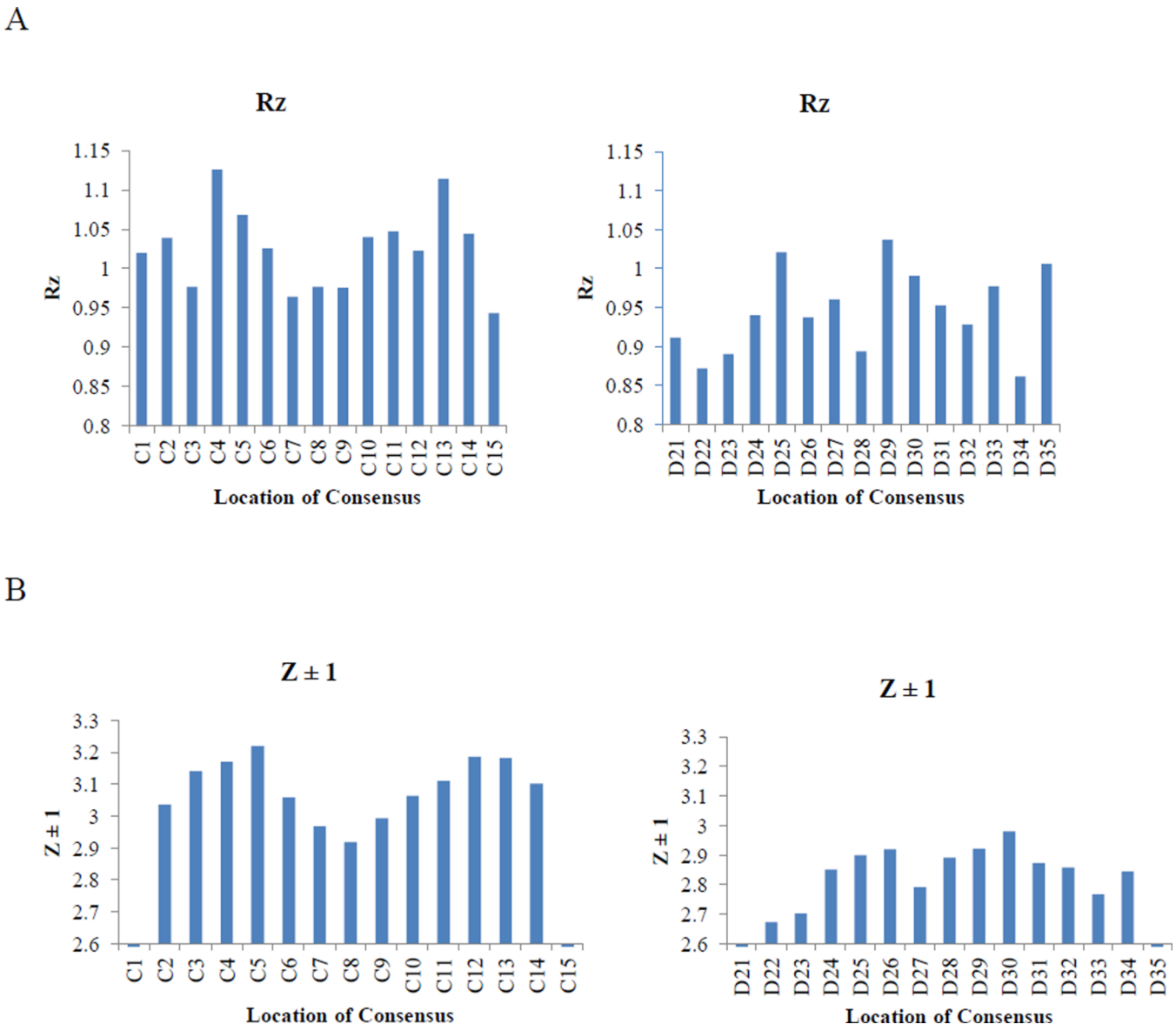

Figure 4. Calculated conserved movement scores

A) The relative conserved movement (R_z) for each site of the consensus on the C DNA strand (left) or the D strand (right). **B)** Adding the R_z to the consensus shifted to the left or right by one shows two highly conserved favorable binding sites on the C strand (left) while none on the D (strand). A value of 1 for **A** or 3 for **B** would be the score of the control.

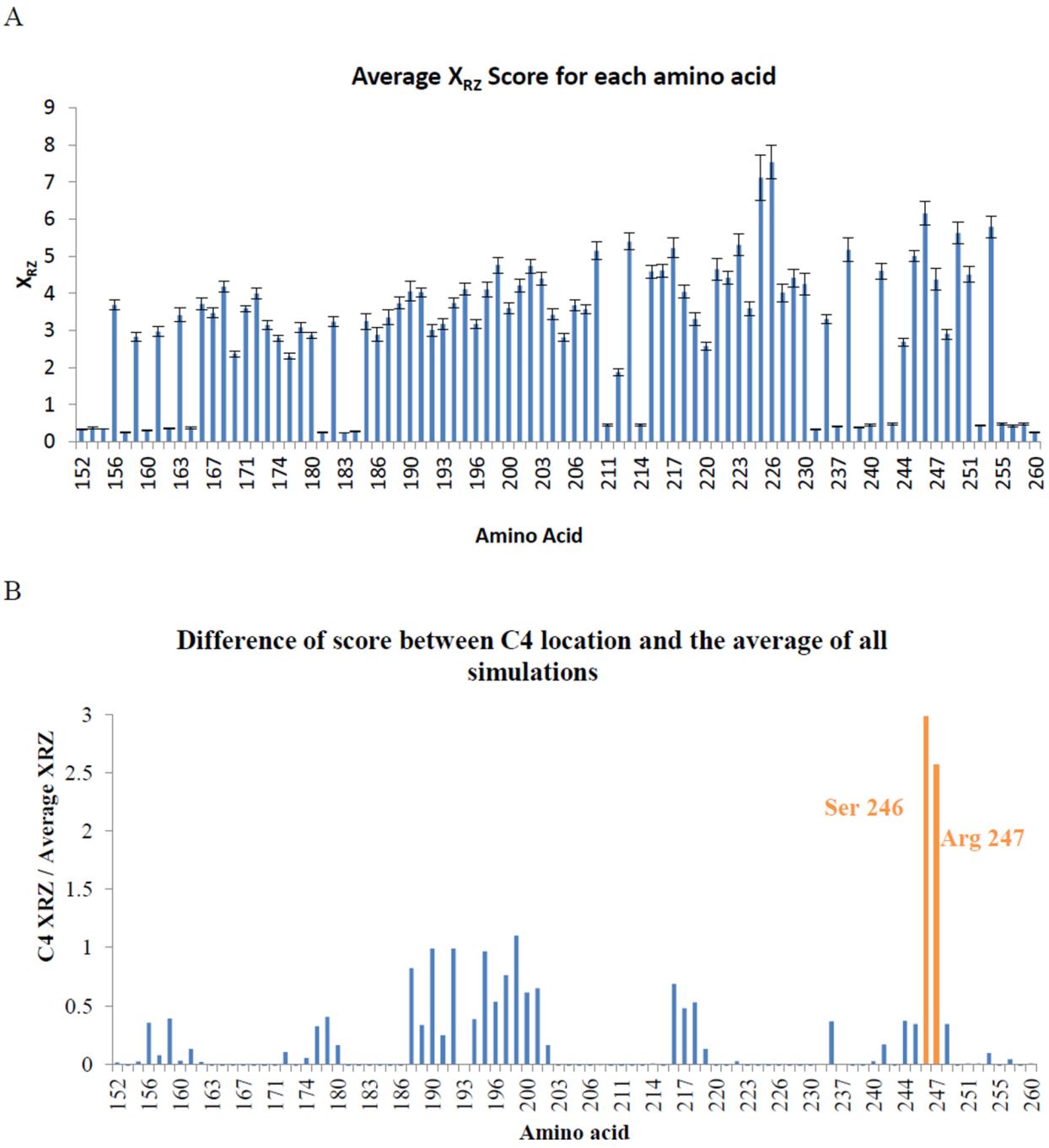

Figure 5. X_{RZ} score for each amino acid

A) Average X_{RZ} for each amino acid over all locations of the DNA consensus sequence with the error bars representing plus or minus the standard error times two of the 30 locations of the consensus as well as the no consensus control. B) X_{RZ} score for the C4 location of the consensus minus the average for each amino acid shown in A. The largest differences were seen at amino acids 246 and 247 (orange).

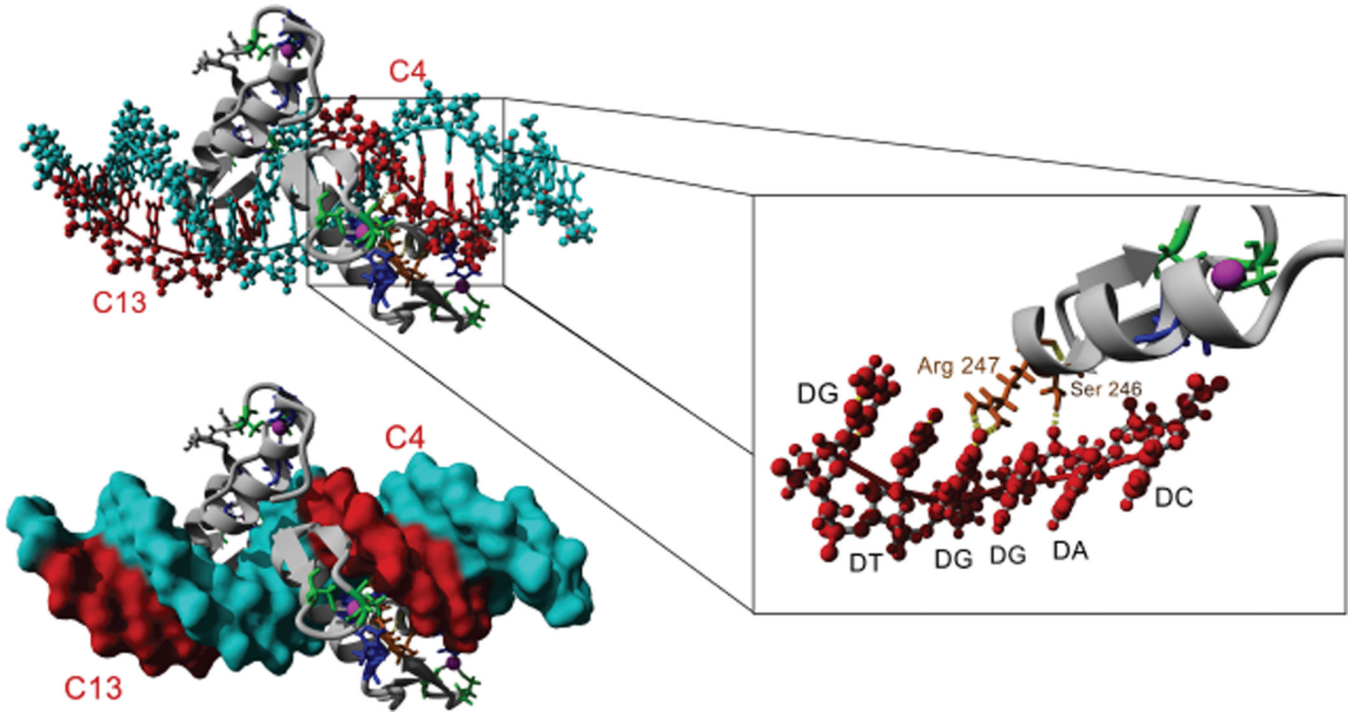

Figure 6. C4 and C13 consensus locations on the DNA

Top shows the DNA in ball and sticks while the bottom shows the DNA molecular surface. Highlighted in the box is the location of the Arg 247 interacting with the DNA consensus sequence of the C4 location while Ser 246 interacts with the phosphate backbone allowing for stability of Arg 247.

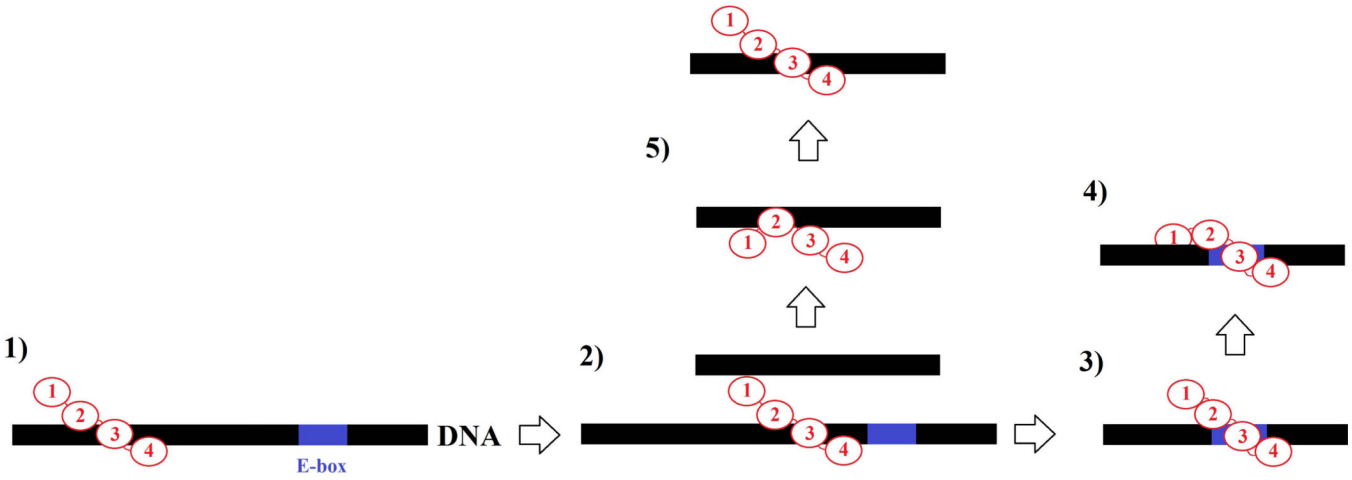

Figure 7. Potential DNA binding mechanism of SNAIL family

1) Fingers 2–3 of SNAIL form weak complex with DNA with no sequence specificity. **2)** SNAIL moves on the DNA. **3)** SNAIL binds to the E-box sequence and stabilizes. **4)** Decreased dynamics allows finger 1 to stabilize and tightly bind the E-box sequence. **5)** With finger one unbound, it may be able to translocate SNAIL to another DNA strand similar to mechanisms seen in Egr-1.

Cancer Research

LIMD2 Is a Small LIM-Only Protein Overexpressed in Metastatic Lesions That Regulates Cell Motility and Tumor Progression by Directly Binding to and Activating the Integrin-Linked Kinase

Hongzhuang Peng, Mehdi Talebzadeh-Farrooji, Michael J. Osborne, et al.

Cancer Res 2014;74:1390-1403.

Cancer Research

LIMD2 Is a Small LIM-Only Protein Overexpressed in Metastatic Lesions That Regulates Cell Motility and Tumor Progression by Directly Binding to and Activating the Integrin-Linked Kinase

Hongzhuang Peng[1], Mehdi Talebzadeh-Farrooji[6], Michael J. Osborne[6], Jeremy W. Prokop[4], Paul C. McDonald[7], Jayashree Karar[1], Zhaoyuan Hou[1], Mei He[3], Electron Kebebew[3], Torben Orntoft[5], Meenhard Herlyn[1], Andrew J. Caton[1], William Fredericks[2], Bruce Malkowicz[2], Christopher S. Paterno[1], Alexandra S. Carolin[1], David W. Speicher[1], Emmanuel Skordalakes[1], Qihong Huang[1], Shoukat Dedhar[7], Katherine L.B. Borden[6], and Frank J. Rauscher III[1]

Abstract

Proteins that communicate signals from the cytoskeleton to the nucleus are prime targets for effectors of metastasis as they often transduce signals regulating adhesion, motility, and invasiveness. LIM domain proteins shuttle between the cytoplasm and the nucleus, and bind to partners in both compartments, often coupling changes in gene expression to extracellular cues. In this work, we characterize LIMD2, a mechanistically undefined LIM-only protein originally found to be overexpressed in metastatic lesions but absent in the matched primary tumor. LIMD2 levels in fresh and archival tumors positively correlate with cell motility, metastatic potential, and grade, including bladder, melanoma, breast, and thyroid tumors. LIMD2 directly contributes to these cellular phenotypes as shown by overexpression, knockdown, and reconstitution experiments in cell culture models. The solution structure of LIMD2 that was determined using nuclear magnetic resonance revealed a classic LIM-domain structure that was highly related to LIM1 of PINCH1, a core component of the integrin-linked kinase–parvin–pinch complex. Structural and biochemical analyses revealed that LIMD2 bound directly to the kinase domain of integrin-linked kinase (ILK) near the active site and strongly activated ILK kinase activity. Cells that were null for ILK failed to respond to the induction of invasion by LIMD2. This strongly suggests that LIMD2 potentiates its biologic effects through direct interactions with ILK, a signal transduction pathway firmly linked to cell motility and invasion. In summary, LIMD2 is a new component of the signal transduction cascade that links integrin-mediated signaling to cell motility/metastatic behavior and may be a promising target for controlling tumor spread. *Cancer Res; 74(5); 1390–403. ©2014 AACR.*

Introduction

Defining the complex biology and the cascade of events that lead to metastatic spread of primary tumors, both locally and to distant sites, continue to be major unmet needs in cancer biology (1). Moreover, defining which molecular events in both the metastatic cascade and in the maintenance of tumor dormancy are targetable for therapeutic or preventative benefit is an even more daunting task. These results have defined molecules that control a large array of cellular phenotypes, including cell motility, cell–cell and cell–matrix interactions, and immune evasion (2, 3). Generally, it seems that rare metastatic variants appear stochastically in a genetically heterogeneous primary tumor, occurring quite early in tumor progression, and that normal developmental processes, such as epithelial–mesenchymal transition (EMT), mesenchymal-epithelial transition (MET), and angiogenic cascades, are often ectopically activated to achieve tumor spread (1, 4, 5). Both forward genetic and descriptive experimental approaches have been utilized to identify genetic and epigenetic determinants for metastatic capability, largely by selection and analysis of metastatic variants in populations, or by comparing the expression and mutation profiles of matched primary and metastatic lesions using the vast spectrum of "-omic" technologies currently popular (6–8). The finding of metastasis-associated antigens and transcriptional and/or

Authors' Affiliations: [1]The Wistar Institute; [2]Department of Surgery, University of Pennsylvania and Veterans Affairs Medical Center, Philadelphia, Pennsylvania; [3]Center for Cancer Research, National Cancer Institute, Bethesda, Maryland; [4]University of Akron, Akron, Ohio; [5]Aarhus University Hospital at Skejby Sygehus, Denmark; [6]Department of Pathology and Cell Biology, University of Montreal, Institute for Research in Immunology and Cancer; and [7]Department of Integrative Oncology, British Columbia Cancer Research Centre, Vancouver, British Columbia, Canada

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

H. Peng and M.T.Z. Farrooji contributed equally to this work.

Corresponding Authors: Katherine L.B. Borden, Department of Pathology and Cell Biology, Institute for Research in Immunology and Cancer, University of Montreal, Canada. E-mail: katherine.borden@umontreal.ca; and Frank J. Rauscher, Wistar Institute, 3601 Spruce Street, Philadelphia, PA. E-mail: rauscher@wistar.org.

doi: 10.1158/0008-5472.CAN-13-1275

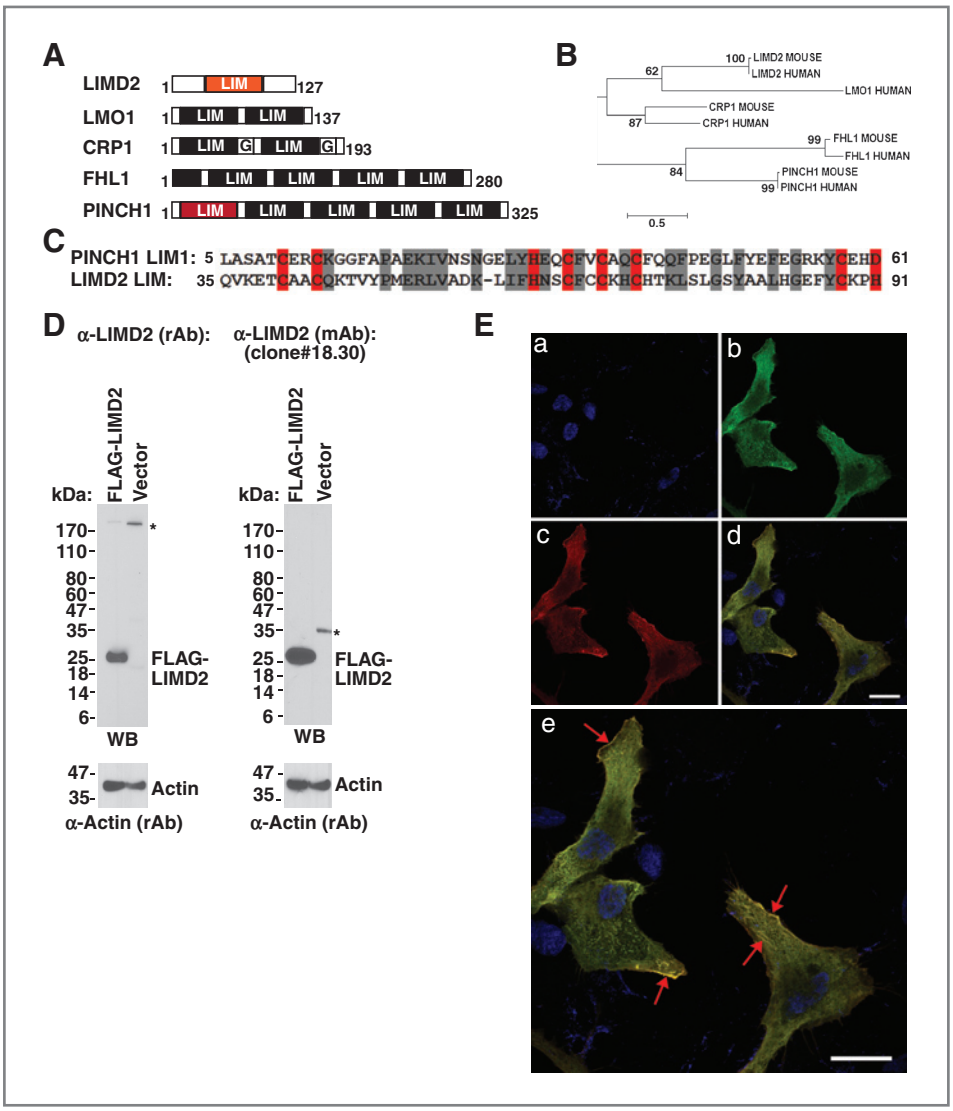

Figure 1. A, LIMD2 and the LIM-only protein family. B, LIMD2 is most closely related to CRP1. C, the PINCH1-LIM1 and LIMD2 LIM domains are homologous. The zinc-chelating residues are highlighted in red; the conserved amino acids are highlighted in gray. D, antibodies robustly detect LIMD2 protein. *, nonspecific binding. E, TPC1 cells were transfected with myc-LIMD2 then fixed and stained with both anti-myc tag antibody (red) or anti-LIMD2 mAb (green). The cells were counterstained with DAPI to highlight the nucleus (blue). The red arrows indicate concentrations of Myc- and LIMD2 costaining signal, which is at the leading edge of the cells and also is present in streaks reminiscent of vinculin-containing adhesion plaques.

genetic signatures in these comparisons has been a useful exercise, and may be useful in the clinic, but these approaches often fail to distinguish drivers of the metastatic phenotype from passenger/markers of the phenotype and, moreover, rarely led to specific mechanisms. In this study, we characterized the LIMD2 protein, which originally identified as highly and exclusively overexpressed in metastatic lesions but absent in matched normal tissue or primary tumor (9).

LIMD2 is a LIM-only domain protein that was identified as a biomarker for papillary thyroid carcinoma (PTC) lymph node metastasis (LNM) from molecular profiling of matched samples (9). LIMD2 was found to be highly expressed in LNM but absent from the primary tumor or normal thyroid tissue in matched patient PTC samples, suggesting that LIMD2 expression could provide an improved means of detecting potentially metastatic PTC cells during initial staging of a newly diagnosed carcinoma. In the human genome, there are 135 identifiable LIM-encoding sequences located within 58 genes. The LIM domain is organized as a tandem zinc-finger structure that functions as a modular protein-binding interface (Fig. 1A). LIM domain–containing proteins have diverse cellular roles such as regulators of gene expression, cyto-architecture, cell adhesion, cell motility, and signal transduction. LIM domain proteins are emerging as key molecules in a wide variety of human cancers (10). In particular, all members of the human LIM domain–only (LMO) proteins, LMO1 to LMO4, which are required for many normal developmental processes, are implicated in the onset or progression of several cancers, including T-cell leukemia, breast cancer, and neuroblastoma. Here, we report that LIMD2 regulates cell motility and is a novel effector of tumor progression via its role in the integrin-linked kinase (ILK) pathway.

Materials and Methods

Plasmids, antibodies, cell culture, transfection, migration assays, soft-agar assays, GST-binding assays, immunofluorescence, and immunohistochemistry

The plasmids pET-28a(+)-LIMD2, pGEX-4T1-LIMD2, pcDNA3.1-Flag-LIMD2, pcDNA3.1-Flag-LIMD1, and pcDNA3.1-

Flag-AJUBA were constructed through PCR-based cloning (cloning details available on request). Sport6-LIMD2 (human) and pLKO.1-shLIMD2 (human) plasmids were purchased from Open Biosystems. pET-32-TT-ILK, 1–174 (His), and pCOLA-Duet-PINCH1-LIM1 (GST) plasmids are kindly provided by David Calderwood and Titus Boggon (11). Preparation of the His-LIMD2 and GST-LIMD2 fusion proteins were performed as described previously (12) and yielded soluble protein purified under native conditions. The DNA region encoding human LIMD2 amino acids 1 to 127 was isolated via PCR with synthetic oligonucleotides containing 5′ *Bam*H1 and 3′ *Hin*dIII and the digested DNA fragment cloned into pQE30 vector at the corresponding restriction sites. This produced a protein that encoded the N-terminal vector-encoded peptide of MRGSHHHHHHGS-. Followed by AA1-127 of hLIMD2 followed by a stop codon. This was used to produce rabbit polyclonal antibody using standard techniques. For monoclonal antibody, the DNA region encoding LIMD2 (AA1-127) was isolated via PCR with synthetic oligonucleotides containing 5′ *Eco*RI and 3′ *Xho*I, and the digested DNA fragment cloned into pGEX4T-1 vector at the corresponding restriction sites. This produced a protein that encoded the N-terminal vector-encoded GST, followed by AA1-127 of hLIMD2 followed by a stop codon. The mAb used in this study was given the Wistar designation "LIMD2 clone #18.30" and was immunoglobulin G (IgG) isotype. Both sources of crude antibodies were affinity purified using either the T-Gel or the SulfoLink Kits (Pierce) using full-length, native LIMD2 protein coupled to the resin.

HEK293, U2OS, MCF-7, MDA-MB-231, MDA-MB-435, and mouse fibroblast ($ILK^{+/+}$ and $ILK^{-/-}$) cells were maintained in Dulbecco's modified Eagle medium (DMEM) containing 10% FBS, 2 mmol/L L-glutamine, and penicillin (50 units/mL)/streptomycin (50 μg/mL; P/S) at 37°C under 5% CO_2 in a humidified chamber. MCF-10A mammary epithelial cells were maintained in DMEM:Ham's F12 containing 5% house serum, epidermal growth factor (10 ng/mL), hydrocortisone (0.5 μg/mL), cholera toxin (100 ng/mL), insulin (10 μg/mL), and P/S. Melanoma cell lines (Fom196-1, RGP-WM35, RGP-WM3211, VGR-WM793, and Met1205 Lu) were maintained in Tu 2% medium containing 80% MCDB153, 20% Leibovitz's L-15, 2% FBS, insulin (bovine; 5 μg/mL), and 1.68 mmol/L $CaCl_2$. Bladder cancer cell lines (T24 and RT4) were maintained in McCoys 5A containing 10% FBS, 1.5 mmol/L L-glutamine, 2.2 g/L sodium bicarbonate, and P/S. Bladder cancer cell lines (HT1376, UM-UC3, ScaBer, J82, and TCCSUP) were cultured in Eagle's Minimum Essential Medium containing 10% FBS, 2 mmol/L L-glutamine, 1.5 g/L sodium bicarbonate, and P/S. Thyroid cancer cell lines (XTC-1, FTC133, FTC236, FTC238, and TPC-1) were maintained in DMEM containing 10% FBS, thyroid-stimulating hormone (TSH; bovine; 0.01 units/mL), insulin (human; 10 μg/mL), P/S, and amphotericin B (250 ng/mL). Transient transfection in HEK293 and U2OS cells using lipofectamine were performed as described (13). Short hairpin RNA (shRNA)–mediated knockdown using lentiviral vectors was performed as described. The siRNAs directed against human LIMD2 were obtained from Open Biosystems. The infected cells were grown in puromycin selection medium for 1 week, and the pool was used for Western blotting and for functional assays. Whole cell extracts were prepared for Western blotting.

Cell migration assays were performed using standard transwell chambers (8-μm pore size; Corning). The cells were serum starved for 4 hours before seeding for the transwell assay. After detachment with trypsin, the cells were washed gently with PBS and resuspended in serum-free medium. A 250 μL cell suspension (2×10^5 cells/mL) was added to the upper chamber. Triplicate wells were done for each assay. The complete medium was added to the bottom chamber. After overnight incubation, cells on the filters were fixed (glutaraldehyde) and stained with 0.5% toluidine blue. The number of migrated cells on each filter was quantified manually by counting all cells in 4 separate fields under ×10 magnification using a microscope. The mean number of cells in triplicate filters was calculated and is reported in each bar of the graph, ±SD for 1 experiment. Each experiment was repeated at least twice. *T* tests were performed and *P* values were obtained to compare the different cell lines used.

For immunofluorescence, TPC1 cells grown on sterile coverslips were transfected with pcDNA3.1 myc-LIMD2; 24 hours later the media was aspirated off and the cells were rinsed 2 times in 1× PBS. The cells were fixed with 4% formaldehyde for 15 minutes at room temperature (RT). Then the cells were rinsed 3 times with 1× PBS for 5 minutes each. The cells were permeated with 0.2% Triton X-100 for 5 minutes. The cells were rinsed 3 times with 1× PBS for 5 minutes each. The cells were blocked for 1 hour in 5% goat serum. The cells were then incubated overnight at 4°C with the following dilutions of primary antibodies (made in 5% goat serum): 1:1,000 LIMD2 (mAb), 1:200 Myc-tag antibody (Cell Signaling Inc.). The cells were rinsed 3 times with 1× PBS. Then the cells were incubated with 1:1,000 dilution of secondary anti-mouse Alexa Fluor 594 and anti-rabbit Alexa Fluor 647 antibodies (diluted in 5% goat serum) for 30 minutes at RT (in the dark). The cells were rinsed 3 times with 1× PBS for 5 minutes each. The cells were counterstained with DAPI for 2 minutes and rinsed 3 times in 1× PBS for 5 minutes each. The coverslip was mounted on a glass slide with a drop of Prolong Gold antifade reagent and observed by confocal microscopy.

The soft agar assay was performed as described by the Wallert and Provost Lab. The base agar is composed of 0.5% agar, 1× DMEM, and 10% FBS. The top agar is formed by mixing equal volume of 2× DMEM, 20% FBS, 0.7% agar, and the cell suspension to give 1× DMEM, 10% FBS, 0.35% agar, and 2×10^4 or 4×10^4 cells/well (6-well plate). The plates were incubated at 37°C in humidified incubator for 2 to 3 weeks. The cells were fed 2 times per week with cell culture media. The plates were stained with 0.5 mL of 0.005% Crystal Violet for 1 hour and the colonies in each well were quantified by counting 4 fields under 2× magnification using a microscope. Each condition was done in triplicate and the average number of colonies per well is represented. Each experiment was repeated 3 times.

The GST association assays using *in vitro* transcribed and translated (IVT) protein was performed as described (12). Full-length, ^{35}S-labeled ILK protein was used for binding to the

bacterially produced GST fusion protein, and the association and washing buffer was 1× radioimmunoprecipitation assay buffer. Standard IHC techniques were used to stain paraffin-embedded tissues, except that sections were subjected to antigen retrieval by pressure boiling in 10 mmol/L sodium citrate, 0.05% Tween 20, pH 6.0, for 3 minutes at 120°C. The sections were blocked with 1% normal horse serum for 30 minutes, incubated with αLIMD2 (mAb) overnight at 4°C (1:100 dilution of a 1 mg/mL IgG stock) and followed by 30 minutes incubation with horse anti-mouse biotinylated immunoglobulin (1:200). Subsequent steps followed the Vectastain Elite ABC Kit. The staining intensity and percentage of positive cells were scored using an ordinal scale blinded to the clinical and pathologic data. Staining intensity score was 0 for absent, 1 for faint, 2 for moderate, and 3 for strong. The percentage of positive cells score was 0 for 0%, 1 for less than 33%, 2 for ≥33% and ≤66%, and 3 for ≥66%.

Nuclear magnetic resonance spectroscopy

The hLIMD2 cDNA was cloned into pMAL-c5TEV (modified version of pMAL-c5X from NEB) and expressed in *Escherichia coli* BL21 (DE3). The cells were grown at 37°C until OD_{600} 0.7 and then induced with 0.5 mmol/L of isopropylthio-β-galactoside at 25°C for 20 hours. The harvested cells were suspended in lysis buffer [20 mmol/L Tris pH 7.4, 200 mmol/L NaCl, 100 μmol/L $ZnSO_4$, 0.2% IGEPAL, complete protease inhibitor (1 tablet per liter of the culture), 0.1 mg/mL lysozyme, 0.5 mmol/L phenylmethylsulfonyl fluoride, and 10 mmol/L β-mercaptoethanol]. The cells were disrupted by sonication, lysate was cleared by centrifugation at 18,000 rpm, and applied to amylose resin (NEB) equilibrated with the equilibrate buffer (20 mmol/L Tris pH 7.4, 200 mmol/L NaCl, 100 μmol/L $ZnSO_4$, and 10 mmol/L β-mercaptoethanol) at 4°C for 1 hour. The resin was washed 5 times with 20 mL of equilibration buffer. MBP-fusion protein was eluted with 10 mmol/L maltose and cleaved by 100 μg of TEV protease per liter of the culture at 16°C overnight. The cleaved protein was subsequently purified by MonoS 4.6/100 PE and HiLoad 16/60 Superdex 75 pg [pre-equilibrated with 50 mmol/L phosphate pH 7.2, 100 mmol/L NaCl, and 100 μmol/L TCEP (Amersham Biosciences)]. Isotopically enriched LIMD2 was prepared from cells grown on minimal media (supplemented with 100 μmol/L of $ZnSO_4$) containing 1 g/L of [^{15}N] ammonium chloride with 2 g/L of either unlabeled glucose or [^{13}C6] glucose (Cambridge Isotopes Laboratory). The ankyrin repeat domain (ARD; 1-174aa), kinase domains (182-452aa) of ILK were cloned into pET 32a (Invitrogen) and pGEX-6P1 (GE), respectively. Growth and induction were carried out as described above, except that the expression temperature was reduced to 15°C. The lysis buffer contained 50 mmol/L phosphate pH 7.2, 200 mmol/L NaCl, 10 mmol/L β-mercaptoethanol, and 10% glycerol. The protein used for nuclear magnetic resonance (NMR) was characterized by analytical ultracentrifugation and circular dichroism to determine its suitability for NMR.

Protein concentrations of LIMD2 (300–400 μmol/L) were used for NMR experiments for resonance assignment and structure calculation. The NMR spectra for assignment were collected in 90% H_2O/10% D_2O containing 50 mmol/L sodium phosphate, pH 7.2, 100 mmol/L NaCl, and 1 mmol/L TCEP at 600 MHz on a Varian INOVA spectrometer equipped with an HCN cold probe at 25°C. Backbone assignment was carried out using HNCACB and CBCA(CO)NH, whereas side chain assignments were determined using HCCH-TOCSY, CC(CO)NH, HBHA(CO)NH, as well as 2D (HB)CB(CGCD) HD and (HB)CB(CGCDCE)HE. The protonation state of the histidine residues was examined using a modified HMQC experiment (14). Distance restraints were derived from ^{15}N-edited NOESY in 90% H_2O/10% D_2O, and ^{13}C-edited NOESY and 2D NOESY in D_2O. The spectra were processed by NMRPipe (15) and analyzed by SPARKY (16). The molecular dynamics based model of LIMD2–ILK–Parvin was obtained by AutoDOCK (17) analysis of fragmented LIMD2 structure against that of the pdb structure 3kmw (ILK/α-Parvin), rethreading the LIMD2 structure through the top 2 binding sites of the Autodock analysis and energy minimizations performed in water at a pH of 7.4.

The ILK kinase assay was performed as described previously (18) with modifications. ILK and LIMD2 were mixed and preincubated for 30 minutes at room temperature before addition of the substrate, myosin phosphatase target subunit 1 (MYPT1; Millipore). The amounts of ILK and MYPT1 were kept constant at 33.3 nmol/L (50 ng/20 μL reaction) and 417 nmol/L (500 ng/20 μL reaction), respectively, and increasing amounts of LIMD2 were added as indicated. Kinase reactions were initiated by the addition of ATP to a final concentration of 500 nmol/L. Reactions were carried out in 1× kinase reaction buffer (25 mmol/L Tris-HCl pH 7.5, 5 mmol/L β-glycerophosphate, 2 mmol/L dithiothreitol, 0.1 mmol/L Na_3VO_4, 10 mmol/L $MgCl_2$) for 30 minutes at 30°C, terminated with the addition of sample buffer and proteins separated by SDS-PAGE. ILK and P-MYPT were assessed by Western blot analysis. MYPT concentration was assessed by Coomassie blue staining.

Results

Characterization of LIMD2 protein and production of αLIMD2 antibodies

LIMD2 belongs to the LIM-only family of LIM proteins, which contain between 1 and 5 LIM domains (Fig. 1A). A distance-based phylogenetic analysis of amino acid sequences among these LIM-only protein family members was performed using MEGA program (Fig. 1B). The LIMD2, CRP1, and LMO1 segregate from FHL1 and PINCH1, indicating that they are more similar to one another than they are to FHL1 and PINCH, which are more closely related, implying that there are shared ancestral relationships between the individual LIM domains (Fig. 1C; ref. 19). To define the targets and functions of LIMD2, rabbit polyclonal (rAb) and mouse monoclonal antibodies (mAb; Supplementary Fig. S1A) were produced and their specificity was determined. Both αLIMD2 (rAb) and αLIMD2 (mAb) detect the exogenous and endogenous LIMD2 protein in HEK293 and U2OS cell lines (Fig. 1D) but do not interact with AJUBA and LIMD1 (Supplementary Fig. S1B–S1D). By transient transfection, the Myc-LIMD2 protein [detected by both Myc (rAb) and LIMD2 (mAb)] is found predominantly in cytoplasm in a smooth granular distribution (Fig. 1E), a result confirmed by subcellular fractionation (data not shown). However,

LIMD2 is concentrated in membrane ruffles and in streaks reminiscent of focal adhesion plaques (Fig. 1E). Interestingly, the protein is not excluded from the nucleus, consistent with many other reports of nucleocytoplasmic shuttling by LIM proteins. LIMD2 is robustly detected in formalin fixed, paraffin-embedded tissue specimens as a cytoplasmic protein, and no signal was seen in cells that lack LIMD2 (Supplementary Fig. S1E). Thus, these antibodies are specific and will be useful in examining the role of LIMD2 in cell motility and tumor metastasis.

LIMD2 transcript and protein levels correlate with malignant potential in breast, bladder, melanoma, and thyroid cancer cell lines and tumors

LIMD2 levels are well correlated with the higher degree of invasiveness among immortalized mammary epithelial cells (MCF-10A), nonmetastatic breast cancer cells (MCF-7), and breast cancer cells (MDA-MB-231 and MDA-MB-435; Fig. 2A). The LIMD2 RNA levels measured by quantitative PCR (qPCR) largely correlate with the protein levels (Supplementary Fig. S2A). MCF-7 cells expressing exogenous LIMD2 were obtained using lentivirus infection: this LIMD2 protein was readily detected by Western blot analysis (Fig. 2B, top). LIMD2 expression stimulated the migration of the MCF-7 cells compared with the control cell line (pLU vector control; Fig. 2B, bottom). Conversely, the knockdown of LIMD2 in MDA-MB-231 cells strongly impaired migration in these highly aggressive breast cancer cells (Fig. 2C). Restoration of wtLIMD2 expression in Si-5 knockdown MDA-MB-231 cells (using a LIMD2 cDNA with silent mutations in the coding sequence that render it resistant to the LIMD2-Si RNA) resulted in a regaining of the same migration ability as parental cells (Supplementary Fig. S3A and S3B), suggesting that the effects are because of LIMD2 expression and not to off-target effects. Depletion of LIMD2 strongly diminished colony formation in soft agar for all cell lines, but was efficiently restored by complementation by wtLIMD2 expression (Fig. 2E). Major differences in cell morphology between parental, control shRNA, LIMD2-Si2, and LIMD2-Si5

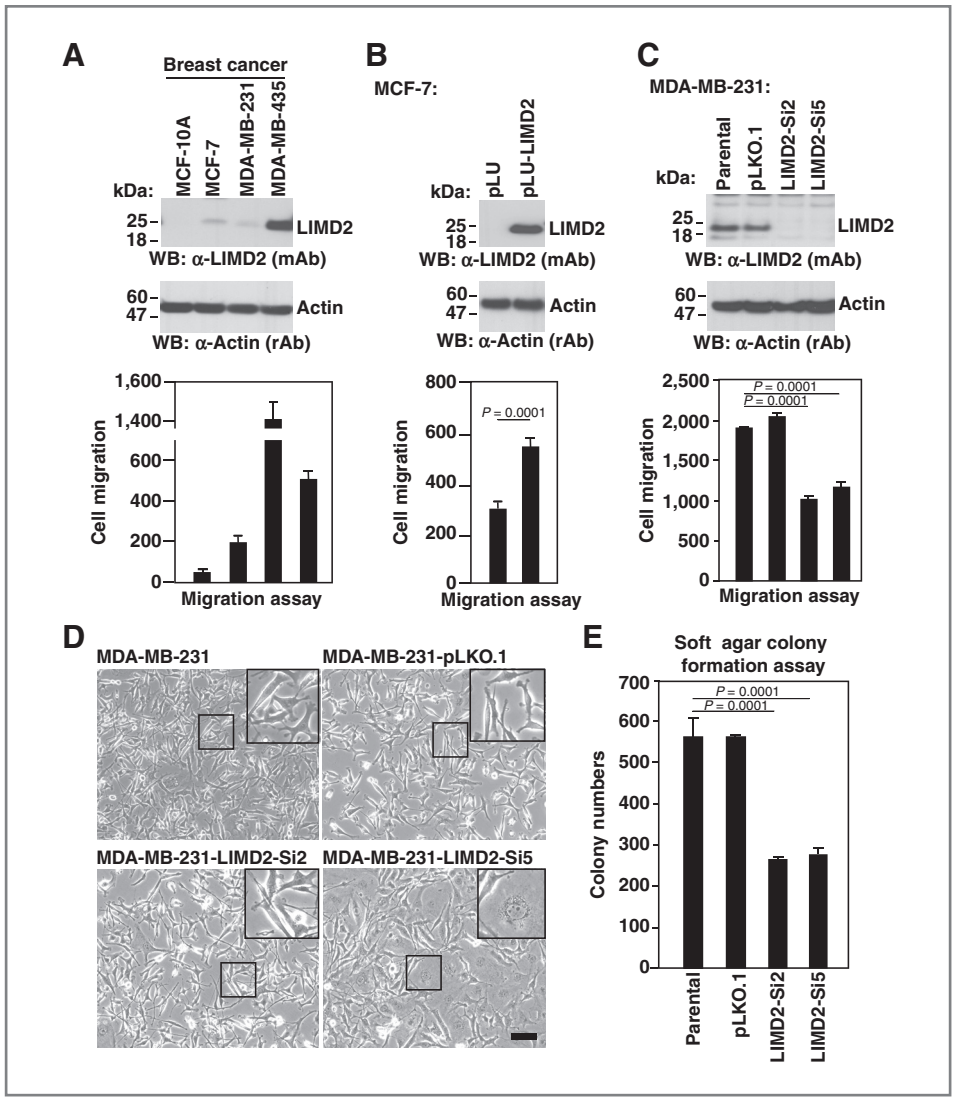

Figure 2. A, breast cancer cell lines express endogenous LIMD2. B, ectopic expression of LIMD2 in MCF-7 cells. C, siRNA-mediated LIMD2 knockdown in MDA-MB-231 cells (the Si2 and Si5 are 2 independent cell clones). Migration potentials of each cell line. Data are presented as mean ± SD. *T* test was performed and *P* value is indicated. D, morphology of the cells. Bar, 100 μm. E, soft agar colony formation assays of each cell line.

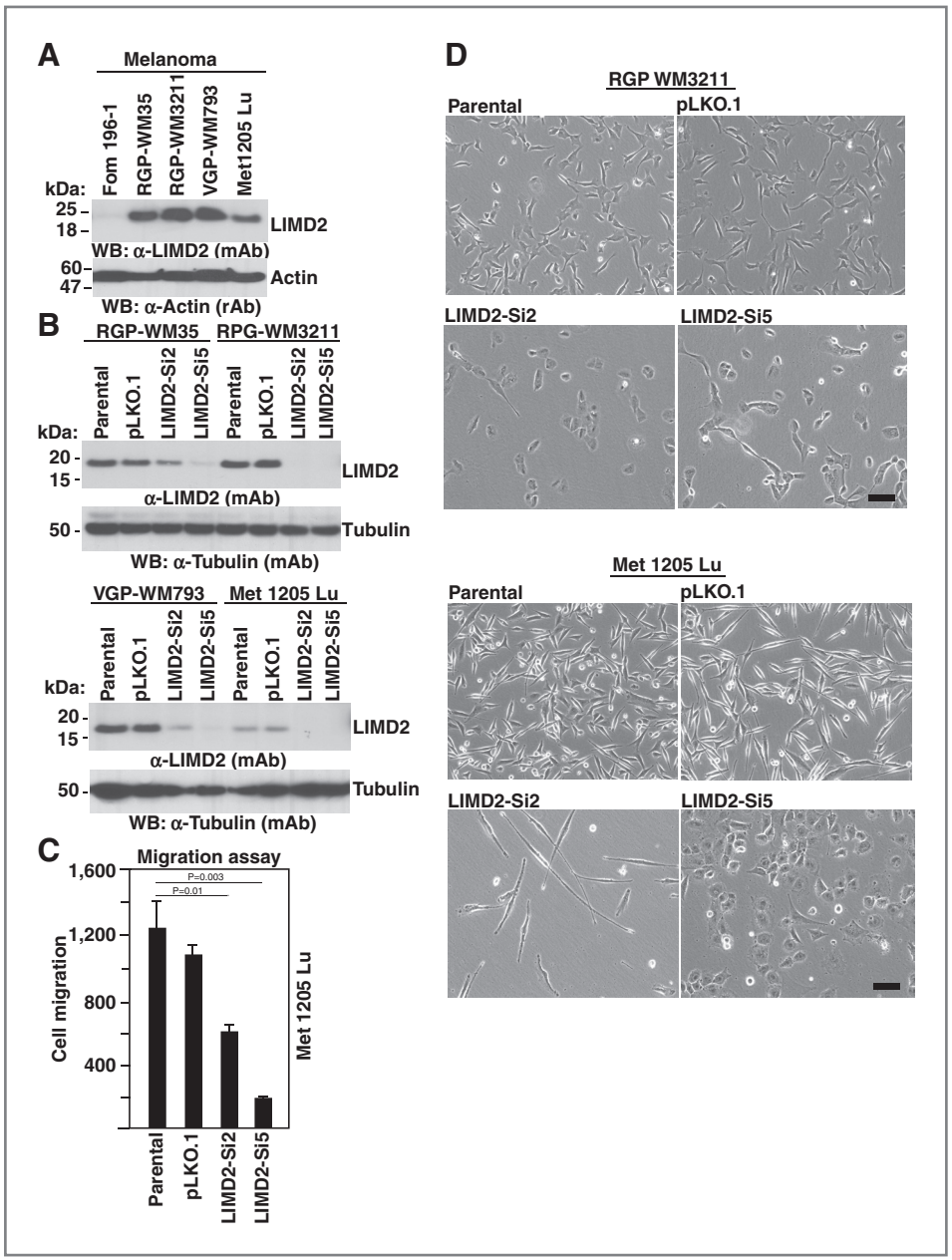

Figure 3. A, LIMD2 expression in a spectrum of melanoma cancer cells by Western blot analysis. B, knockdown of LIMD2. C, migration potentials of melanoma cells. The statistics were performed and *P* value is indicated. D, morphologic changes in melanoma cancer cells. Bars,100 μm.

cells were also observed: the vector cells exhibited a spindle-like fibroblastic morphology, whereas the LIMD2-Si2 and LIMD2-Si5 cells showed an enlarged and flattened morphology (Fig. 2D). LIMD2 knockdown cells are enlarged and flattened as indicated with white arrows (Supplementary Figs. S3C and S4). LIMD2 overexpression restored the parental cells phenotype, showing a spindle-like fibroblastic morphology as indicated with black arrows (Supplementary Figs. S3C and S4). Thus, LIMD2 controls multiple hallmarks of oncogenic growth: anchorage-independent growth, migration, and altered cell morphology.

LIMD2 was analyzed in a series of melanoma [Fom196-1 cells (normal melanocytes), RGP-WM35, RGP-WM3211 cells (radial growth-phase melanoma), VGP-WM793 cells (vertical growth-phase melanoma), and Met1205 Lu cells (metastatic melanoma); Supplementary Fig. S2B]. LIMD2 protein was undetectable in normal melanocyte Fom196-1 cells, whereas the expression significantly increased in the RGP, VGP, and metastatic melanoma cells in a somewhat graded fashion (Fig. 3A). Following introduction of knockdown siRNA lentiviral vectors for LIMD2, protein levels and the migration rates were significantly reduced in all melanoma cell lines (Fig. 3B and C), whereas the cell morphology was widely variable (Fig. 3D). The RGP WM3211 cells with LIMD2 knockdown were flat compared with the spindle-shaped parental cells (Fig. 3D, RGP WM3211 panels). The metastatic melanoma cells (Met1205 Lu) morphology was changed dramatically showing a highly elongated shape reminiscent

of neuronal cells, whereas the LIMD2-Si5 cells became flat and larger. Loss of LIMD2 significantly reduced Met1205 Lu cell migration (Fig. 3C).

In the series of bladder cancer cells [RT4, HT1376, UM-UC3, ScaBer, J82, TCCSUP, and T24 (listed from least to most aggressive)], LIMD2 levels coordinately increased with the known malignant characteristics of these cells (Supplementary Figs. S2C and S6A), with the more aggressive cell lines showing higher migration rates (Supplementary Fig. S6B and S6C). Ectopic LIMD2 expression in the more benign bladder cancer cell line RT4, which showed undetectable endogenous protein expression levels significantly increased migration potential (Supplementary Fig. S6D), indicating that the gain-of-function LIMD2 promotes bladder cancer cell migration. LIMD2 knockdown in T24 cells efficiently reduced LIMD2 protein levels to almost undetectable levels (Supplementary Fig. S6E), and these cells showed significant decreased numbers of colony formation in soft agar compared with the parental and nontargeted siRNA cells.

Because LIMD2 was identified from a papillary thyroid cancer metastasis, we evaluated LIMD2 expression in thyroid carcinoma cells lines: XTC-1, FTC133, FTC236, FTC238, and TPC-1. LIMD2 levels are elevated in the thyroid papillary cell line TPC-1, which is the most aggressive thyroid cancer cell line (Fig. 4A and Supplementary Fig. S2D) and shows the highest migration rate (Fig. 4B and C). We also evaluated LIMD2 expression in primary thyroid cancer specimens using qPCR. A total of 252 papillary thyroid cancer samples were evaluated for LIMD2 mRNA level by qPCR. We found that the LIMD2 mRNA level was significantly higher in papillary thyroid cancer with extrathyroidal invasion ($P = 0.001$; Fig. 4D and E). This highly significant finding further underscores the role of LIMD2 in invasion and metastasis processes.

Structural studies into LIMD2 give insights into molecular function

By analogy to the known functions of other LIM domains, LIMD2 likely binds to cellular proteins. We sought a structural biology solution to identify its targets by taking advantage of the fact that LIMD2 is well folded, globular, monomeric, and monodisperse in solution at high concentrations, making it suitable for NMR. Complete backbone assignments of the protein were determined (Fig. 5A), however analysis of NOESY spectra (^{13}C- and ^{15}N-edited NOESY-HSQC as well as 2D-NOESY) indicated a lack of signals for the first 35 and last 23 amino acids of the protein consistent with these being unstructured. Consequently, only residues 33 to 104 were used for structure determination. This protein is composed of two treble clef fingers (19, 20) folded around two zinc atoms (Fig. 5B and C). ICP-EOS studies also indicate that there are 2.4 zinc molecules per LIMD2 domain. The zinc-ligated residues are Cys40, Cys43, His61, and Cys64 (CCHC) for site I, and Cys67, Cys70, Cys88, and His91 (CCCH) for site II. The first finger is composed of 2 antiparallel β-sheets with zinc-coordinated residues located on the loop connecting two strands in each antiparallel β-sheet. The second finger is composed of a long loop followed by an antiparallel β-sheet and ends in an approximately 2-turn helix. The zinc-ligating residues in the second site reside on the long loop and the beginning of a 2-turn helix (Fig. 5B). The $N^{\delta 1}$ atoms of the 2 histidines (His61 and His91) coordinate the zinc. The zinc-ligating residues comprise one of the key structural elements of LIMD2. In addition, structural analysis indicates that there is a small hydrophobic core likely involved in the stabilization of the protein, including the aromatic side chains of residues Phe60, Phe65, and Phe86, which form a hydrophobic cluster between the 2 zinc fingers and are buried (Fig. 5E).

We assessed the structural similarity between LIMD2 and other reported LIM structures present in the PDB. The superposition of LIMD2 onto the 3-dimensional crystal structures of PINCH1-LIM1 (PDB code 3F6Q; ref. 11), and PINCH2-LIM1 (PDB code 3IXE; refs. 21 and 22) yielded an RMSD of 2.9 and 2.8 Å, respectively, for the backbone of the LIM consensus sequence (residues 39–96 in LIMD2). The secondary structure elements are well conserved, although the antiparallel β-sheet of the second finger of the LIMD2 adopts a flat conformation, and the c′ strand rotates around 90° along the axis of the sheet compared with other structures (Fig. 5C and D). The electrostatic surface of LIMD2 is very positive, as expected, given its p*I* of 8.7; however, the charge distribution is quite asymmetric, with both positive and hydrophobic patches (Fig. 5F). Comparison of LIMD2 to PINCH-LIM1 reveals that although the latter has a net negative charge (p*I* of 4.9), once again the charge distribution is asymmetric and contains hydrophobic patches (Supplementary Fig. S5). Similar to LIMD2, PINCH-LIM4 is a positively charged domain (p*I* 8.7); however, in contrast to LIMD2, the C-terminal helix is both negatively and positively charged on the corresponding faces based on alignment with the LIMD2 domain, respectively. The asymmetric charge distribution and positions of hydrophobic patches may supply various interfaces to accommodate specific binding proteins.

The LIMD2 structure was highly similar to the first LIM domain (LIM1) of PINCH1. PINCH1-LIM1 binds to the ARD of the ILK (11). ILK is a 452 amino acid, multidomain protein consisting of 5 N-terminal ankyrin repeats, and a C-terminal kinase domain (Figs. 6A and 8B). ILK is a core component of the IPP complex, which contains PINCH proteins and Parvin (Fig. 8B). The IPP complex binds to the C-terminal tails of β1-integrins, thus serving to communicate cell–cell and cell–matrix adhesion signals to the nucleus (23). To determine if PINCH1-LIM1 and LIMD2 bound to ILK, we used GST-LIMD2, GST-PINCH1-LIM1 and ^{35}S-labeled ILK protein produced by IVT. Indeed, the GST-LIMD2 protein robustly binds to ILK under the same conditions as PINCH-LIM1; neither protein binds to GST alone (Fig. 6B). Identical results were observed for ILK and LIMD2 interaction using fully recombinant proteins, isolated from *E. coli* (Fig. 6C). These studies strongly suggest that ILK is a candidate target protein for binding by LIMD2. GST-PINCH1-LIM1 bound to the ARD robustly, as has been shown by other investigators (11), but the LIMD2 protein did not bind to the ILK-ARD (data not shown). Moreover, LIMD2 did not compete for PINCH1-LIM1 for binding to ILK (data not shown).

To determine if endogenous cellular ILK is required for stimulation of migration by LIMD2, we used a matched set of

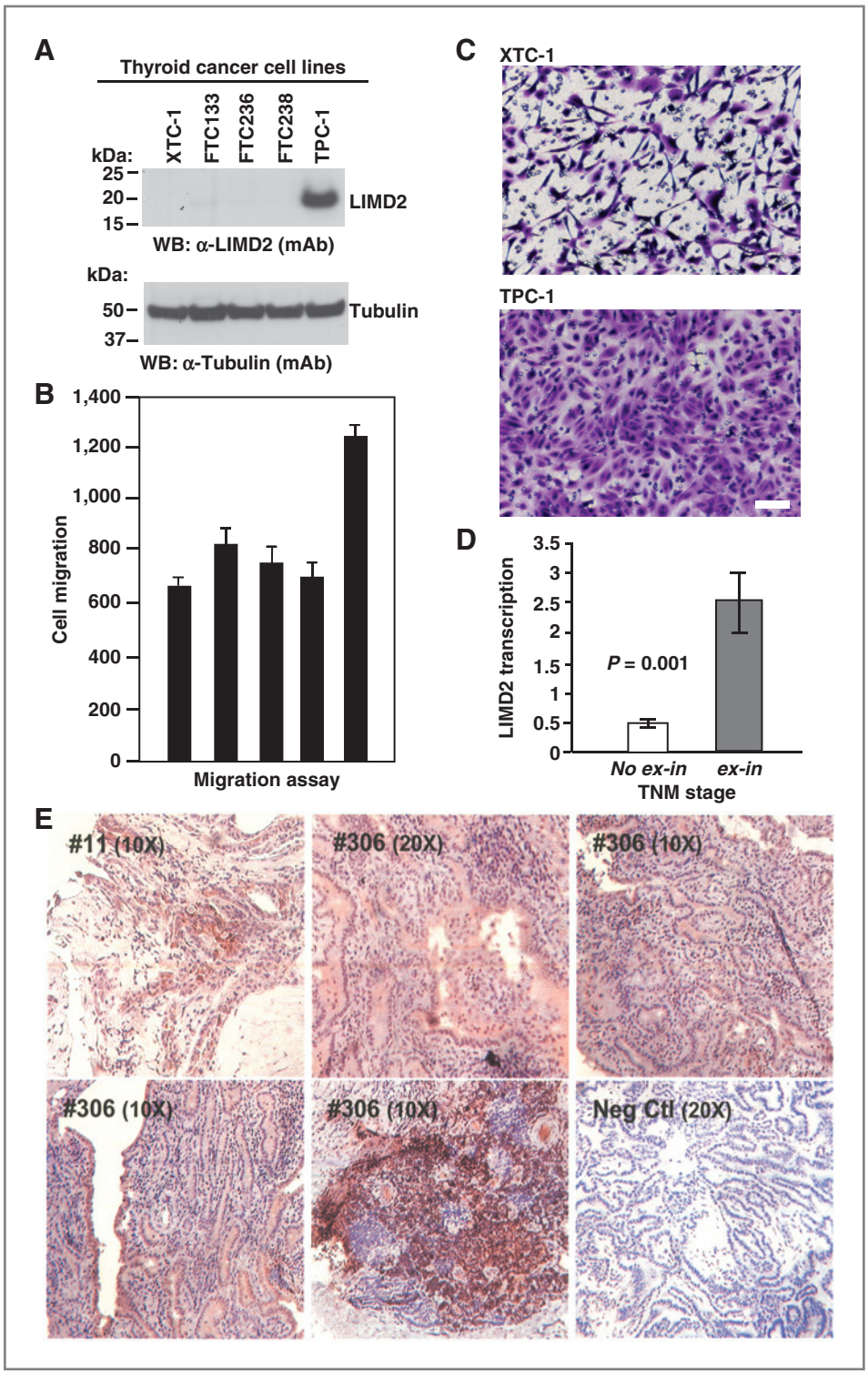

Figure 4. A, LIMD2 is highly expressed in thyroid cancer cells. B, migration potentials of the thyroid cancer cells. Data are presented as mean ± SD. C, images of thyroid carcinoma cell line XTC-1 and thyroid papillary carcinoma cell line TPC-1 taken from the migration assay. Bar, 100 μm. D, LIMD2 mRNA is higher in thyroid cancer samples with extrathyroidal invasion ($n = 49$). Total RNA was extracted from 252 papillary thyroid cancer samples using the TRIzol reagent (Invitrogen). TaqMan gene expression assay (Applied Biosystems) was used to quantitate LIMD2 mRNA expression levels. E, representative immunostaining results in papillary thyroid cancer samples using the αLIMD2 (mAb).

fibroblasts that are either wild type for ILK (+/+) or have been rendered ILK null (−/−) via CRE-mediated excision of floxed ILK alleles (24). Both of these cell types grow robustly in culture, show similar morphologies, similar doubling times, and are equally infectable by lentiviruses (Fig. 6D). The −/− cells lack ILK by Western blot analysis as expected (Fig. 6E). LIMD2 is robustly, and equally expressed in both cell types (Fig. 6E) when infected with lentiviruses. LIMD2 expression in $ILK^{+/+}$ cells stimulates cell migration almost 4-fold. However, there is no stimulation of migration by LIMD2 in cells lacking

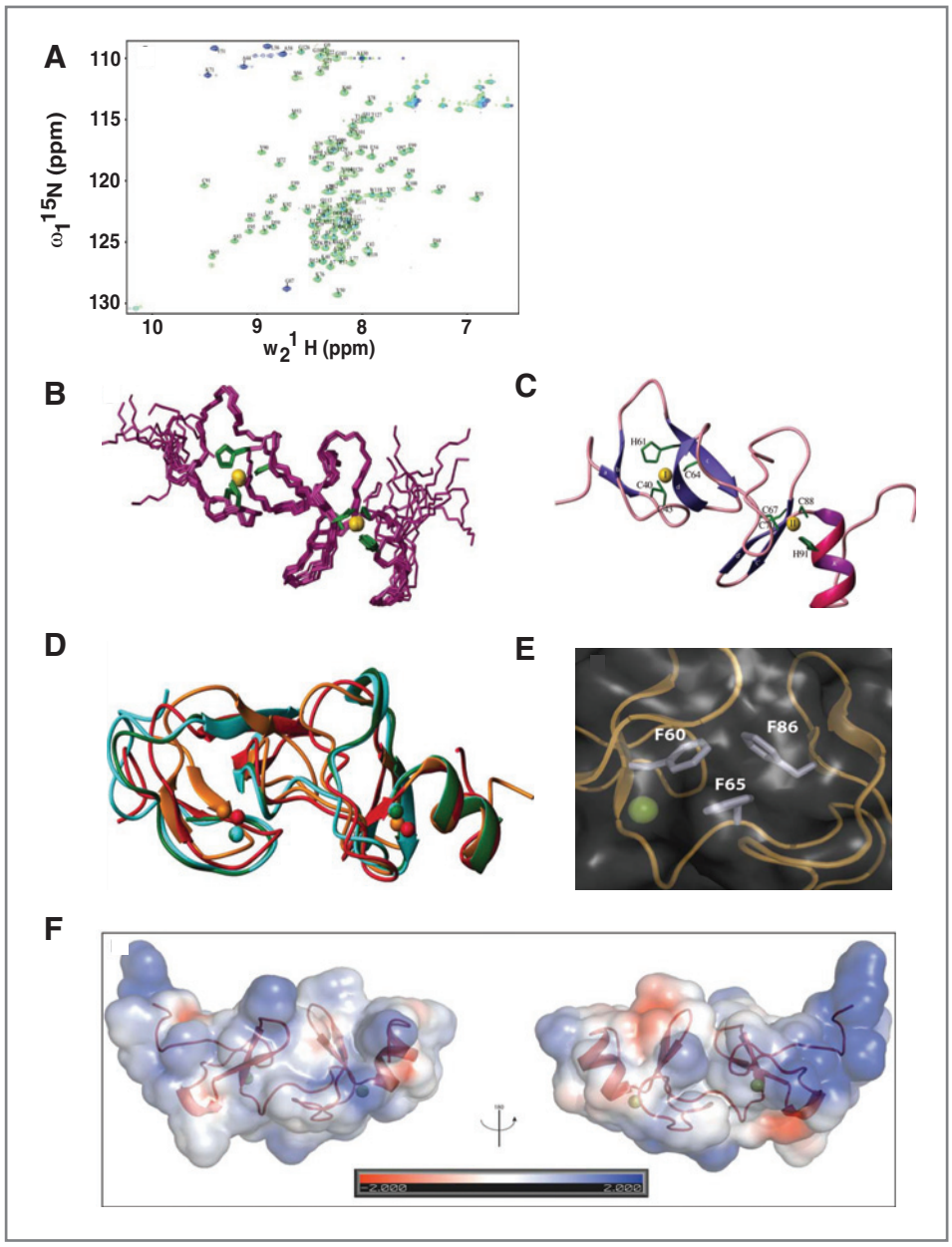

Figure 5. A, ^{1}H-^{15}N HSQC spectrum of LIMD2 with backbone assignment: BMRB accession number 18778. Purple represents aliased signals. B, superimposition of the 10 best structures calculated by ARIA (PDB accession code 2LZU) with an RMSD of 0.32 A for LIMD2 residues 39-96aa over backbone atoms. See Supplementary Table S1 for structure statistics. Zinc is colored in gold and side chains of the zinc-coordinated residues are shown in green. C, ribbon representation of LIMD2. D, alignment of LIMD2 (red) with PINCH1-LIM1 (cyan PDB code 3F6Q), PINCH2-LIM1 (green PDB code 3IXE), and PINCH1-LIM4 (orange PDB code 1NYP). E, side chains of F60, F65, and F86 from a LIMD2 hydrophobic core. F, electrostatic surface of LIMD2. The asymmetric charge distribution in LIMD2 is shown.

the −/− cells that lack ILK (Fig. 6F). Thus, ILK is likely a key effector for LIMD2 stimulated cell migration.

Because LIMD2 bound tightly to the region of ILK containing the kinase domain, we determined if LIMD2 could alter the kinase activity of ILK toward a model peptide substrate *in vitro*. Although there is considerable controversy in the literature about whether ILK is a kinase or a pseudo-kinase, the observation that baculovirus-expressed and purified protein contains this activity warranted a look at this question: we engaged the investigators who have championed this claim to perform these experiments. Baculovirus-produced full-length ILK (Fig. 7A) showed considerable kinase activity toward the MYPT peptide as has previously been shown (18). The addition of LIMD2 to the *in vitro* assay resulted in a dose-dependent increase in kinase activity, showing almost a 3-fold stimulation at the highest levels of LIMD2 added (Fig. 7B–D). This apparent activation of ILK by LIMD2 was abolished when the ILK small-molecule inhibitor kinase inhibitor QLT0267 was added to the reaction (Fig. 7E). Thus, binding of ILK by LIMD2 *in vitro* results in stimulation of the kinase activity exhibited by the recombinant ILK under these conditions. Together, these data strongly suggest that LIMD2 is a new positive regulator of ILK activity.

Discussion

Characterization of the novel, metastasis-associated LIMD2 protein described herein supports the following conclusions: (i) LIMD2 is a small cytoplasmic protein whose endogenous

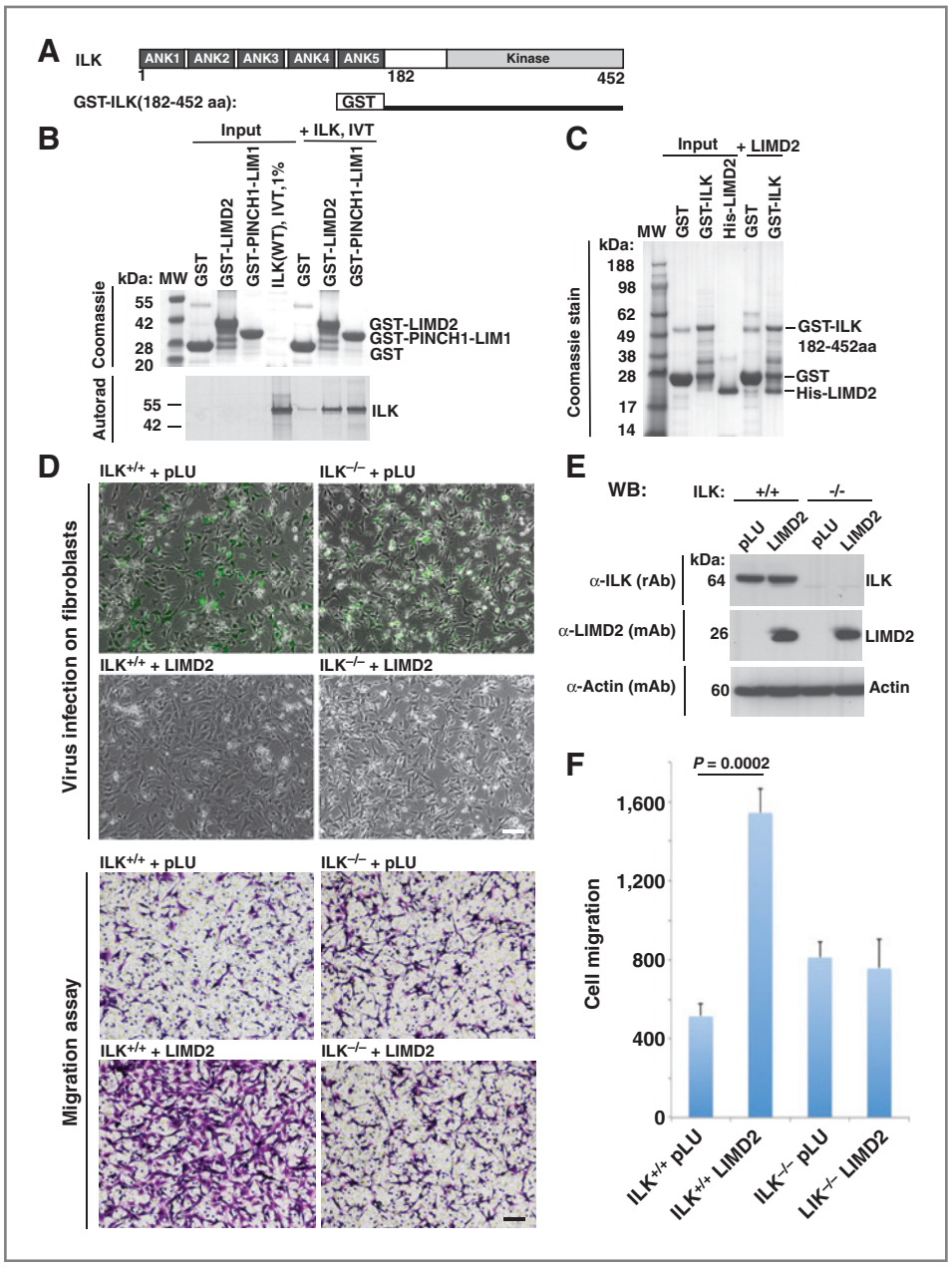

Figure 6. LIMD2 and PINCH1 directly associate with ILK, LIMD2 affects cell migration through ILK. A, domain architecture of ILK. The N-terminus of ILK contains ankyrin repeat domain (ARD) and the C-terminus of ILK contains a kinase domain. B, ILK protein binds to GST-LIMD2 and GST-PINCH1-LIM1 but not GST alone. C, LIMD2 protein binds to GST-ILK (182-452aa). D, LIMD2 stimulates cell migration in ILK$^{+/+}$ fibroblast but not in ILK$^{-/-}$ fibroblast. Lentivirus pLU vector (GFP) and pLU-LIMD2 were infected on ILK$^{+/+}$ and ILK$^{-/-}$ fibroblast cells. Migration assay was performed on pLU and LIMD2 virus infected ILK$^{+/+}$ and ILK$^{-/-}$ fibroblast cells. Images were taken from the migrated cells. Bar, 100 μm. E, endogenous ILK protein were detected in ILK$^{+/+}$ fibroblast cells but not in ILK$^{-/-}$ cells by Western blot analysis using αILK (rAb). Exogenous LIMD2 proteins were detected in LIMD2 virus–infected fibroblast cells by Western blot analysis using αLIMD2 (mAb). F, ectopic LIMD2 expression promotes ILK$^{+/+}$ fibroblast cells migration. Three independent experiments were performed. Data, mean ± SD. *P* value is indicated.

expression levels correlate with the known malignant properties of cells derived from breast, melanoma, bladder, and thyroid cancers; (ii) ectopic expression or knockdown of LIMD2 stimulates or abrogates cell invasiveness, motility, and cell morphology consistent with tumor grade; (iii) primary human thyroid cancers that express LIMD2 show much greater extrathymic invasion; (iv) the LIMD2 structure is homologous to the PINCH1-LIM1 domain that binds to ILK; (v) both PINCH1-LIM1 and LIMD2 proteins bind to ILK directly but at nonoverlapping sites; (vi) LIMD2 stimulates the kinase activity of ILK, and endogenous ILK is required for the ability of LIMD2 to stimulate cell migration. Thus, in LIMD2, we have discovered a new target and function for a LIM-only protein in tumor progression using a combination of tumor biology and structural biology experimental approaches.

The finding of a direct connection between LIMD2 and the integrin signaling apparatus is wholly consistent with the promotion of cell migration and invasion seen in LIMD2-expressing cell lines and tumors. The multiprotein IPP complex (ILK–PINCH–Parvin) as a core complex links transmembrane integrin adhesion receptors to the actin cytoskeleton and intracellular signaling cascades (11, 25). The ILK localizes to sites of integrin-mediated cell adhesion (26, 25, 27). The presence of intrinsic catalytic activity as a kinase in ILK is still controversial, but the domain clearly binds to the CH2 domain of Parvin (28). Formation of the tripartite IPP complex is

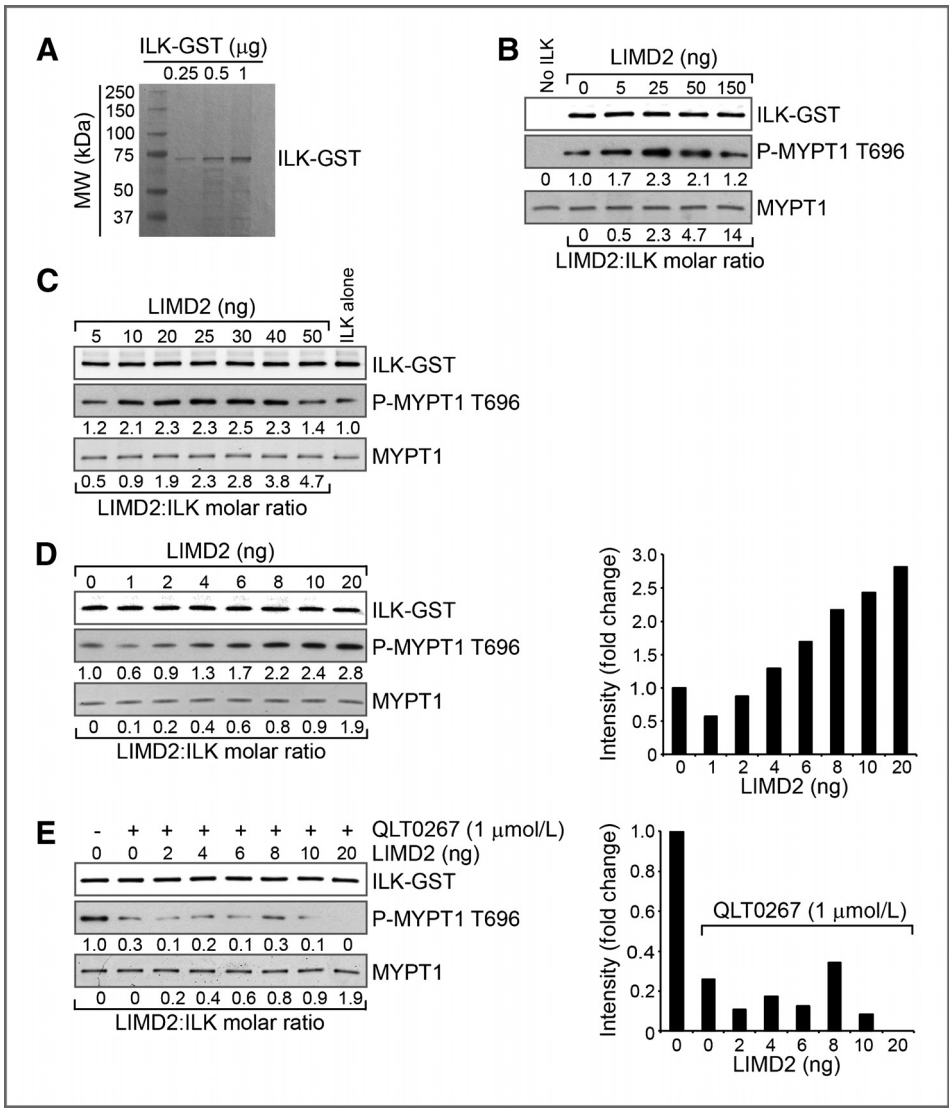

Figure 7. LIMD2 increases ILK kinase activity in a dose-dependent manner. A, coomassie blue-stained polyacrylamide gel showing highly purified ILK–GST fusion protein used in the ILK kinase assay. B, dose-dependent phosphorylation of the ILK substrate myosin phosphatase target subunit 1 (MYPT1) on Thr696 in the presence of increasing concentrations of purified LIMD2. ILK and LIMD2 were mixed and preincubated for 30 minutes at room temperature before addition of the substrate. The amounts of ILK and MYPT1 were kept constant at 50 and 500 ng/reaction, respectively, and increasing amounts of LIMD2 were added as indicated. C, concentration curve similar to that shown in B, with LIMD2 concentrations increased stepwise from 5 to 50 ng. Reactions were carried out as described for B. D, concentration curve similar to that shown in C, with LIMD2 concentrations increased stepwise from 1 to 10 and 20 ng (molar ratio of 0.1–1.9). Reactions were carried out as described for B. Data in the bar graph are normalized to the signal observed in the presence of ILK alone. E, effect of the specific small molecule inhibitor of ILK, QLT0267, on the phosphorylation of MYPT by ILK-LIMD2. ILK-LIMD2 complexes were preincubated at the indicated concentrations as described above, and QLT0267 was added at a final concentration of 1.0 μmol/L just before initiation of the kinase reaction. Data in the bar graph are normalized to the signal observed in the presence of ILK alone.

absolutely required for integrin signaling, migration, and invasion in both normal developmental and malignant programs. The PINCH1 LIM1 domain binds directly to the ARDs of ILK and stimulates integrin signaling (whereas the highly related PINCH2 protein binds to the ILK ARD and competes for PINCH1 binding; Figs. 22 and 29–31). Other proteins that bind to the core IPP complex, such as Nck, ERK1/2, GSK3β, and Rsu1, serve to remodel the IPP complex to respond to different growth factor signaling inputs emanating from the cell membrane (32) and largely control cell growth and differentiation pathways, most notably EMT. In addition to the direct effects of LIMD2 on ILK kinase function as described below, LIMD2 analogously may adapt the IPP complex to other as yet undefined inputs.

The finding that LIMD2 binds to ILK and promotes cell migration and invasion is, of course, inconsistent with it being a competitor for PINCH1 LIM1 binding, even though they share high structural similarity. That LIMD2 binds to ILK at a site distinct from PINCH1 is consistent with a model that LIMD2 may be a new positive regulator of the IPP complex (Fig. 8B). In fact, LIMD2 and PINCH1 likely bind simultaneously to ILK at widely dispersed sites. At first blush, it may seem surprising that LIMD2 and PINCH1 would target separate regions of ILK; however, given the vastly different charge distributions on the same structural scaffold (Fig. 5F), it is reasonable that these would bind different surfaces (Supplementary Fig. S5). In contrast, our data are consistent with binding of LIMD2 near the active site and thus directly influencing its enzyme activity. An *in silico* molecular dynamics–based docking analysis strongly suggests that LIMD2 binds to a site in ILK very near the potential active site but separate from the binding site for α-Parvin (Fig. 8A). It is noteworthy that when ILK is coexpressed with α-Parvin in bacteria, the complex is enzymatically dead and in fact α-Parvin is an inhibitor of kinase activity when added to *in vitro* ILK kinase assays. Thus, α-Parvin and LIMD2 may serve as negative and positive regulators (respectively) of kinase activity by binding the 2 different lobes of the kinase domain, which border the active site cleft (Fig. 8A). The model

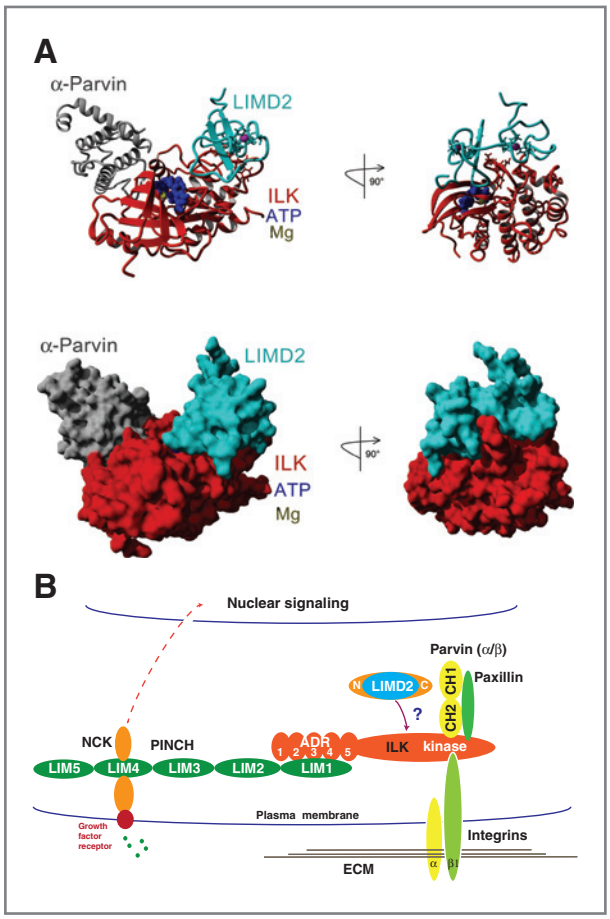

Figure 8. A, computer molecular modeling of the LIMD2–ILK–αParvin protein complex. B, a proposed working model for LIMD2 functional roles in assembly of and normal function of the ILK–PINCH–Parvin complex in the integrin-linked signaling pathway.

also suggests that LIMD2 forms a well-defined pocket/groove bordered by the 2 zinc-binding fingers, which directly interact with an active site loop in ILK: this pocket may be a promising target for small molecule–based inhibitor design.

Like LIMD2, almost all components of the IPP complex have been observed to be deregulated in human cancer (32) and are largely correlated with the later metastatic phenotype. PINCH1 expression is highly correlated with tumor grade and invasiveness in gliomas, gastric cancer, and squamous cell carcinoma, suggesting that LIMD2 overexpression may phenocopy PINCH1 overexpression (33). It may be that optimal activation of ILK requires binding of both PINCH1 and LIMD2. The negative regulator of IPP, PINCH2, is often silenced by methylation in tumors. Moreover, the PINCH1 LIM domain 5 binds directly to an altered form of the Ras suppressor, Rsu1, thereby directly participating in Ras-mediated transformation (33). Thus, IPP deregulation is a common denominator among tumors arising from different lineages, which supports the observations that LIMD2 expression occurs in multiple tumor types.

Does LIMD2 function as a classic "adaptor" protein as has been demonstrated in other LIM-only proteins? One of the best examples of this function is the LIM-only protein LMO2, which uses its 2 LIM domains to bind the GATA1 and TAL1 transcription factors in the nucleus (34, 35). In this model, each LIM domain in LMO2 is specific for each partner. Similar specificity for protein-protein interactions are demonstrated by the individual LIM domains in Paxillin, TRIP6, Ajuba, and of course PINCH1, whose 5 LIM domains are likely specialized for different partners. Thus, LIM domain proteins with 2 or more LIM domains function as true scaffolds to combinatorially assemble multiprotein complexes (10). The case of LIMD2 is problematic to propose as a classic scaffold in that it encodes only a single LIM domain. However, the unstructured NH2- and COOH-terminal protein sequence in LIMD2 may provide this function. The pronounced charge differences observed on each "face" of the LIMD2 LIM domain may indicate that the LIM domain itself can provide separate surfaces for protein–protein interactions. It is formally possible that LIMD2 functions as a positive regulator of the IPP complex by simply using its single LIM domain to compete for a negative regulator of ILK. LIMD2 may also directly modulate the activity of the kinase domain of ILK. It is noteworthy that LIMD2 has not been observed in multiple proteomics–based approaches to the composition of the ILK complex and the proteins, which inhabit focal adhesion plaques (36). This is surprising as our preliminary analyses of LIMD2 distribution in cells by confocal microscopy show it to partially overlap with focal adhesion streaks, and the fact that LIMD2 seems to bind ILK directly and with considerable affinity using recombinant proteins. An ongoing dynamic approach using videomicroscopy of genetically tagged fluorescent proteins should address this.

Does LIMD2 carry out multiple independent functions in different subcellular compartments? Many multi-LIM domain proteins shuttle between the cytoplasm and nucleus. These studies were initiated by the pioneering work of Beckerle showing that the LIM domain protein paxillin rapidly shuttled between the cytoplasm and the nucleus and had key regulatory roles in both compartments (19, 37). LIMD2 may function like its closest evolutionary homologue, CRP1. CRP1 is a 2 LIM protein that shuttles between the cytoplasm and nucleus. In the nucleus it binds a complex of the SRF and GATA transcription factors, whereas in the cytoplasm, CRP1 binds to α-actinin and zyxin to reorganize the cytoskeleton (38). The Ajuba LIM domain protein illustrates a similar pathway highly relevant to cancer progression. Our laboratory has shown that the 3 LIM domain protein Ajuba binds to the SNAG repression domain in the snail/slug transcription factors and plays a key role in the EMT program during tumor progression by repressing E-cadherin and other genes (39, 40). However, when in the cytoplasm, Ajuba is found at cell adhesion junctions and binds to CTNNA, F-actin, and GRB2. It is noteworthy that PINCH1 can be found in both the cytoplasm and the nucleus, and this location has prognostic value in primary human tumors (33).

In summary, the study of LIMD2 as a marker identified in a simple profiling study of primary and metastatic tumors has yielded a new player in the ILK signaling pathway. It is conceivable that small molecule approaches to disrupting

LIMD2 binding to its partners could influence the metastatic process.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: H. Peng, M. Talebzadeh-Farrooji, M.J. Osborne, Z. Hou, M. He, E. Kebebew, K.L.B. Borden, F.J. Rauscher

Development of methodology: H. Peng, M. Talebzadeh-Farrooji, M.J. Osborne, Z. Hou, M. He, M. Herlyn, A.J. Caton, C.S. Paterno, Q. Huang, S. Dedhar

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M. Talebzadeh-Farrooji, M.J. Osborne, P.C. McDonald, J. Karar, Z. Hou, M. He, E. Kebebew, W. Fredericks, B. Malkowicz, A.S. Carolin, E. Skordalakes, Q. Huang, K.L.B. Borden, F.J. Rauscher

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): H. Peng, M. Talebzadeh-Farrooji, M.J. Osborne, J.W. Prokop, P.C. McDonald, M. He, E. Kebebew, T. Orntoft, D.W. Speicher, Q. Huang, S. Dedhar, K.L.B. Borden, F.J. Rauscher

Writing, review, and/or revision of the manuscript: H. Peng, M. Talebzadeh-Farrooji, M.J. Osborne, J.W. Prokop, P.C. McDonald, J. Karar, M. He, E. Kebebew, W. Fredericks, B. Malkowicz, D.W. Speicher, S. Dedhar, K.L.B. Borden, F.J. Rauscher

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M. Talebzadeh-Farrooji, M. He, T. Orntoft, M. Herlyn, F.J. Rauscher

Study supervision: M. Talebzadeh-Farrooji, S. Dedhar, K.L.B. Borden, F.J. Rauscher

Acknowledgments

The authors thank F. Keeney (Wistar) for help with the imaging studies and The Wistar Institute Cancer Center Shared Facilities. They also thank R. Fassler and colleagues for the gift of the ILK$^{-/-}$ cells and D. Calderwood and T. Boggon (Yale University) for ILK and PINCH1 reagents.

Grant Support

Work in the Rauscher laboratory is supported by NIH grants CA129833, CA010815, CA163761, and DOD-BCRP W81XWH-11-1-0494, The Samuel Waxman Cancer Research Foundation, Susan G. Komen for the Cure, and The Noreen O'Neill Foundation for Melanoma Research. K.L.B. Borden holds a Canada Research Chair and is supported by NIH grants NIH RO1-80728 and NIH 98571.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 3, 2013; revised November 26, 2013; accepted December 9, 2013; published online March 3, 2014.

References

1. Fidler IJ. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nat Rev Cancer 2003;3:453–8.
2. Sethi N, Kang Y. Unravelling the complexity of metastasis—molecular understanding and targeted therapies. Nat Rev Cancer 2011;11:735–48.
3. Scott J, Kuhn P, Anderson AR. Unifying metastasis—integrating intravasation, circulation and end-organ colonization. Nat Rev Cancer 2012;12:445–6.
4. Hunter K. Host genetics influence tumour metastasis. Nat Rev Cancer 2006;6:141–6.
5. Alderton GK. Metastasis: epithelial to mesenchymal and back again. Nat Rev Cancer 2013;13:3.
6. Gumireddy K, Huang Q. Identification of metastasis genes by a functional genomics approach. Cell Cycle 2010;9:423.
7. Gumireddy K, Li A, Gimotty PA, Klein-Szanto AJ, Showe LC, Katsaros D, et al. KLF17 is a negative regulator of epithelial-mesenchymal transition and metastasis in breast cancer. Nat Cell Biol 2009;11:1297–304.
8. Nguyen DX, Bos PD, Massague J. Metastasis: from dissemination to organ-specific colonization. Nat Rev Cancer 2009;9:274–84.
9. Cerutti JM, Oler G, Michaluart P Jr, Delcelo R, Beaty RM, Shoemaker J, et al. Molecular profiling of matched samples identifies biomarkers of papillary thyroid carcinoma lymph node metastasis. Cancer Res 2007;67:7885–92.
10. Matthews JM, Lester K, Joseph S, Curtis DJ. LIM-domain-only proteins in cancer. Nat Rev Cancer 2013;13:111–22.
11. Chiswell BP, Zhang R, Murphy JW, Boggon TJ, Calderwood DA. The structural basis of integrin-linked kinase-PINCH interactions. Proc Natl Acad Sci U S A 2008;105:20677–82.
12. Peng H, Begg GE, Harper SL, Friedman JR, Speicher DW, Rauscher FJ 3rd. Biochemical analysis of the Kruppel-associated box (KRAB) transcriptional repression domain. J Biol Chem 2000;275:18000–10.
13. Peng H, Ivanov AV, Oh HJ, Lau YF, Rauscher FJ 3rd. Epigenetic gene silencing by the SRY protein is mediated by a KRAB-O protein that recruits the KAP1 co-repressor machinery. J Biol Chem 2009;284:35670–80.
14. Pelton JG, Torchia DA, Meadow ND, Roseman S. Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. Protein Sci 1993;2:543–58.
15. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–93.
16. Goddard TD, Kneller DG (2003), SPARKY (University of California, San Francisco, CA).
17. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 2009;30:2785–91.
18. Maydan M, McDonald PC, Sanghera J, Yan J, Rallis C, Pinchin S, et al. Integrin-linked kinase is a functional Mn2+-dependent protein kinase that regulates glycogen synthase kinase-3β (GSK-3β) phosphorylation. PLoS ONE 2010;5:e12356.
19. Kadrmas JL, Beckerle MC. The LIM domain: from the cytoskeleton to the nucleus. Nat Rev Mol Cell Biol 2004;5:920–31.
20. Grishin NV. Treble clef finger—a functionally diverse zinc-binding structural motif. Nucleic Acids Res 2001;29:1703–14.
21. Chiswell BP, Stiegler AL, Razinia Z, Nalibotski E, Boggon TJ, Calderwood DA. Structural basis of competition between PINCH1 and PINCH2 for binding to the ankyrin repeat domain of integrin-linked kinase. J Struct Biol 2010;170:157–63.
22. Velyvis A, Vaynberg J, Yang Y, Vinogradova O, Zhang Y, Wu C, et al. Structural and functional insights into PINCH LIM4 domain-mediated integrin signaling. Nat Struct Biol 2003;10:558–64.
23. Rooney N, Streuli CH. How integrins control mammary epithelial differentiation: a possible role for the ILK-PINCH-Parvin complex. FEBS Lett 2011;585:1663–72.
24. Sakai T, Li S, Docheva D, Grashoff C, Sakai K, Kostka G, et al. Integrin-linked kinase (ILK) is required for polarizing the epiblast, cell adhesion, and controlling actin accumulation. Genes Dev 2003;17:926–40.
25. Hannigan G, Troussard AA, Dedhar S. Integrin-linked kinase: a cancer therapeutic target unique among its ILK. Nat Rev Cancer 2005;5:51–63.
26. Legate KR, Montanez E, Kudlacek O, Fassler R. ILK, PINCH and parvin: the tIPP of integrin signalling. Nat Rev Mol Cell Biol 2006;7:20–31.
27. Wu C. PINCH, N(i)ck and the ILK: network wiring at cell-matrix adhesions. Trends Cell Biol 2005;15:460–6.
28. Hannigan GE, McDonald PC, Walsh MP, Dedhar S. Integrin-linked kinase: not so 'pseudo' after all. Oncogene 2011;30:4375–85.
29. Fukuda K, Gupta S, Chen K, Wu C, Qin J. The pseudoactive site of ILK is essential for its binding to α-Parvin and localization to focal adhesions. Mol Cell 2009;36:819–30.
30. Velyvis A, Yang Y, Wu C, Qin J. Solution structure of the focal adhesion adaptor PINCH LIM1 domain and characterization of its interaction with the integrin-linked kinase ankyrin repeat domain. J Biol Chem 2001;276:4932–9.
31. Zhang Y, Chen K, Tu Y, Velyvis A, Yang Y, Qin J, et al. Assembly of the PINCH-ILK-CH-ILKBP complex precedes and is essential for localization of each component to cell-matrix adhesion sites. J Cell Sci 2002;115:4777–86.

32. Persad S, Dedhar S. The role of integrin-linked kinase (ILK) in cancer progression. Cancer Metastasis Rev 2003;22:375–84.
33. Kovalevich J, Tracy B, Langford D. PINCH: more than just an adaptor protein in cellular response. J Cell Physiol 2011;226:940–7.
34. Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell 2012;149:1233–44.
35. Jing H, Vakoc CR, Ying L, Mandat S, Wang H, Zheng X, et al. Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. Mol Cell 2008;29:232–42.
36. Kuo JC, Han X, Hsiao CT, Yates JR 3rd, Waterman CM. Analysis of the myosin-II-responsive focal adhesion proteome reveals a role for β-Pix in negative regulation of focal adhesion maturation. Nat Cell Biol 2011;13:383–93.
37. Hervy M, Hoffman L, Beckerle MC. From the membrane to the nucleus and back again: bifunctional focal adhesion proteins. Curr Opin Cell Biol 2006;18:524–32.
38. Chang DF, Belaguli NS, Iyer D, Roberts WB, Wu SP, Dong XR, et al. Cysteine-rich LIM-only proteins CRP1 and CRP2 are potent smooth muscle differentiation cofactors. Dev Cell 2003;4: 107–18.
39. Hou Z, Peng H, White DE, Negorev DG, Maul GG, Feng Y, et al. LIM protein Ajuba functions as a nuclear receptor corepressor and negatively regulates retinoic acid signaling. Proc Natl Acad Sci U S A 2010;107:2938–43.
40. Hou Z, Peng H, White DE, Wang P, Lieberman PM, Halazonetis T, et al. 14-3-3 binding sites in the snail protein are essential for snail-mediated transcriptional repression and epithelial-mesenchymal differentiation. Cancer Res 2010;70:4385–93.

Cancer Research

Snail Recruits Ring1B to Mediate Transcriptional Repression and Cell Migration in Pancreatic Cancer Cells

Jiangzhi Chen[1,2], Hong Xu[1], Xiuqun Zou[1], Jiamin Wang[1], Yi Zhu[1], Hao Chen[1], Baiyong Shen[1], Xiaxing Deng[1], Aiwu Zhou[3], Y. Eugene Chin[3], Frank J. Rauscher, III[4], Chenghong Peng[1], and Zhaoyuan Hou[1,3]

Abstract

Transcriptional repressor Snail is a master regulator of epithelial–mesenchymal transition (EMT), yet the epigenetic mechanism governing Snail to induce EMT is not well understood. Here, we report that in pancreatic ductal adenocarcinoma (PDAC), elevated levels of the ubiquitin E3 ligase Ring1B and Snail, along with elevated monoubiquitination of H2A at K119 (H2AK119Ub1), are highly correlated with poor survival. Mechanistic investigations identified Ring1B as a Snail-interacting protein and showed that the carboxyl zinc fingers of Snail recruit Ring1B and its paralog Ring1A to repress its target promoters. Simultaneous depletion of Ring1A and Ring1B in pancreatic cancer cells decreased Snail binding to the target chromatin, abolished H2AK119Ub1 modification, and thereby compromised Snail-mediated transcriptional repression and cell migration. We found that Ring1B and the SNAG-associated chromatin modifier EZH2 formed distinct protein complexes with Snail and that EZH2 was required for Snail-Ring1A/B recruitment to the target promoter. Collectively, our results unravel an epigenetic mechanism underlying transcriptional repression by Snail, suggest Ring1A/B as a candidate therapeutic target, and identify H2AK119Ub1 as a potential biomarker for PDAC diagnosis and prognosis. *Cancer Res; 74(16); 4353–63. ©2014 AACR.*

Introduction

Snail is a member of SNAG domain containing zinc finger proteins and a master regulator of epithelial–mesenchymal transition (EMT) and metastasis in various tumor types (1–3). Snail can directly bind to the E-boxes of E-cadherin gene promoter to repress its transcription and convert normal epithelial cells into mesenchymal cell phenotype (4, 5). Mechanistically, Snail recruits multiple repressive protein complexes involving histone deacetylation and methylation as well as DNA methylation to its target promoters and exerts its repressive function (6–10). However, how these protein complexes are assembled at the target chromatin regions remains elusive.

Ring1A and Ring1B belong to the RING domain containing ubiquitin E3 ligase family and are crucial components of the polycomb repressive complex 1 (PRC1) by catalyzing monoubiquitination of histone H2A at lysine 119 (H2AK119Ub1; ref. 11). H2AK119Ub1 status is associated with gene silencing (12), chromatin remodeling (13), and X chromosome inactivation (14). Genetic disruption of Ring1B in mice causes embryonic lethality due to gastrulation arrest and defective mesoderm formation (15), which are reminiscent of the Snail null mice. Snail-deficient mouse embryos also die early in gestation, displaying defects in gastrulation and mesoderm formation (16, 17). The striking genetic evidence strongly indicates that Snail very likely correlates with Ring1B in regulation of gastrulation and mesoderm formation during embryo development.

A recent study found an elevated Ring1B expression in high-grade pancreatic ductal adenocarcinoma (PDAC; ref. 18); however, no biologic and mechanistic studies on Ring1B were performed. In this present study, we report that Ring1A and Ring1B interact with Snail and are important coregulators for Snail-mediated transcriptional repression and cell migration in pancreatic cancer cells. Snail enhances the binding of Ring1A and Ring1B to the promoter region of E-cadherin to increase H2AK119Ub1 at this locus, and an elevated Snail, Ring1B, and H2AK119Ub1 modification in PDAC is highly correlated with poor prognosis.

Materials and Methods

IHC and tissues microarray

IHC staining was performed as described (18) with specific antibodies against Snail, Ring1B, or H2AK119Ub1. Tissue

[1]Department of Surgery, Ruijin Hospital, Department of Biochemistry and Molecular Cell Biology, Shanghai Jiaotong University School of Medicine, Shanghai, China. [2]Department of Hepatobiliary Surgery, Union Hospital, Fujian Medical University, Fuzhou, China. [3]Shanghai Key Laboratory for Tumor Microenvironment and Inflammation, Shanghai Jiaotong University School of Medicine, Shanghai, China. [4]Institute of Health Sciences, Chinese Academy of Sciences, Shanghai, China. [5]The Wistar Institute, Philadelphia, Pennsylvania.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

J. Chen and H. Xu contributed equally to this article.

Corresponding Authors: Zhaoyuan Hou, Shanghai Jiaotong University, 280 South Chongqing Road, Building #7, Room 110, Shanghai 200025, China. Phone: 15121148204; Fax: 86-021-64661525; E-mail: houzy@sjtu.edu.cn; and Chenghong Peng, No.197, Ruijin 2nd Road, Building #36, Floor 5, Shanghai 200025, China. E-mail: chhpeng@yearh.net

doi: 10.1158/0008-5472.CAN-14-0181

microarray (TMA) chips that contain 90 cases of paired tumor and peri-tumor specimens were purchased (HPan-Ade180 Sur-01; ShGnghGi Outdo Biotech Company). All specimens spotted on TMA chips were well documented including complete postoperative fellow-ups for periods from 3 to 7 years. Tumor staging was evaluated according to the TNM Classification of Malignant Tumors. After IHC staining, all specimens were strictly evaluated by a senior pathologist and only those with a tumor content of more than 50% were considered as tumor samples and the peri-tissues showing sign of chronic pancreatitis were excluded, and total of 60 cases are qualified for further analysis. Peri-tumor tissues were defined as that more than 1 cm distant from the tumor edge. TMA chips were scanned by Aperio Scanscope XT, and the whole field of each tissue spot was obtained for IHC evaluation. The immunoreactive score system (IRS; refs. 19, 20) was used to evaluate the staining of each sample as follow: negative, 0 to 1 point; mild (+), 2 to 3 points; moderate (++), 4 to 8 points; strongly positive (+++), 9 to 12 points. IHC staining showing IRS scores of more than 4 points was considered as high.

Plasmids

Ring1A and Ring1B cDNAs and their mutants were cloned into pCMV4-Flag vector between *Hin*d III and *Eco* RI sites. Snail and its mutants were cloned into pCMV5-HA vector between *Hin*dIII and *Xba*I sites. pLKO.1-shRNAs targeting Ring1A were ATAGATCTTAGAGATCAGGGC and ATCGTTGTGGTCTGA-TCTGAC; targeting Ring1B were ATTGTGCTTGTTGAT-CCTGGC and TTCTAAAGCTAACCTCACAGC, respectively. All point mutants were made using the QuikChange Site-Directed Mutagenesis procedures (Stratagene), and were confirmed by DNA sequencing.

Cell culture and transfections

HEK-293T cells and pancreatic cancer cells PanC1 and AsPC1 were obtained from the ATCC and were tested and authenticated by DNA typing at the Shanghai Jiao Tong University Analysis Core. The cells were maintained in DMEM supplemented with 10% FBS, 2 mmol/L L-glutamine, and penicillin (50 U/mL)/streptomycin (50 μg/mL) at 37°C under 5% CO_2 in a humidified chamber.

Transfection of PanC1 and HEK-293T cells was performed using Lipofectamine 2000 as described (8). The viral supernatants were generated in HEK-293T cells, and were infected into PanC1 and AsPC1 cells. Puromycin was added into the media to generate stable knockdown of Ring1A and Ring1B in PanC1 and AsPC1 cells. FACS was performed to sort the cells stably expressing Flag-Snail.

Affinity purification of Snail-interacting protein complex

A Flag-tagged, full-length Snail cDNA in the pcDNA3.1- vector was stably expressed in HEK-293T cells. Single-cell clones were selected with G418 and screened by Western blot assays using anti-Flag antibody. The method used for affinity purification was previously described (8). A total of 5×10^9 cells were used for affinity purification, and the eluted proteins were resolved on 4% to 12% SDS-PAGE gels (Invitrogen) for Western blot and colloidal staining analyses. The proteins were excised from the gel and identified by standard mass spectrometry.

Coimmunoprecipitation, Western blot, immunofluorescence, and antibodies

Plasmids encoding Flag-Ring1A, Flag-Ring1B, hemagglutinin (HA)-Snail proteins were transiently expressed in HEK-293T cells, and 24 hours after transfection, cells were lysed in buffer containing 20 mmol/L Tris-HCl (pH 8.0), 150 mmol/L NaCl, 2.5 mmol/L EDTA, 0.5% NP40, 0.1 mmol/L phenylmethylsulfonylfluoride, and protease inhibitor cocktail. Method for total histones extraction was as described (12). The whole-cell extracts were precleared with protein A/G beads, and coimmunoprecipitation (co-IP) assays were performed with either Flag or HA antibodies. The methods used for Western blot and immunofluorescence were previously described (8). Antibodies for Flag (Sigma-Aldrich; F 7425), HA (COVANCE; MMS-101P), Ring1A, Ring1B, H2A, ubiquityl-Histone H2A-lys119 and E-cadherin (Cell Signaling Technology; #2820, #5694,#2578,#8240, #3195), Snail (Santa Cruz; sc-28199); and β-actin (Proteintech; 60008–1-Ig) were purchased.

Chromatin immunoprecipitation and qPCR

The chromatin immunoprecipitation (ChIP) experiments were carried out in PanC1 cells and derivatives. To prepare cells for ChIP assays, the PanC1 cells were grown in 10 cm plates to 70% to 90% confluency and were processed as described (8). The immunoprecipitated DNA fragments were detected by qPCR assays. The primer sets that amplify the DNA fragment flanking the known E-boxes in the E-cadherin promoter are as follows: forward, 5-GCAGGTGAACCCTCAGC-CAA-3; reverse, 5-CACAGGTGCTTTGCAGTTCC-3.

Total RNA was isolated from cells with TRIzol reagent (Invitrogen). qRT-PCR was performed on a 7500 Fast Real-time PCR system (Applied Biosystem) using SYBR Green agent. Primers used for qRT-PCR assay were listed in Supplementary information. All RT-PCR assays were repeated three times.

Transwell cell migration assays

PanC1 cells were harvested after serum-free starvation for 12 hours, and were resuspended in plain DMEM media. Ten thousand cells were applied to 8-μm pore transwell filters (Corning). DMEM media containing 10% FBS were added to the bottom chamber as attractants. After incubation for 24 hours, the nonmigrated cells at the top of the filter were removed and the migrated cells at the bottom of the filter were fixed with 4% paraformaldehyde and were stained with colloidal staining method. The number of migrating cells in each chamber was quantified by counting nine randomly chosen fields under ×20 magnification using a bright-field microscope. Each condition was performed in duplicate, and the average number of cells per field was represented. Experiments were repeated three times.

Statistical analysis

Data shown as mean ± SD were analyzed by the independent Student *t* test. The distribution of the IHC scoring results

of each protein on TMA chips was analyzed by the McNemar test. The correlation between the expression of Snail and Ring1B in PDAC was analyzed by the Spearman rank correlation coefficient test. Spearman ρ are categorized as moderate to strong correlations according to Dancey and Reidy's categorization: 0 (zero); 0.1 to 0.3 (weak); 0.4 to 0.6 (moderate); 0.7 to 0.9 (strong); and 1 (perfect; ref. 21). The postoperative survival of patients with PDAC was analyzed by Kaplan–Meier estimator and tested by the log-rank. P value of <0.05 was considered statistically significant. Statistical analysis was performed using IBM SPSS Statistics Version 20 software.

Study approval

The study was approved by the Ethic and Research Committees of Ruijing Hospital, Shanghai Jiaotong University School of Medicine and was conducted in accordance with the Declaration of Helsinki Principles. The procedures for pancreatic tumor resection were described in detail to all patients before admission, and informed consent of patients was obtained.

Results

Expression of Snail and Ring1B proteins in PDAC is positively correlated

To examine the clinical relevance of Snail, Ring1B and level of H2AK119Ub1 in PDAC, IHC assays were performed on human pancreatic ductal tumor resection specimens and paired peri-tumor tissues were used as normal controls. The staining for Snail was strong in tumor cells surrounding acinus and lumen, but was barely detected in epithelial cells of normal tissues (Fig. 1A, top). Similarly, the staining for Ring1B was strong in tumor cells especially in high-grade differentiated cells, whereas in normal tissues the staining was weak or not detectable except in islet cells that showed positive staining (Fig. 1A, middle). Notably, the staining for H2AK119Ub1 was strong in tumor cells, but remained low level in islet cells and acinar cells of the normal tissues (Fig. 1A, bottom).

To further elucidate the clinical correlation of Snail, Ring1B, and level of H2AK119Ub1 in PDAC, we performed IHC on TMA chips containing strictly selected 60 cases of paired tumor and peri-tumor specimens of PDAC. Of the 60 patients, 38 cases are males and 22 cases are females with an average age of 62 ± 11 years old. These patients have undergone pancreatectomy between September 2004 and December 2008, and the complete postoperative follow-up was finished on December 2011, with duration of 3 to 7 years. The overall median survival time of 60 patients after pancreatectomy was 15 months. The McNemar test was conducted to analyze the symmetrical distribution of the IHC staining for Snail, Ring1B, and H2AK119Ub1 among the 60 cases of samples according to the IRS scores (high: >4 vs. low: $\leq$4). We found that specimens showing high levels of Snail, Ring1B, and H2AK119Ub1 were distributed asymmetrically with much more cases found in tumor samples (Fig. 1B, $P < 0.05$). We next performed the Spearman rank correlation coefficient analysis to determine whether elevated expression of Snail and Ring1B in PDAC samples cooccurs. Indeed, there was a positive correlation between Snail and Ring1B (Fig. 1C, $P < 0.05$). Taken together, these data suggest Snail and Ring1B may functionally correlate in PDAC development and progression.

High level of H2AK119Ub1 predicts poor survival of patients with PDAC

To determine whether there is a correlation between high levels of Snail, Ring1B, H2AK119Ub1 and survival rate of postoperative patients with PDAC, Kaplan–Meier analysis was conducted. We found that patients with PDAC expressing high level of Snail and high level of H2AK119Ub1 showed tendency of higher mortality and was positively correlated with poor survival (Fig. 1D, median survival time: 9 months; Fig. 1F, median survival time: 11 months); expression level of Ring1B displayed similar tendency, but was not statistically significant (Fig. 1E). Remarkably, patients with PDAC showing both high level of Snail and Ring1B strongly predicted poor survival (Fig. 1G, median survival time: 9 months). Taken together, these data indicate that high level of H2AK119Ub1 in PDAC is a potential epigenetic biomarker for prediction of PDAC prognosis.

Ring1A and Ring1B are Snail-interacting proteins

To examine if there is a physical interaction between Snail and Ring1B, we stably expressed N-terminal epitope-tagged Snail protein (Flag-Snail) in HEK-293T cells, and performed affinity purification of the Snail-interacting proteins with Flag antibody. The coeluted proteins along with Snail were identified by mass spectrometry analysis. Notably, Ring1B was found in the protein band migrating at 43 kDa (Fig. 2A). Next, we coexpressed HA-Snail and Flag-Ring1B in HEK-293T cells and performed co-IP assays to verify their interaction. Indeed, Snail and Ring1B interacted with each other (Fig. 2B). Ring1A also interacted with Snail (Fig. 2C). To further confirm the interaction of Snail with Ring1B, we attempted to detect the endogenous interaction. The endogenous Ring1B and Snail were readily detected by Western blot assays using specific antibodies in PanC1 cells. Co-IP assays indicated Ring1B interacted with Snail at endogenous level (Fig. 2D). Collectively, these data clearly demonstrate that Ring1A and Ring1B are novel Snail-interacting proteins.

Ring1A and Ring1B bind to the carboxyl tandem zinc fingers of Snail

To identify the regions in the Snail protein that interact with Ring1A and Ring1B, plasmids encoding the N-terminal SNAG domain (Snail-NT) or C-terminal zinc fingers (Snail-CT) were constructed respectively and were expressed in HEK-293T cells alone with Ring1A or Ring1B (Fig. 3A). Co-IP assays demonstrated that the zinc finger region retained the binding activity with Ring1A and Ring1B (Fig. 3B), whereas the SNAG domain and the linker region of Snail showed no binding activity.

The RING domains of Ring1A and Ring1B are required for the interaction with Snail

To identify the regions of Ring1A and Ring1B responsible for the interaction with Snail, we generated a series of

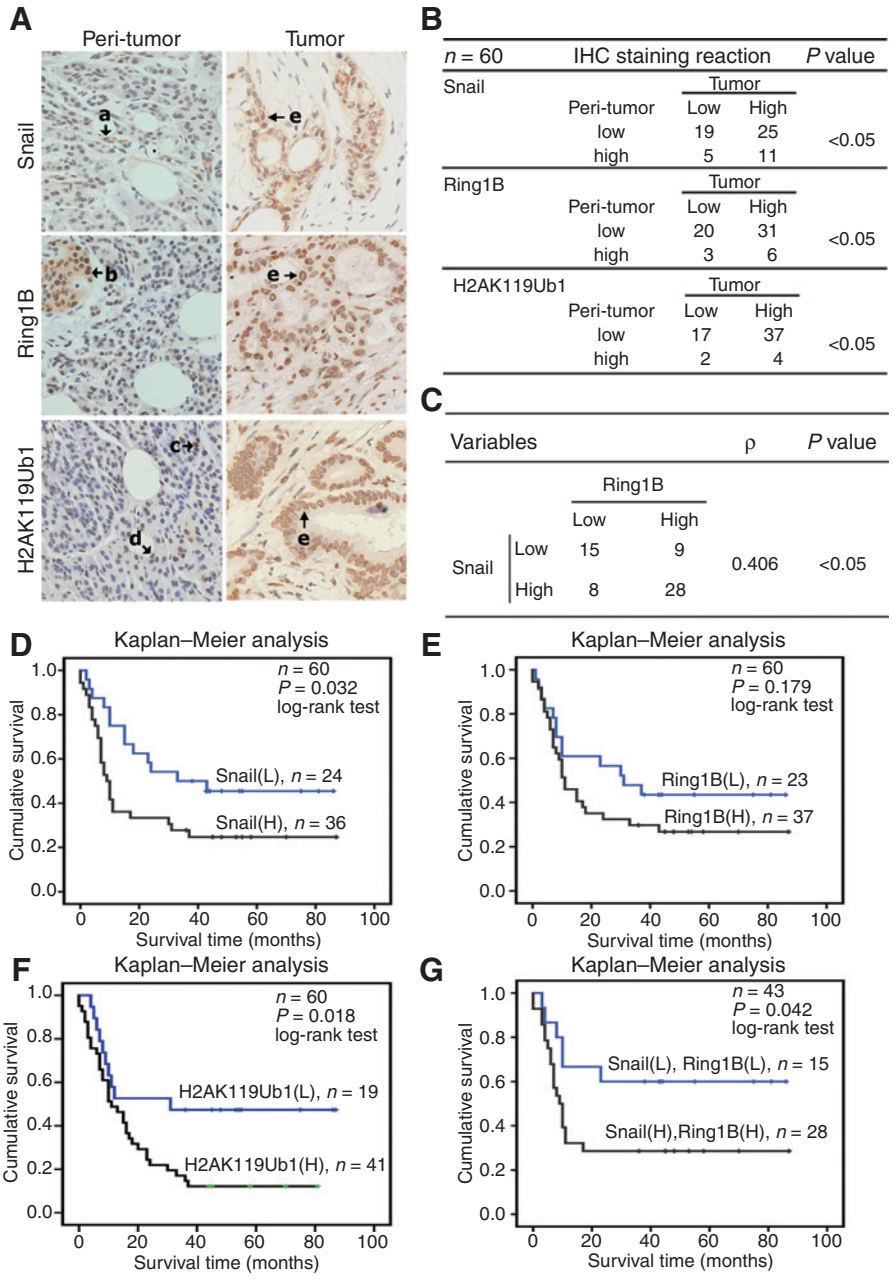

Figure 1. High expression of Snail and Ring1B is positively correlated and predicts poor prognosis in PDAC. A, IHC staining to detect Snail, Ring1B, and H2AK119Ub1 in paired human PDAC specimens (tumor vs. peri-tumor). Arrows, positively stained cells. a and d, duct epithelial cells; b, islet cells; c, acinar cells; e, tumor cells. Original magnifications, ×40. B, the McNemar test of the IHC staining of 60 cases of the paired PDAC specimens on TMA chips. Each group was shown by the distribution of IHC staining scores for each case. Only IHC scores ≥ 4 point (++) was considered high. $n = 60$; *, statistical significance, $P < 0.05$. C, Spearman ρ analysis of correlation of Snail and Ring1B expression in PDAC. $n = 60$. $\rho > 0$, positive correlation; $\rho < 0$, negative correlation. D–G, survival analysis of patients with PDAC patients by Kaplan–Meier plots and log-rank tests. Patients were categorized by high and low expression of the Ring1B, Snail, and H2AK119Ub1 based on IHC staining scores. H, high; L, low. Average follow-up time of 60 patients was 26.2 months (range, 0–87 months). The median survival time of each group is as follows: Snail (L), 33 months; Snail (H), 9 months; Ring1B (L), 31 months; Ring1B (H), 11 months; H2AK119Ub1 (L), 31 months; H2AK119Ub1 (H), 11 months; Snail (L) + Ring1B (L), the median survival time cannot be estimated because more than 50% of cases are censored in this group; Snail (H) + Ring1B (H), 9 months.

deletion mutants of Ring1A (Fig. 3C, left). Next, we coexpressed the deletion mutants for Ring1A and Snail in HEK-293T cells and performed co-IP assays with Flag antibody. We observed that the N-terminal Ring1A containing the conserved RING domain retained the binding ability to Snail (Fig. 3C, right), but the C-terminal region of Ring1A was unable to bind Snail.

To further determine if the RING domain contributes to the interaction with Snail, simultaneous mutations of the 66th histidine (H66A) and 69th cysteine (C69A) to alanines in Ring1A were made to disrupt the RING structure (Fig. 3D). Co-IP assays showed that the double mutation of the key residues H66 and C69 of Ring1A resulted in loss of the binding activity to Snail (Fig. 3D), suggesting that the RING domain is critical for Snail binding. Similarly, mutation of H69 and C72 to alanines in the RING domain of Ring1B abolished its binding to Snail (Fig. 3E). Further, we showed that mutations of these core amino acid residues in the RING domains did not change the subcellular localization of Ring1A and Ring1B proteins (Supplementary Fig. S1). Taken together, these results demonstrate that the RING domains of Ring1A and Ring1B are required for their interaction with Snail.

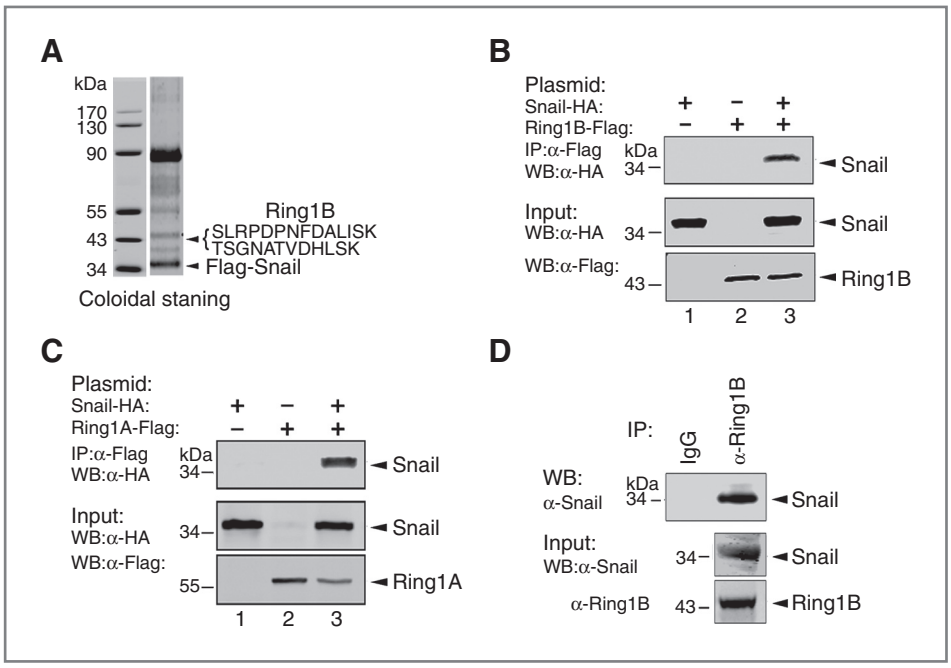

Figure 2. Ring1A and Ring1B are Snail-interacting proteins. A, mass spectrometry analysis identifies Ring1B as a Snail-interacting protein. Affinity purification of lysate from HEK-293T cell stably expressing Flag-Snail was performed with Flag antibody. SDS-PAGE and colloidal staining show proteins coeluted with Snail. Mass spectrometry analysis discovered the band migrating at 43 kDa contains Ring1B and the major band migrating at 34 kDa is Flag-Snail protein. B, Ring1B interacts with Snail. Lysates from HEK-293T cells transfected with plasmids encoding HA-Snail and/or Flag-Ring1B are incubated with Flag M2 beads, and the coeluted proteins were blotted with HA antibody. C, Ring1A interacts with Snail. D, Snail interacts with Ring1B at endogenous level. Lysates from PanC1 cells were incubated with anti-Ring1B, and the coeluted proteins were blotted with Snail antibody. Ten percent of the input was loaded. α, anti; WB, Western blot.

Ring1A and Ring1B are required for Snail-mediated E-cadherin gene repression and cell migration

To examine the role of Ring1A and Ring1B in Snail-mediated transcriptional repression, we first depleted their expression in PanC1 cells, a widely used pancreatic cancer cells, using specific shRNAs targeting Ring1A and Ring1B simultaneously to overcome the redundancy effect (Fig. 4A). Western blot assays indicated that depletion of Ring1A and Ring1B resulted in global decrease of H2AK119Ub1, but no apparent change on H3K27Me3 (Fig. 4A). qRT-PCR assays showed that expression of the Snail target genes E-cadherin and Cyclin D2 and the known Ring1A/B target genes HOXC5 and HOXB4 was increased in PanC1-shRing1A/B cells (Fig. 4B and Supplementary Fig. S2). Next, we stably expressed Flag-Snail in PanC1-shLuc or PanC1-shRing1A/B cells using lentiviral system (Fig. 4C). The expression of Snail, Ring1A, and Ring1B was validated by Western blot assays. The protein and mRNA levels of E-cadherin were examined by Western blot and qRT-PCR approaches. As expected, expression of Snail markedly reduced both the protein and mRNA levels of E-cadherin in PanC1-shLuc cells, but still slightly decreased the E-cadherin expression in PanC1-shRing1A/B cells (Fig. 4C and D). In addition, we observed similar effect of depletion of Ring1A and Ring1B on Snail-mediated transcriptional repression in another pancreatic cancer AsPC1 cells (Supplementary Fig. S3).

To examine the role of Ring1A and Ring1B in Snail-induced cell migration, we performed standard transwell assays to measure the migration capability of the PanC1 cells and its derivatives. Expression of Snail robustly enhanced migration capability of the PanC1-shLuc cells. However, expression of Snail in PanC1-shRing1A/B cells increased cell migration, but at lesser extent compared with that of in PanC1-shLuc-Snail cells (Fig. 4E and F). Collectively, these data suggest that Ring1A and Ring1B are essential for Snail to maximally repress E-cadherin and to induce cell migration.

Snail recruits Ring1A and Ring1B to the promoter locus of E-cadherin

To examine if Snail recruits Ring1A and Ring1B to the E-cadherin promoter, we performed ChIP assays in PanC1-derivated cells with antibodies specific to Snail, Ring1A, Ring1B, or H2AK119Ub1. The coeluted DNA fragments were amplified using primer pairs flanking the E-boxes located in the proximal promoter of E-cadherin (Fig. 5A). In PanC1-shLuc cells, expression of Snail resulted in an increased binding of Snail, Ring1A, and Ring1B to the E-cadherin promoter, and a concomitant increase in H2AK119Ub1 modification at this locus (Fig. 5B–E). Depletion of Ring1A and Ring1B simultaneously resulted in loss of their binding to the E-cadherin promoter, diminished H2AK119Ub1 modification, and decreased binding of Snail to the E-cadherin promoter, suggesting Ring1A and Ring1B may stabilize Snail/DNA complexes at the target promoter. However, depletion of Ring1A and Ring1B did not affect EZH2 binding and H3K27 methylation at the E-cadherin promoter locus

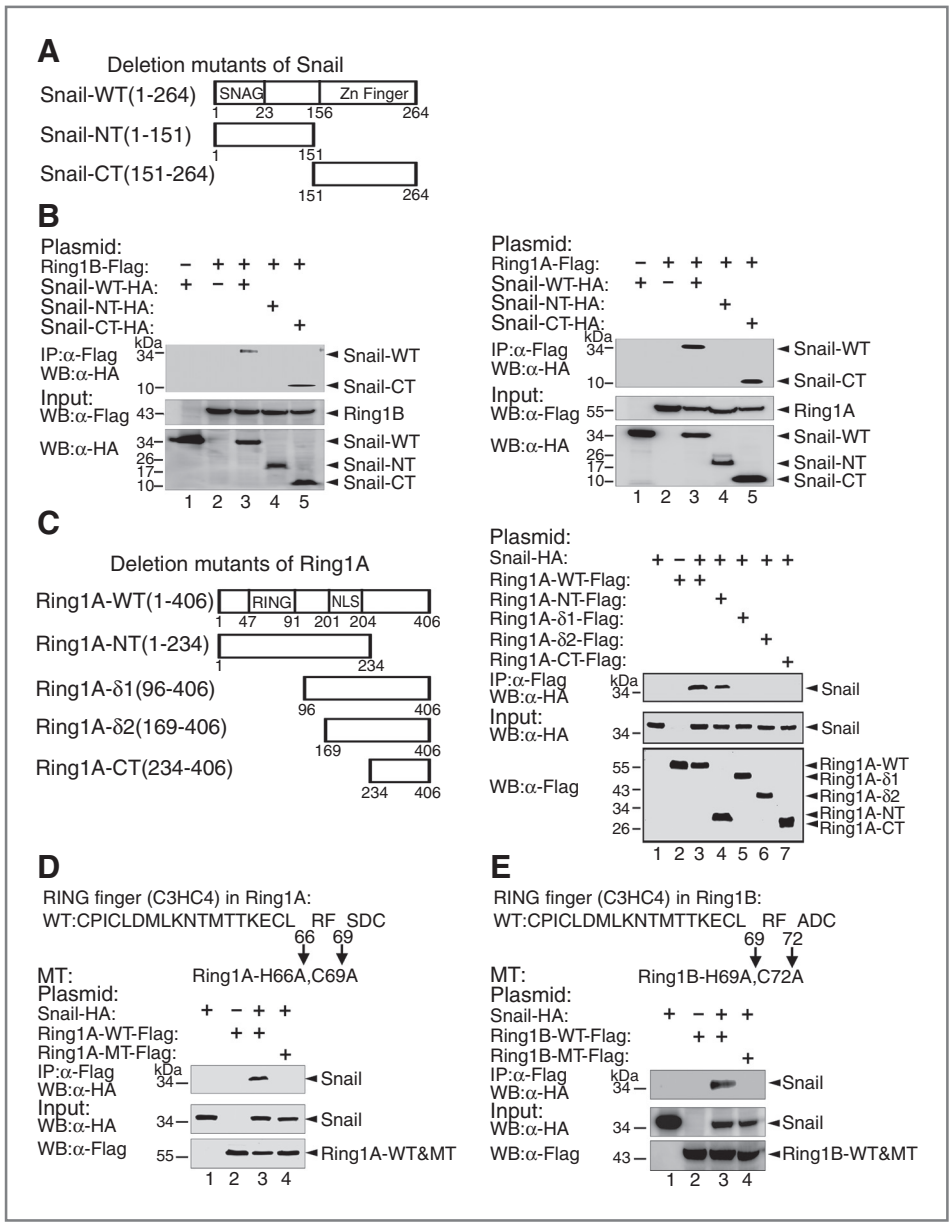

Figure 3. A, the RING domains of Ring1A and Ring1B and the zinc fingers of Snail are required for their interaction. Schematic representation of deletion mutants of Snail. B, the C-terminal zinc fingers of Snail interact with Ring1A and Ring1B. Lysates prepared from HEK-293T cells transfected with plasmids encoding HA-Snail and its truncation mutants, together with Flag-Ring1A or Flag-Ring1B were incubated with anti-Flag M2 beads, and coeluted proteins were analyzed by anti-HA. C, N-terminal (1–96 aa) of Ring1A interacts with Snail schematic representation of deletion mutants of Ring1A. N-terminus of Ring1A contains a conserved RING finger domain. D, mutation of the 66th histidine and 69th cysteine in the RING finger of Ring1A abolishes the binding between Snail and Ring1A. E, mutation of the 69th histidine and No. 72 cystenie to alanines in the RING finger of Ring1B abolishes the interaction of Snail and Ring1B. WT, wild-type; MT, mutant-type; NT, Amino terminus; CT, carboxyl terminus.

(Fig. 5F and G). Taken together, these data demonstrate that Snail recruits Ring1A and Ring1B to modify H2A at K119 at the target chromatin region.

Depletion of EZH2 reduces Snail/Ring1A/B multiprotein complex binding at the E-cadherin promoter locus

Previous studies demonstrated that EZH2 binds to the SNAG domain of Snail to repress E-cadherin expression by trimethylation of H3 at K27 (22, 23). To examine if EZH2 affects Ring1A/B-mediated H2AK119Ub1 modification, we generated PanC1-shEZH2 and PanC1-shLuc cells. Depletion of EZH2 resulted in markedly global decrease of H3K27Me3 and slight decrease of H2AK119Ub1 (Fig. 6A). Similar to that observed in PanC1-shRing1A/B cells, depletion of EZH2 increased E-cadherin expression as evidenced by Western blot and indirect immunofluorescence assays (Fig. 6B and C). Consistently, expression of Snail in PanC1-shLuc cells effectively repressed E-cadherin, but only showed weak repression on E-cadherin expression in PanC1-shEZH2 cells (Fig. 6B and C), indicating EZH2 is important for Snail-mediated E-cadherin repression. Next, ChIP assay was performed in these resulting cells. Depletion of EZH2 in PanC1 cells decreased the binding of Snail, Ring1A, and Ring1B to the E-cadherin promoter, diminished H3K27Me3, and greatly reduced H2AK119Ub1 modifications induced by Snail at this locus (Fig. 6D–I). These observations indicate that EZH2 functions upstream of Ring1A and Ring1B and promotes Snail to recruit Ring1A and Ring1B more efficiently at the E-cadherin promoter locus.

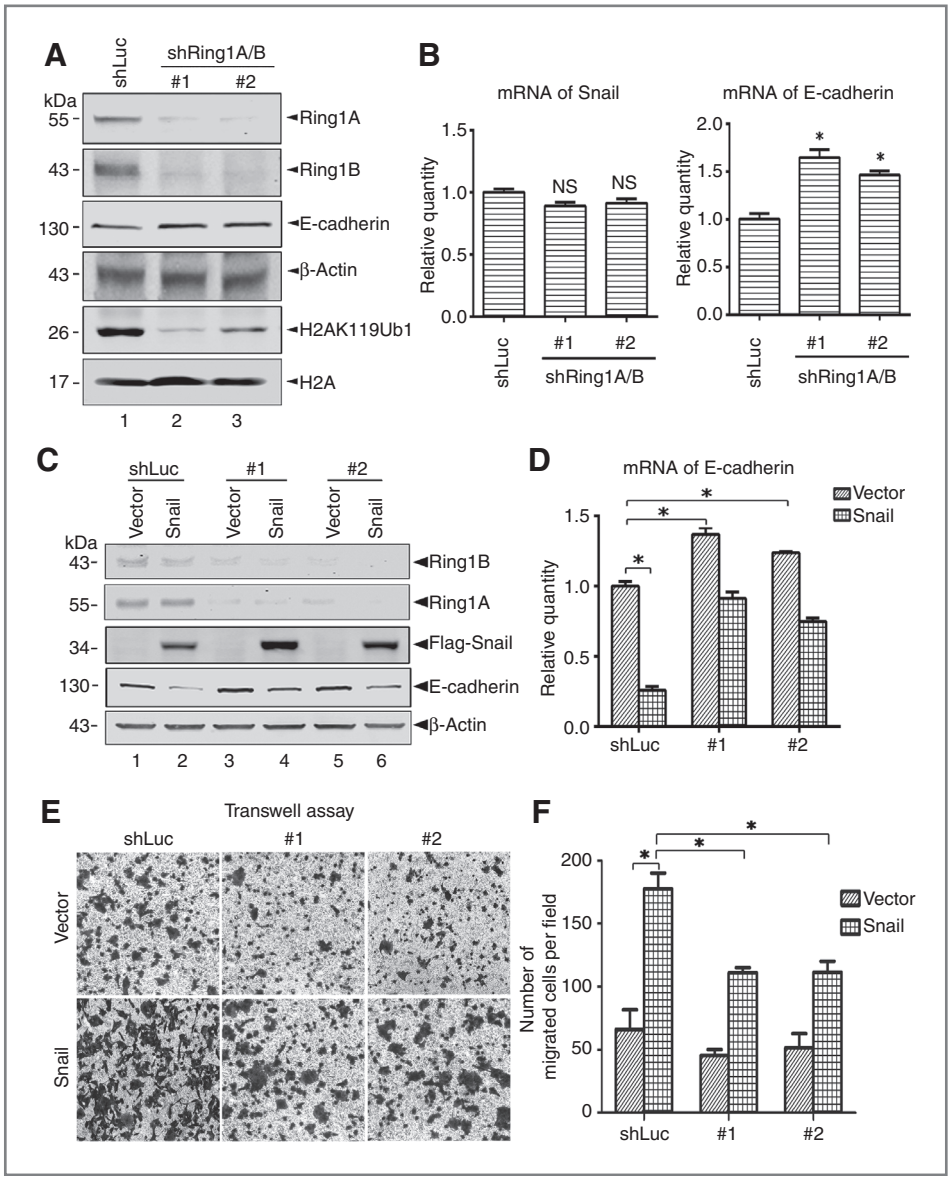

Figure 4. Ring1A and Ring1B are required for Snail-mediated E-cadherin gene repression and cell migration. A, simultaneous depletion of Ring1A and Ring1B in PanC1 cells results in global decrease of H2AK119Ub1. Double depletion of Ring1A and Ring1B in PanC1 cells was made using two sets of specific shRNAs targeting Ring1A and Ring1B, respectively, and created two pools of PanC1-shRing1A/B cells. Western blots were performed to examine the expression of Ring1A, Ring1B, H2AK119Ub1, and H3K27Me3. B, E-cadherin is upregulated in PanC1-shRing1A/B cells. Total RNAs were isolated from the resulting PanC1 cells, and expression of Snail and E-cadherin was examined by qRT-PCR. #1 and #2, different cell pools generated from Ring1A and Ring1B double knocking down by shRNAs; GAPDH was used as loading reference. Data, mean ± SD from three independent experiments. *, statistical significance, $P < 0.05$. NS, statistical insignificance, $P > 0.05$. C, Snail was stably expressed in PanC1-shLuc or PanC1-shRing1A/B cells. Plu-Snail-Flag was introduced into the resulting PanC1 cells vial lentiviral infection, and the cells were sorted by FACS. The expression of Snail-Flag, Ring1A, Ring1B, and E-cadherin was examined by Western blot assays. D, depletion of Ring1A and Ring1B compromises Snail's repression on E-cadherin. qRT-PCR assays were conducted to examine the mRNA of E-cadherin. GAPDH was used as loading reference. E, Snail induces cell migration in transwell assays. F, statistical analysis for the migrating cells from transwell assays was shown in the bar graphs. Nine randomly chosen fields were counted for statistic analysis. Data, mean ± SD from three independent experiments. #1 and #2, different cell pools generated from Ring1A and Ring1B double knocking down by shRNAs; *, $P < 0.05$.

EZH2 and Ring1B form distinct complexes with Snail

To determine if EZH2, Snail, and Ring1B form one functional complex, we coexpressed HA-Snail, Myc-EZH2, and Flag-Ring1B in HEK-293T cells and performed co-IP assays with Flag antibody. When these three proteins were coexpressed, Ring1B coimmunoprecipitated Snail, but not EZH2, indicating Snail, EZH2, and Ring1B do not coexist in the same complex (Fig. 7A). To further strengthen this observation, we performed size exclusion fractionation of whole-cell extracts prepared from HEK-293T cells expressing Snail, Ring1B, and EZH2. EZH2 was eluted in a single peak (fractions 12–14), Ring1B was detected in fraction 15 and was peaked at

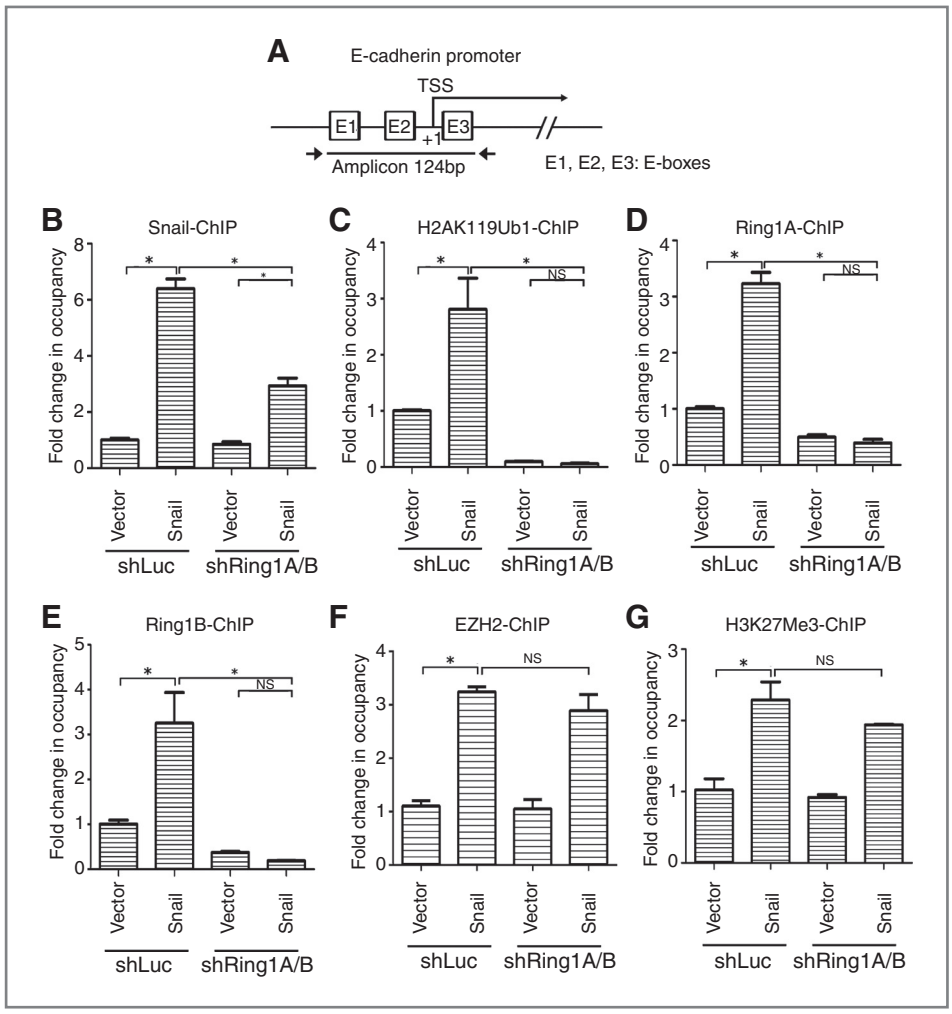

Figure 5. A, Snail recruits Ring1A and Ring1B to modify H2A at K119 at the target chromatin region. Schema shows PCR primers used for amplifying the E-cadherin promoter flanking the functional E-boxes. B–G, ChIP DNA fragments from PanC1cells were analyzed by qPCR assays. For comparison between cells, PCR amplifications were normalized to their inputs respectively. The units on the *y* axis were arbitrary; data, mean ± SD from three independent experiments. *, $P < 0.05$. NS, statistical insignificance, $P > 0.05$. Enrichment of Snail (B), H2AK119Ub1 (C), Ring1A (D), and Ring1B (E) at the E-cadherin promoter is increased when Snail is expressed, whereas depletion of Ring1A and Ring1B results in loss of Snail occupancy and diminished H2AK119Ub1 modification at the E-cadherin promoter. Snail increases EZH2 binding and H3K27Me3 at E-cadheirn promoter locus, but depletion of Ring1A and Ring1B does not affect EZH2 binding (F) and H3K27Me3 modification (G).

fractions 16 to 17; Snail was eluted in one peak (fractions 12–16; Fig. 7B). These observations indicate that Snail forms two complexes, one with EZH2 and another one with Ring1B. Together, these data suggest that EZH2, Snail, and Ring1B do not coexist in the same protein complex; rather, EZH2 and Ring1B form two distinct protein complexes with Snail respectively (Fig. 7C).

Discussion

PDAC is one of the most malignant human cancers, with a 5-year survival rate of only 5% and a median survival of less than 6 months (24, 25). Almost all patients with PDAC develop metastases due to its properties of rapid progression, potent invasiveness, and profound resistance to treatments. Identifying biomarkers for early diagnosis and prognosis is especially important to reduce the death rate of patients with PDAC.

H2AK119Ub1 modification is important during the embryonic development and organogenesis (12–15); however, the biologic significance of H2AK119Ub1 in tumor progression has not yet been elaborated. In this study, we found that H2AK119Ub1 is markedly elevated in tumor cells of PDAC, and the level of H2AK119Ub1 is positively associated with poor survival, indicating that H2AK119Ub1 is a potential epigenetic biomarker for diagnosis and prognosis of PDAC. However, we observed that levels of Snail, Ring1B, and H2AK119Ub1 are not correlated in some cases of PDAC specimens; for example, of the 60 cases, 36 cases show Snail (H), 37 cases show Ring1B (H), but only 28 cases show both Snail (H) and Ring1B (H; Fig. 1). The discrepancy between the level of Snail, Ring1B, and H2AK119Ub1 modification suggests additional mechanisms that are independent of the Snail/Ring1B complex and might be involved in Snail and Ring1B functions as well as H2AK119Ub1 modification in PDAC. To clarify this puzzle, ChIP-seq analyses of the occupancy of Snail and Ring1B, and the loci with high level of H2AK119Ub1 in the genome of PDAC specimens, will help to obtain global understanding of Snail, Ring1B, and H2AK119Ub1-dependent or -independent repressive pathways.

A number of the SNAG-associated histone modification complexes such as Sin3A-HDAC1/HDAC2, EZH2/SUZ12, LSD1-CoREST, and Ajuba-PRMT5 have been identified (6–10), but how these complexes are orchestrated and

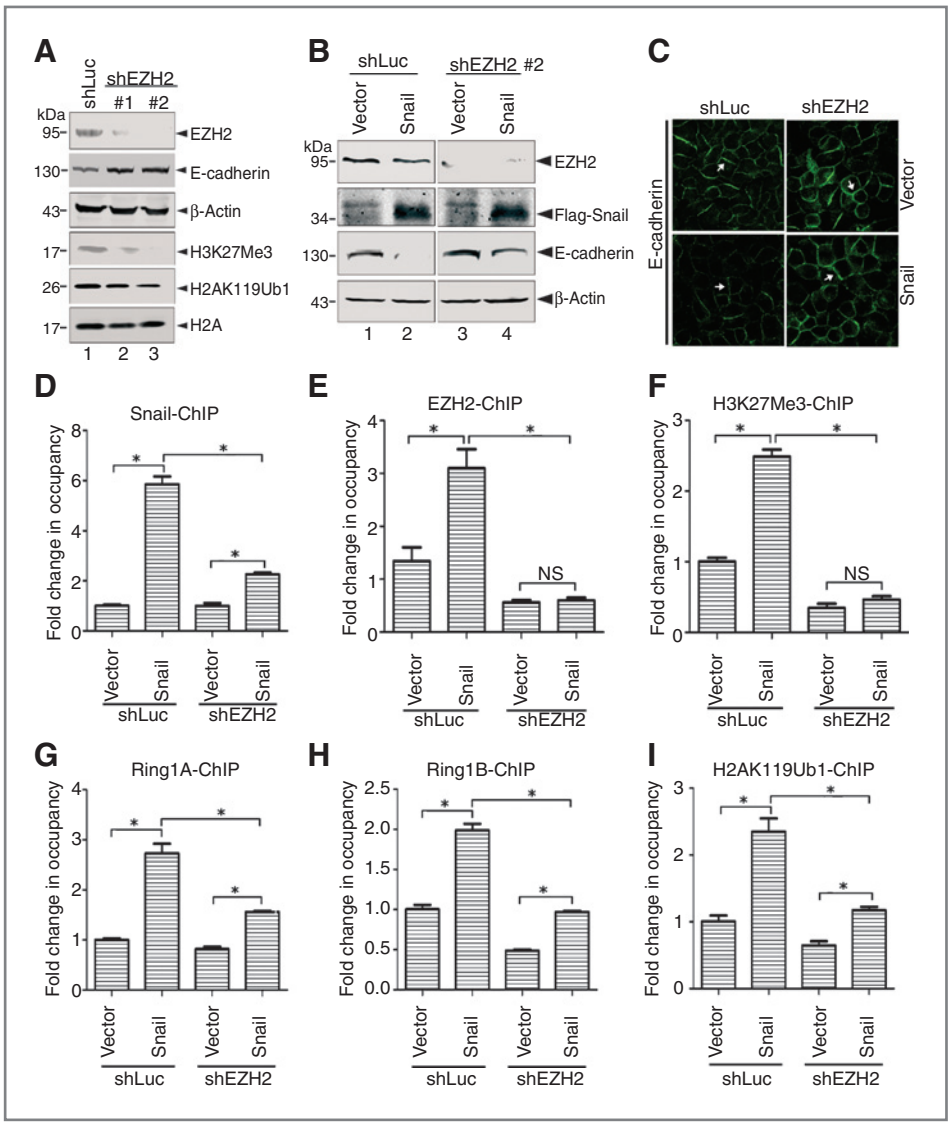

Figure 6. Depletion of EZH2 reduces Snail/Ring1A/B multiprotein complex binding at the E-cadherin promoter locus. A, depletion of EZH2 in PanC1 cells by specific shRNAs results in global decrease of H3K27Me3 and elevated E-cadherin. Western blot showed EZH2 is effectively depleted in PanC1 cells, global H3K27Me3 is reduced, and H2AK119Ub1 is partially decreased in these cells. #1 and #2, different cell pools generated from EZH2 knocking down by two specific shRNAs. B, Snail-Flag is stably expressed in PanC1-shLuc or PanC1-shEZH2 (pool #2 in A) cells. Expression of Snail effectively represses E-cadherin in PanC1-shLuc cells, but only weakly represses E-cadherin in PanC1-shEZH2 cells. Whole-cell lysate (100 μg) was loaded and β-actin was used as loading control. C, immunofluorescence staining (green) of E-cadherin. E-cadherin is increased in PanC1-shEZH2 cells. Expression of Snail markedly decreases E-cadherin in PanC1-shLuc cells, but only slightly reduces E-cadherin in PanC1-shEZH2 cells. Arrows, E-cadherin localization on cell boundary. Original magnification, ×63. D–I, ChIP DNA fragments from PanC1-shLuc or PanC1-shEZH2 cells overexpressing Snail were analyzed by qPCR assays. PCR amplifications were normalized to their inputs respectively. Enrichment of Snail enhances the binding of EZH2 (E), Ring1A (G), and Ring1B (H) to E-cadherin promoter, increases H3K27Me3 (F) and H2AK119Ub1 (I) levels, whereas depletion of EZH2 decreases binding of Snail, Ring1A, and Ring1B to E-cadherin promoter, abolishes H3K27Me3, and decreases H2K119Ub1. The units on the *y* axis are arbitrary. Data, mean ± SD from three independent experiments. *, $P < 0.05$. NS, without statistical significance, $P > 0.05$.

assembled at the target chromatin is largely unknown. Here, we provided an example that how the SNAG-associated EZH2 and the zinc finger–associated Ring1A/B complexes cooperate at the target chromatin to regulate gene expression. EZH2 can trimethylate H3 at K27, which is required for Snail-dependent E-cadherin repression during cancer progression (22, 23). Consistently, we showed that depletion of EZH2 in PanC1 cells results in loss of Snail-mediated repression on E-cadherin, and a concomitant decrease in the binding of Snail, Ring1A, and Ring1B to the E-cadherin promoter, indicating that EZH2 functions upstream of Ring1A/B and promotes Snail to recruit Ring1A and Ring1B more efficiently at the E-cadherin promoter locus. This observation is well supported by previous findings that H3K27me3 is thought to involve in the initiation of gene repression and serve as a docking site for the recruitment of

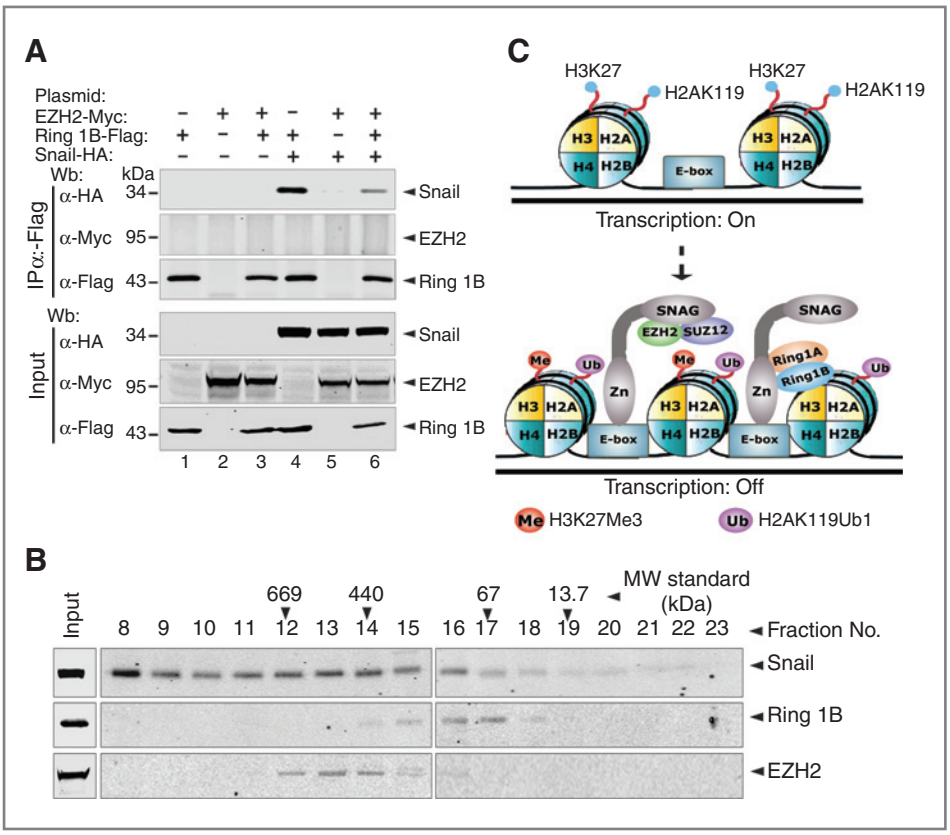

Figure 7. EZH2 and Ring1B form two distinct complexes with Snail. A, Snail, Ring1B, and EZH2 do not coexist in the same complex. Lysates from HEK-293T cells transfected with plasmids of HA-Snail, Flag-Ring1B, and Myc-EZH2 were incubated with anti-Flag M2 beads, and coeluted proteins were analyzed by HA and Myc antibodies. Co-IP showed that Ring1B does not interact with EZH2. B, size exclusion fractionation of whole-cell extracts prepared from HEK-293T cells expressing HA-Snail, Flag-Ring1B, and Myc-EZH2 was performed on Superose 6 10/300 GL column, and fractions of proteins were collected at 1 mL per fraction at 0.35 mL per minute in the cold room. Western bolt analysis of the proteins from each fraction with HA, Myc, or Flag antibodies showed that Ring1B and EZH2 form two distinct complexes with Snail. C, model for Snail-mediated repression. Snail recruits Ring1A and Ring1B via the carboxyl tandem zinc fingers to monoubiquitylate H2AK119, whereas Snail recruits EZH2 via SNAG domain to trimethylate H3K27, which may enhance Snail/Ring1A/B complex recruitment. Snail/Ring1A/B and Snail/EZH2 complexes may bind to different E-boxes or dynamically bind to the same E-boxes at the target promoters.

PRC1 proteins (26–29), whereby PRC1 is recruited to maintain the stable repression of genes. However, one remaining question is that how exactly Snail/EZH2 and Snail/Ring1A/B complexes are recruited to the target promoters because these two complexes do not coexist. Notably, we also observed that Snail occupancy at the target chromatin is affected by Ring1A and Ring1B, suggesting Ring1A and Ring1B may stabilize the Snail complex at the target chromatin. Although we showed that Snail binds to the RING domains of Ring1A and Ring1B, Snail is not a substrate of Ring1A and Ring1B (unpublished data). The detailed mechanism by which Ring1A and Ring1B regulate chromatin binding of Snail is not clear and needs to explore further. Moreover, we observed that depletion of Ring1A and Ring1B in PanC1 cells abolishes Ring1A and Ring1B binding and H2AK119Ub1 modification at the E-cadherin promoter, but expression of Snail in PanC1 cells still represses E-cadherin and induces cell migration in these cells (Fig. 4), indicating Snail may recruit additional repression complexes to repress transcription even in the absence of Ring1A and Ring1B. We postulate that the binding of Snail-associated complexes to Snail target chromatin may be dynamically and temporally regulated during the development and EMT events, which need to be dissected further in a system so that expression of Snail is finely controlled.

In summary, these observations demonstrate that the repressive activity of Snail is contributed not only by SNAG domain but also by the carboxyl zinc fingers; the histone ubiquitination also plays an important role in Snail-mediated transcriptional repression. Together, this sheds new light on the epigenetic machinery in Snail-induced EMT and cancer metastasis. These results provide unequivocal evidence to support that Ring1A and Ring1B could be explored for new therapeutics and that H2AK119Ub1 could be a potential biomarker of pancreatic cancer prognosis.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: J. Chen, H. Xu, F.J. Rauscher, III, C. Peng, Z. Hou
Development of methodology: J. Chen, H. Xu, Z. Hou

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J. Chen, H. Xu, X. Zou, J. Wang, A. Zhou, E.Y. Chin, Z. Hou
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J. Chen, H. Xu, B. ShenZ. Hou
Writing, review, and/or revision of the manuscript: J. Chen, H. Xu, F.J. Rauscher, III, C. Peng, Z. Hou
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H. Xu, Y. Zhu, H. Chen, C. Peng, Z. Hou
Study supervision: JH. Xu, X. DengC. Peng, Z. Hou

Acknowledgments

The authors thank Guiping Li, Han Qin, Xuemei Yang, and Minmin Shi for technical help on histology and imaging and Yanyan Song for statistical analysis.

Grant Support

This work was supported by the Ministry of Sciences and Technology of China (2013CB 910900), the National Science Foundation of China (grant Nos. 81172028 and 81372309), and the Shanghai Committee of Science and Technology (13JC1401302). Z. Hou is a Pujiang Scholar of Shanghai metropolitan. F.J. Rauscher is supported by NIH grants CA129833, CA010815, CA163761, CA167151, DOD-BCRP W81XWH-11-1-0494, The Samuel Waxman Cancer Research Foundation, Susan G. Komen for the Cure, The Noreen O'Neill Foundation for Melanoma Research, and the Ovarian Cancer Research Foundation.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 21, 2014; revised March 24, 2014; accepted May 2, 2014; published OnlineFirst June 5, 2014.

References

1. Nieto MA. Epithelial-mesenchymal transitions in development and disease: old views and new perspectives. Int J Dev Biol 2009;53:1541–7.
2. Moreno-Bueno G, Portillo F, Cano A. Transcriptional regulation of cell polarity in EMT and cancer. Oncogene 2008;27:6958–69.
3. Moody SE, Perez D, Pan T-c, Sarkisian CJ, Portocarrero CP, Sterner CJ, et al. The transcriptional repressor Snail promotes mammary tumor recurrence. Cancer cell 2005;8:197–209.
4. Cano A, Perez-Moreno MA, Rodrigo I, Locascio A, Blanco MJ, del Barrio MG, et al. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat Cell Biol 2000;2:76–83.
5. Batlle E, Sancho E, Franci C, Domínguez D, Monfar M, Baulida J, et al. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat Cell Biol 2000;2:84–9.
6. Lin Y, Wu Y, Li J, Dong C, Ye X, Chi Y-I, et al. The SNAG domain of Snail1 functions as a molecular hook for recruiting lysine-specific demethylase 1. EMBO J 2010;29:1803–16.
7. Hector Peinado, Esteban Ballestar, Manel Esteller, Cano aA. Snail mediates E-cadherin repression by the recruitment of the Sin3A-histone deacethylase1–2 complex. Mol Cell Biol 2004;24:306–19.
8. Hou Z, Peng H, Ayyanathan K, Yan KP, Langer EM, Longmore GD, et al. The LIM protein AJUBA recruits protein arginine methyltransferase 5 to mediate SNAIL-dependent transcriptional repression. Mol Cell Biol 2008;28:3198–207.
9. Dong C, Wu Y, Yao J, Wang Y, Yu Y, Rychahou PG, et al. G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. J clin invest 2012;122:1469–86.
10. Tong Z, Cai M, Wang X, Kong L, Mai S, Liu Y, et al. EZH2 supports nasopharyngeal carcinoma cell aggressiveness by forming a co-repressor complex with HDAC1/HDAC2 and Snail to inhibit E-cadherin. Oncogene 2011;31:583–94.
11. Vidal M. Role of polycomb proteins Ring1A and Ring1B in the epigenetic regulation of gene expression. Int J Dev Biol 2009;53:355–70.
12. Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, et al. Role of histone H2A ubiquitination in Polycomb silencing. Nature 2004;431:873–8.
13. Cao R, Tsukada Y, Zhang Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol cell 2005;20:845–54.
14. de Napoles M, Mermoud JE, Wakao R, Tang YA, Endoh M, Appanah R, et al. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev cell 2004; 7:663–76.
15. Voncken JW, Roelen BA, Roefs M, de Vries S, Verhoeven E, Marino S, et al. Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc Natl Acad Sci U S A 2003;100:2468–73.
16. Alberga A, Boulay J-L, Kempe E, Dennefeld C, Haenlin M. The snail gene required for mesoderm formation in Drosophila is expressed dynamically in derivatives of all three germ layers. Development 1991;111:983–92
17. Carver EA, Jiang R, Lan Y, Oram KF, Gridley T. The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition. Mol Cell Biol 2001;21:8184–8.
18. Martinez-Romero C, Rooman I, Skoudy A, Guerra C, Molero X, Gonzalez A, et al. The epigenetic regulators Bmi1 and Ring1B are differentially regulated in pancreatitis and pancreatic ductal adenocarcinoma. J pathol 2009;219:205–13.
19. Remmele W, Stegner HE. Recommendation for uniform definition of an immunoreactive score (IRS) for immunohistochemical estrogen receptor detection (ER-ICA) in breast cancer tissue. Der Pathologe 1987; 8:138–40.
20. Maacke H, Opitz S, Jost K, Hamdorf W, Henning W, Krüger S, et al. Over-expression of wild-type Rad51 correlates with histological grading of invasive ductal breast cancer. Int J Cancer 2000;88: 907–13.
21. Dancey CP, Reidy J. Estatística sem matemática para psicologia: usando SPSS para Windows. Brazil: Artmed/Bookman; 2006.
22. Herranz N, Pasini D, Diaz VM, Franci C, Gutierrez A, Dave N, et al. Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. Mol Cell Biol 2008;28:4772–81.
23. Cao Q, Yu J, Dhanasekaran SM, Kim JH, Mani R-S, Tomlins SA, et al. Repression of E-cadherin by the polycomb group protein EZH2 in cancer. Oncogene 2008;27:7274–84.
24. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics. CA Cancer J Clin 2010;60:277–300.
25. Li D, Xie K, Wolff R, Abbruzzese JL. Pancreatic cancer. Lancet 2004; 363:1049–57.
26. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. Drosophila enhancer of Zeste/ESC complexes have a Histone H3 methyltransferase activity that marks chromosomal polycomb sites. Cell 2002; 111:185–96.
27. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of histone H3 lysine 27 methylation in polycomb-group silencing. Science 2002;298:1039–43.
28. Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev 2003;17:1870–81.
29. Min J, Zhang Y, Xu R-M. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev 2003;17:1823–8.