

**A TALE OF TWO EPIGENETIC PROTEIN PAIRS:
CHARACTERIZING THE SIN3A-TET1 COMPLEX
AND THE DEACETYLATION ACTIVITY OF
HDAC1/2.**

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

Aditya Chandru M.Sc. (Sheffield Hallam University)
Department of Biochemistry/Molecular and Cell Biology
University of Leicester

November 2017

Abstract

A Tale Of Two Epigenetic Protein Pairs: Characterizing the Sin3a-TET1 Complex and the Decrotonylation Activity of HDAC1/2.

Characterizing the Sin3a-TET1 Complex:

TET1 permits active DNA demethylation, promoting a transcriptionally permissible environment by preventing aberrant methylation. Contrastingly, Sin3a is the scaffold-protein central to the eponymous co-repressor complex. Histone deacetylases 1 and 2 (HDAC1/2) form the enzymatic core of this protein-multiplex where they function to increase DNA:histone cohesion, thereby repressing gene expression. Despite these divergent activities, TET1 was found to bind Sin3a and recruit it to target genes. Mapping this interaction was integral to characterizing the TET1:Sin3a complex. An evolutionarily conserved 892-VAIEALTQLS-901 region in TET1 was identified. This α -helix was confirmed to bind to Sin3a's PAH1 domain. Complex formation was found to be equimolar and NMR illustrated the association to the resolution of individual residues. The TET1 Sin3a interaction domain (SID)'s bulky hydrophobic residues, and smaller adjacent alanines, bound to 18 PAH1 amino acids in a manner that resembles a key fitting a lock. The SAP25-SID interacts with these same PAH1 residues and binds in an identical orientation to TET1. While it initially seems curious that Sin3a and TET1 should bind, TET1 operates as a net transcriptional repressor in cells. This activity appears independent of TET1's enzymatic function and instead likely relies on HDAC recruitment through the Sin3a co-repressor complex. This was confirmed in a reporter assay where Gal4 TET1-SID fusions repressed gene expression in a manner that was proportional to the fusion's ability to bind Sin3a.

Characterizing the Decrotonylation Activity of HDAC1/2:

In addition to Lys-Acetylation, a constellation of different Lys-Acylation have recently been discovered, including crotonylation. The ubiquity and nuclear localization of HDAC1/2 suggested their involvement in decrotonylation. *HDAC1/2* DKO ESCs confirmed HDAC1/2's absence led to hypercrotonylation of the core histones. A novel decrotonylase assay established HDAC1/2 as potent lysine decrotonylases and that this activity could be recruited to lysine residues, by LSD1, via the CoREST complex.

Acknowledgements

“It was the best of times,
it was the worst of times...”¹

I submit to you the results attained over the last four years for the assessment of my PhD. Presenting the data contained within this thesis as my own, however, belies a greater truth. This is not only the product of individual effort, but rather the distillate of many thousands of small acts of kindness visited upon me by family, friends and co-workers; though, through the strange crucible of time, a kinship formed that made the distinguishing between these groups rather difficult.

Firstly, I would like to acknowledge my family, especially my parents, aunt, uncle and Dr. Pooyeh Lopu Farahmand – for their unyielding support. I thank my Ph.D. and M.Sc. supervisors Prof. Shaun Cowley, Dr. Benjamin Abell, Dr. David Smith and Dr. Adrian Hall for their camaraderie and for affording me a series of incredible opportunities which culminated in the writing of this document, which itself serves as a prologue of sorts.

I am thankful to Chris Busch and Dr. Alina Jaworska-Sobiesiak for their initial letters of recommendation that permitted me to embark on this voyage. A path that presented me, midway through the journey of life, some of my dearest of friends, including: Dr. Li Chen, Dr. Nico Portolano, Dr. Miran Rada, Vinod Reddy, Dr. Andy Jobbins, Dr. Sarder Hasib, Dr. Waqar Ahmed, Nazia Tabassum, Sachin Ramesh and Caryn Yeoh².

A particularly large debt is owed to my labmates Marcel Hildebrand, Sandra Germano, Chih-Hsiang Huang, Helen Calder, Dr. Grace Hodson, Dr. Claire Barnes, Dr. Gregg Hudson, Dr. Emma Kelsall, Dipti Vashi, Christian Lucas, Dr. Diana Marouco, Justyna Skorezeweska, Dr. Patrick McIntyre, Dr. Elvis Shilton, Anjali Pillai, Dr. Lyndsey Wright and Dr. Akang Basse, all of whom patiently taught me a multitude of

¹ DICKENS, C. 2004. *A Tale of Two Cities*. Edited by Boss, J. and Widger, D. The Project Gutenberg. Available at: <http://www.gutenberg.org/files/98/98-h/98-h.htm>.

² ALIGHIERI, D. 1997. *The Divine Comedy (Longfellow's Translation, Hell)*. Edited by Longfellow, H.W. and McCarthy, D. The Project Gutenberg. Available at: <https://www.gutenberg.org/files/1001/1001-h/1001-h.htm>.

techniques. Dr. Richard Kelly's encyclopedic knowledge, and help interpreting heat-maps, made many of the experiments in this thesis possible. Dominique Pope was the remarkably bright summer student who assisted me with the final assay. Dr. Hakan Bagci, Dr. Darren John Fitzpatrick and Dr. Adrian Bracken's research cast a light that illuminated our path. The assistance of Dr. Ralf Schmid, Dr. Jin-Li Luo, Dr. Nicholas Sylvius and especially Dr. Spencer Gibson enabled me to learn programming. I thank Dr. Peter Watson, Prof. John Schwabe, Dr. Olga Makarova and for their assistance with my project and for their depth of insight.

I would like to express my eternal gratitude to Mr. Peter. J. Conboy, Dr. Srinivas, Dr. Faraz Anjum, Dr. Julie L. Andrews, Dr. Bindy Sahota, Ali Naqvi, Alex, Zoya, Prof. Sue Shackleton and Prof. Ian Epperon for aiding me when the clear path was overgrown, through the darkest of forests².

I am fortunate to have met Dr. Salvador Macip, Mr. Konstantinos Polymeros, and Dr. Jesvin Samuel as their enthusiasm for science and literature extended well beyond the realm of the laboratory; though, perhaps more mundanely, I am also thankful for their assistance.

No thesis would be possible without the tireless efforts of the University of Leicester Library Content Services Team, including: Jonathan Bennett, Julie Lynch, Andrew McGee, Helen Newcombe, Joan Palmer, Lynne Shelton, Theresa Smith and Ushma Sharma. I am similarly grateful to Morag Clarke and the University of Leicester Printing Services staff.

The contributions of my friends and co-workers Chris Gott, Dr. Kayoko Tanaka, Dr. Vincenzo Marra, Barrister and Solicitor Ken Hughes, Annabel Cowley and the calfleys, Anna Maria Hodara, Maria Teresa Ortiz Jr., Maria Teresa Ortiz, José Ortiz, Mrs. Angastiniotis, Mrs. Annabel Henriques, Mr. David Hansen, Mr. Timothy Mills, David Brown, Dr. Cyril Dominguez, Dr. Khushwant Sidhu, Dr. Carolyn Jones, Dr. Alia Alfi, Dr. Alaa Khadidos, Dr. Jaswir Basran, Dr. Andrey Revyakin, Dr. Paul Smith, Dr. Amit Sharma, Andrew Dean, Dr. Luke Wendt, Dr. Matthew Adamson, Dr. Adam Lightfoot,

Dr. Charlie Milano, Dr. Rajkumar Reddy, Dr. Fedir Bokhovchuk, Barbara Birch, Amanda Fisher, Kudou Sallah, Kerry Warner, Dr. Jon Tillotson-Roberts, Brankica Besevic, Tom Silver, Bill Hassan, Michael Lee, Andy Bradshaw, Andy Smith, Romilla Parmer, Terry Bailey, Mick Phillips, Wayne Hall, Paula Monaher, Dr. Ben Goult, Jacquie Greenwood, Maureen Wragg, Prof. Martin Dyer, Prof. Don Jones, Dr. Zsolt Szijgyarto, Jasbir Summan and Dr. Alice Haworth aided my personal and professional development. I thank Dr. Lorenzo De la Rica for his insight into L1 elements, and for making time to correspond to a student.

I am grateful to my friends Zarah Qusairy, Nini Karen, Nima Farahmand, Ladan Abassi, Behrouz Farahmand, Faty Daghighi, Dr. Hamideh Parizi, Sarder Hafeez, Nazma Khatun, Wasif Ahad, Elizabeth Razaque, Claire Lee, Michelle Huang, Ling Huang, Kinuyo Ono, Rie Kakui, Saki Fukasawa and Shafiul Rony for their kindness. I especially thank you, dear reader, for making this work relevant.

On an aside, I would like to take a moment to commend Pepceuticals, the college IT services team and the doctoral college for their unparalleled efficiency.

“So it goes...”³

³ VONNEGUT, K. 1991, *Slaughterhouse-Five*. Dell, 22.

Contents

Table of Titles

Table of Titles.....	8
Table of Contents	8
Table of Figures.....	13
Table of Tables.....	18
List of Abbreviations.....	19

Table of Contents

Abstract	1
Acknowledgements.....	3
Contents.....	7
Chapter 1. Introduction	23
1.1. Epigenetics.....	24
1.2. Preface	25
1.3. DNA, Histones and Chromatin	26
1.4. Histone Methylation	30
1.5. Histone Acetylation	31
1.5.1. An Expanded Set of Epigenetic Instructions.....	32
1.6. Histone Acetylation Writers.....	34
1.7. Histone Acetylation Readers.....	35
1.8. Histone Acetylation Erasers: HDACs.....	36
1.8.1. HDAC1/2 Complexes	38
1.8.2. The Sin3a Co-Repressor Complex	39
1.8.3. NuRD, NODE and CoREST	40
1.8.4. Non-Histone Acetylation Targets.....	41
1.9. Alternative Acylations: Crotonylation.....	42
1.9.1. Crotonylation Writers	43
1.9.2. Crotonylation Readers and Function	46

1.9.3. Lysine Acylation Erasers	47
1.10. Sin3a and the TET1 Binding Site.....	50
1.10.1. PAH Domains.....	51
1.10.2. PAH4 and the HCR.....	55
1.11. DNA Methylation	56
1.12. DNA Methylation Writers.....	57
1.13. DNA Methylation Readers.....	58
1.14. DNA Methylation Erasers: the TET Proteins	59
1.14.1. DNA Demethylation.....	61
1.15. Crosstalk Between DNA and Histone Epigenetic Marks and their Writers ..	63
1.16. TET1's Biological Relevance.....	64
1.16.1. An Expanded Repressor Complex.....	66
1.17. Aims	67
Chapter 2. Materials and Methods.....	69
2.1. Agarose Gel-Electrophoresis.....	70
2.2. Cell Culture	70
2.2.1. Cryostorage	71
2.2.2. Passaging.....	71
2.3. Cloning.....	72
2.4. Co-Immunoprecipitation Using IgG antibody and Sepharose Beads	72
2.5. Western Blot.....	74
2.5.1. Whole Cell Extract Preparation.....	74
2.5.2. SDS-PAGE.....	74
2.5.3. Gel to Membrane Transfer (Blotting).....	75
2.5.4. Western Blot Immunizing and Imaging.....	75
2.6. Crystallography Procedures	75
2.7. <i>E. Coli</i> Transformation and Protein Production.....	76
2.8. Gal4 VP16 Reporter Assay	76

2.8.1. β -gal Activity Measurement.....	77
2.8.2. Luciferase Activity Measurement	77
2.9. GST-Pulldown and IVTT.....	78
2.10. HDAC Deacylation Assay.....	79
2.11. Rescue of <i>TET1/2</i> knockout cells with PiggyBAC TET1.....	81
2.11.1. Generating the TET1 Rescue Cells Through PiggyBAC Integration of Flag-TET1 Constructs in <i>TET1/2</i> DKO Cells.....	81
2.12. Protein Purification for Gel-Filtration and NMR.....	82
2.13. Polymerase Chain Reaction (PCR).....	85
2.14. Quantitative Reverse Transcription PCR (qRT-PCR)	86
2.14.1. RNA Extraction	86
2.14.2. Reverse Transcription	86
2.14.3. Gel Extraction and PCR Cleanup and Dilution Series	86
2.14.4. Quantitative Reverse Transcription PCR Reaction	86
2.14.5. qRT-PCR Calculation of Fold Change	89
Chapter 3. Results: Sin3a-TET1.....	90
3.1. A Conserved Sin3a Interaction Domain.....	91
3.1.1. An α -Helical Conserved Domain	94
3.1.2. The α -Helix Appears to be a SID	98
3.1.3. The SID is Correlated with Binding	99
3.1.4. The TET1-SID is Necessary for Sin3a Binding.....	100
3.1.5. The SID is Sufficient for Sin3a Binding.....	102
3.2. Sin3a's TET1 Interaction Domain.....	103
3.2.1. Sin3a's TET1 Binding Region Occurs Within PAH1-3	104
3.2.2. TET1 Weakly Binds Sin3a 1-205 (PAH1), but Strongly Binds Sin3a 1-479 (PAH1-2)	106
3.3. PAH1 is Required for TET1 Binding	107
3.4. PAH1 is Necessary for TET1 Binding	108
3.5. PAH1 is Sufficient for TET1 Binding.....	109

3.6. A Single PAH1 Interaction Site on TET1.....	110
3.7. Characterizing the Residues of the TET1-SID.....	112
3.8. Characterizing Sin3a:TET1 Binding.....	114
3.9. Identifying the Importance of Individual Residues within the SID.....	122
3.10. Characterizing the Full-Length Sin3a:TET1 Interaction.....	135
3.11. Structural Characterization of the Sin3a:TET1 Interaction.....	140
3.11.1. Protein Crystallography Screens.....	140
3.11.2. An NMR Informed Model of the Sin3a:TET1 Interaction.....	142
3.12. Characterizing the Biological Significance of the TET1:Sin3a Interaction .	149
3.12.1. qRT-PCR Reveals <i>TET1</i> ^{WT} Rescue Increases IDAX and LIF Expression.	150
3.13. Sin3a:TET1 Complex Results.....	152
Chapter 4. Results: HDAC1/2 and Histone Decrotonylation.....	154
4.1. Metabolic Effects on Crotonylation.....	155
4.2. Measuring Decrotonylase Activity.....	156
4.3. Measuring Cellular Decrotonylation Activity.....	157
4.4. Are Classical HDACs Involved in Decrotonylation?.....	159
4.5. Are HDAC1/2 Involved in Decrotonylation?.....	160
4.6. HDAC1/2 Decrotonylation Results.....	167
Chapter 5. Discussion: Sin3a-TET1.....	169
5.1. TET1: The Great Epigenetic Loom.....	170
5.2. A Thousand Strings.....	170
5.3. Identification and Characterizing the TET1/Sin3a Interaction.....	171
5.4. Viewing Sin3a:TET1 from a Structural Perspective.....	172
5.5. Relative Strength of the Sin3a:TET1 Interaction.....	175
5.6. A Dissenting View.....	177
5.7. Biological Implications of the Sin3a:TET1 Complex.....	179
5.7.1. TET1 ^{ΔCXXC} 's Biological Consequences.....	181
5.8. Sin3a:TET1 Complex Conclusion.....	181

Chapter 6. Discussion: HDAC1/2's Role in KDCr.....	183
6.1. Crotonylation.....	184
6.2. Crotonylation and Metabolism	184
6.3. Corroboration of HDAC1/2's Role in Decrotonylation	185
6.4. Crotonylation on Non-Histone Proteins	186
6.5. The Regulatory Importance of Cr in Non-Histone Proteins.....	187
6.6. The Role of HDAC1/2 in Non-Histone Protein Decrotonylation	187
6.7. TET1's Role in Decrotonylation	188
Chapter 7. Appendix.....	189
Chapter 8. Bibliography	203

Table of Figures

Figure 1.1: Metabolism and the Epigenetic Mark Precursors.	26
Figure 1.2: Cytosine (Left) and 5-Methylcytosine (Right).	27
Figure 1.3: Nucleosome Post-translational Modifications.	29
Figure 1.4: Histone Modifications, Their Locations and Effect on Gene Expression (Berger, 2007).	30
Figure 1.5: A Diagrammatic List of Lysine Acylations.	33
Figure 1.6: Histone Lysine Acylation (Acetylation) Results in the Loss of Lysine’s Positive Charge.	34
Figure 1.7: The HDAC Family of Enzymes with the Relative Positions of Their Enzymatic Domains (Seto and Yoshida, 2014).	38
Figure 1.8: The HDAC1/2 Complexes: SIN3 Co-Repressor, NuRD and CoREST Complexes (Kelly and Cowley, 2013).	39
Figure 1.9: The Relative Acyl-CoA Concentrations Dictate the Activity of EP300.	44
Figure 1.10: Human H4 Lysine Modifications (Xie et al., 2012, Dai et al., 2014, Camporeale et al., 2004, Wiśniewski et al., 2008, Goudarzi et al., 2016, Arnaudo and Garcia, 2013).	45
Figure 1.11: The Acetylation and Alternative Acylation Model.	46
Figure 1.12: A List of Lysine Acylations and Their Erasers (Choudhary et al., 2014). .	48
Figure 1.13: Localization of the HDAC Enzymes (Heidemann, 2015).	49
Figure 1.14: Comparative Expression of HDACs in ESCs Obtained from Microarray (Jamaladdin et al., 2014).	49
Figure 1.15: A Schematic of SIN3 Serving as a Scaffold-Protein in its Role in the Sin3a HDAC1/2 Repressor Complex (Grzenda et al., 2009).	50
Figure 1.16: Sin3a Has Four PAH Domains, Which Are Each Comprised of Four Alpha Helices.	52
Figure 1.17: Sin3a Binding Motifs.	53
Figure 1.18: Possible Conserved Electrostatic Interaction.	54
Figure 1.19: SAP30’s Interaction with Sin3a’s PAH3.	55
Figure 1.20: Canonical DNMT DNA Methylation.	58
Figure 1.21: The Primary Sequence of the TET Enzymes.	60

Figure 1.22: TET Expression During Development and the Corresponding DNA Methylation/Hydroxymethylation Levels (Tan and Shi, 2012).....	61
Figure 1.23: TET Mediated DNA Demethylation by Iterative Oxidation.	62
Figure 1.24: TET1 is a Net Transcriptional Repressor.	65
Figure 1.25: The Possible Configurations of OGT Complexes (Vella et al., 2013).....	66
Figure 1.26: The Fam60:TET1:OGT:Sin3a Complex (Streubel et al., 2017).....	67
Figure 2.1: Measuring Luminescence Indirectly Permits the Quantification of Relative Luciferase Activity.	78
Figure 2.2: BOC-Cr-Lys-AMC Substrate.	80
Figure 2.3: A PiggyBac System Was Used to Rescue the B8 Cells with TET1.	82
Figure 2.4: Imidazole Gradient of the Histidine Column Used to Extract Protein Fragments for NMR and Gel Filtration.	83
Figure 2.5: His-Sin3a 115-212 Elution in Fractions A4-A7.	84
Figure 2.6: NaCl Gradient, in a Q Column, Purifies Proteins on the Basis of Charge. ...	85
Figure 2.7: Primer Efficiency Curve of GAPDH.	87
Figure 2.8: qRT-PCR Melt Curve of GAPDH.	88
Figure 2.9: qRT-PCR Melt Peak of GAPDH.	88
Figure 3.1: Sequence Alignment of TET1 and TET3.....	92
Figure 3.2: Phylogeny of the Putative TET1-SID (892-901) Sequence.	94
Figure 3.3: Jpred Secondary Structure Prediction of TET1.	98
Figure 3.4: Putative SID Appears to be an Amphipathic Helix.	99
Figure 3.5: SID Containing Fragments Interact with Sin3a.....	100
Figure 3.6: Putative SID Required for Sin3a Binding.	101
Figure 3.7: Putative SID Sufficient for Sin3a Binding.	103
Figure 3.8: Schematic Diagram of the Six Different Myc-tagged Sin3a Constructs Used in Pulldown Experiments (Laherty et al., 1997).	104
Figure 3.9: Sin3a's TET1 Interaction Domain is Between PAH1-3.....	105
Figure 3.10: Indeterminate Result Regarding Myc Sin3a 1-479 and 1-205.....	106
Figure 3.11: GST-TET1 SID Appears to Associate with PAH1.....	107
Figure 3.12: GST-TET1 868-927 Binds to PAH1.....	108
Figure 3.13: PAH1 is Necessary for TET1 Binding.....	109
Figure 3.14: PAH1 is Sufficient for TET1 Binding.....	110

Figure 3.15: Only TET1 746-951 is Capable of Binding to PAH1.	111
Figure 3.16: TET1 384-745 is Unstructured.	112
Figure 3.17: L89A, L900 and I894 Are Critical Residues in the TET1 SID.	114
Figure 3.18: SDS-PAGE of His-PAH1 Protein Production and Purification.	115
Figure 3.19: SDS-PAGE of the Final Purified Proteins.	116
Figure 3.20: Gel-Filtration Demonstrates the His-GB1-TET1-SID and His-PAH1 Interaction.	117
Figure 3.21: NMR Spectra Demonstrates the His-GB1-TET1-SID and His-PAH1 Interaction.	118
Figure 3.22: TET1 and Sin3a Bind in a 1:1 Ratio.	119
Figure 3.23: SAP25-SID Binds Nearly an Order of Magnitude More Sin3a Than the TET1-SID.	120
Figure 3.24: Gel-Filtration Demonstrates the His-GB1-SAP25-SID and His-PAH1 Interaction.	121
Figure 3.25: SAP25 Outcompetes TET1 for Sin3a Binding.	122
Figure 3.26: Gel-Filtration Demonstrates the L897 Residue is Critical for Sin3a Interaction.	123
Figure 3.27: NMR Spectra Demonstrates the L897A Mutant Nearly Prevents PAH1 Binding.	124
Figure 3.28: NMR Spectra Demonstrates the L897A Mutation Results in Weak His- PAH1 Binding.	125
Figure 3.29: SDS-PAGE Demonstrates That the I894A Mutation Weakens His-PAH1 Binding.	126
Figure 3.30: NMR Spectra Demonstrates the I894A Mutation Results in Mildly Weaker His-PAH1 Binding.	127
Figure 3.31: Gel-Filtration Demonstrates T898 is Unimportant in the TET1:Sin3a Interaction.	128
Figure 3.32: NMR Spectra Demonstrates the T898 Residue is Not Critical for His- PAH1 Binding.	129
Figure 3.33: Gel-Filtration Demonstrates T898 is Unlikely to be a Phosphorylation Site.	130

Figure 3.34: NMR Spectra Demonstrates T898E's Effects Are Nearly Identical to That of T898A and WT.	131
Figure 3.35: Gel-Filtration Demonstrates the L900 Residue is Important for Sin3a Interaction, Though Less Important Than L897.	132
Figure 3.36: NMR Spectra Demonstrates the L900A Mutation Results in Weak His-PAH1 Binding.	133
Figure 3.37: His-GB1-I894A and His-GB1-L900A Migrate Through the Column Identically to TET1-SID ^{WT}	134
Figure 3.38: L897 Appears to be at the Center of the Interaction with Sin3a. Both the Residues Preceding (I894A) and Succeeding (L900) L897 Impact PAH1 Binding.	135
Figure 3.39: B8 Cells Are TET1/2 DKO.	136
Figure 3.40: B8 Cells Do Not Express TET1 Protein.	137
Figure 3.41: B8 Cells Do Not Express TET2 Protein.	137
Figure 3.42: All PiggyBAC Rescue Cells Expressed Flag-TET1.	138
Figure 3.43: The L897A/L900A Mutations Prevent Full-Length TET1 From Interacting with Endogenous Sin3a.	139
Figure 3.44: The L897A/L900A Mutations Prevent Full-Length Endogenous Sin3a from Interacting with Full-Length TET1 in ESC.	140
Figure 3.45: Micro-crystals Formed by the His-PAH1 and TET1-SID ^{WT} Solution in Well B4.	142
Figure 3.46: The T898E Mutation Revealed the Angle the TET1-SID Alpha Helix Binds to PAH1.	143
Figure 3.47: NMR Model of Sin3a's PAH1 Binding TET1 SID. TET1 L897A Occupies the Center of PAH1's Hydrophobic Pocket, it is Flanked by I894 and L900.	144
Figure 3.48: TET1 S901 is Exposed to the Aqueous Environment.	145
Figure 3.49: The Sin3a PAH1 (Grey) Residues That Bind to TET1's SID (Blue).	148
Figure 3.50: Gal4-SID Fusions Interact with Sin3a Domains and Indirectly Recruit HDACs to Repress Reporter Expression.	149
Figure 3.51: TET1's Gene Repression is Dependent on Sin3a Recruitment.	150
Figure 3.52: Reintroduction of TET1 ^{WT} Appears to Increase IDAX Expression.	151
Figure 3.53: Reintroduction of TET1 ^{WT} Increases LIF Expression.	152

Figure 4.1: The Addition of Crotonate Increases Histone Cronylation, But Also Histone Acetylation.....	156
Figure 4.2: The BOC-Lys(Ac/Cr)-AMC Molecule Permits the Measurement of Deacylation.....	157
Figure 4.3: Mammalian Cell Lysate Contains a Lysine Decrotonylase.....	158
Figure 4.4: Classical HDACs Are Histone Decrotonylases.....	159
Figure 4.5: Histone Cronylation is Slower than Histone Acetylation.....	160
Figure 4.6: HDAC DKO Cells Contain Negligible HDAC1/2 3 Days After Induction (Jamaladdin et al., 2014).....	161
Figure 4.7: The Absence of HDAC1/2 Reduces Deacetylation and Decrotonylation Activity.....	162
Figure 4.8: Alternative HDACs Can Deacetylate Lysine Residues Upon HDAC1/2 Loss, Albeit Less Efficiently.....	163
Figure 4.9: HDAC1/2 Are Responsible for the Bulk of Decrotonylation Activity in ESCs.....	164
Figure 4.10: HDAC1/2 Control H3K18 Cronylation Twice as Tightly as Acetylation.....	165
Figure 4.11: HDAC1/2 DKO Has the Greatest Effect on H4 Cronylation.....	166
Figure 4.12: LSD1 Appears to Recruit Deacetylase Activity. HDAC2 Demonstrates Decrotonylase Activity.....	167
Figure 5.1: While TET1 Conforms to the Type I Motif, It Binds PAH1 in SAP25's Type II Orientation.....	174
Figure 5.2: SAP25 and MAD1's SID Contains an Extra Hydrophobic Residue That Likely Permits Stronger PAH1 Binding.....	176
Figure 5.3: TET1 May Have a Long Distance Electrostatic Interaction with Sin3a's PAH1, like MAD1 with Sin3b's PAH2.....	177
Figure 5.4: A Co-IP Revealing the C-terminal End of TET1 is Responsible for Sin3a Binding.....	177
Figure 5.5: GST-Pulldown Shows TET1-F3:1418-2138 Interacts with His-Sin3a.....	178
Figure 7.1: Vectors Provided by the University of Leicester Protex Service (Yang, 2017, Zhao et al., 2008).....	198

Table of Tables

Table 3.1: The His-PAH1 and TET1-SID^{WT} Solution Produced Crystals in well B4 at a pH 3.96, with 0.4M Sodium Acetate with 1M Ammonium Sulfate.	141
Table 7.1: Table of TET1-SID Evolutionary Conservation.....	190
Table 7.2: Table of Buffers and Solutions.	192
Table 7.3: Table of Primers and GBlocks.	195
Table 7.5: Table of Cloning Vectors.....	196
Table 7.6: Primers Used for qRT-PCR.	199
Table 7.7: Primers Used for RT-PCR.	199
Table 7.8: Primers Used to Sequence PiggyBAC TET1 Constructs.	200
Table 7.9: Table of Antibodies.....	202

List of Abbreviations

5caC	5-Carboxylcytosine
5fC	5-Formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
α -KG	α -ketoglutarate
β -gal	β -galactosidase
$\epsilon\pi$ i	Epi
Ac	Acetyl
ACL	ATP citrate lyase
ACSS2/AceCS1	Acyl-CoA synthetase 2
AML	Acute myeloid leukemia
BER	Base excision repair
BET	Bromodomain and extra-terminal
Bhb	β -hydroxybutyryl/3-hydroxybutyryl
Bio	Biotynyl
BME	β -mercaptoethanol
BOC	tert-butyloxycarbonyl protecting group
bp	Base pair
Bu	Butyryl
Co-IP	Co-immunoprecipitation
CoA	Coenzyme A
CoREST	Corepressor of RE-1 Silencing Transcription Factor
Cr	Crotonyl
CXXC	Zinc finger domain
DKO	Double Knockout
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DPBS	Dulbecco's phosphate-buffered saline

DSBH	Double-stranded β -helix
DTT	Dithiothreitol
EtBr	Ethidium Bromide
F1	Filial generation 1/Second generation
F2	Filial generation 2/Third generation
F3	Filial generation 3/Fourth generation
Fo	Formyl
GlcNAc	O-Linked N-Acetylglucosamine
Glu	Glutaryl
GNAT	Gcn5-related N-acetyltransferases
H1	Histone 1/Linker Histone
H2A	Histone 2A
H2B	Histone 2B
H3	Histone 3
H3K4	Histone 3 Lysine 4
H3R2	Histone 3 Arginine 2
H4	Histone 4
HAT	Histone Acetyltransferases
HBP	Hexosamine biosynthetic pathway
HCFC1	Host cell factor C1
HCR	Highly conserved region
HDAC	Histone Deacetylase/Lysine Deacetylase
HDAC1/2	Histone Deacetylase 1 and 2
Hib	2-hydroxyisobutyryl
HID	Histone deacetylase interaction domain
IGF2	Insulin-like growth factor 2
iPSCs	Induced Pluripotent Stem Cells
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITR	Inverted terminal repeat
IVTT	In vitro transcription and translation
JNET	Joint Network prediction of secondary structure

Jnethmm	Joint Network hidden Markov model
KDCr	Lysine decrotonylase
LBH	Panobinostat/LBH589
Lex-Gal-Luc	LexA response element–Gal4 response element–Luciferase
MBD	Methyl-CpG-binding domain
Me	Methyl
MLL	Mixed-lineage leukemia
mM	Millimolar
MTA	Metastasis associated protein
MYST	MOZ, Ybf2 (Sas3), Sas2 and Tip60-like acetyltransferase
NAD	Nicotinamide adenine dinucleotide
nm	nanometer
NMR	Nuclear magnetic resonance
NODE	Nanog/Oct4-associated deacetylase
NuRD	Nucleosome remodeling and deacetylase
OGT	O-Linked N-Acetylglucosamine Transferase
OKSM	OCT4, KLF4, SOX2, and C-MYC
ONPG	o-nitrophenyl- β -galactopyranoside
P-TEF	Positive transcription elongation factor
PAH	Paired amphipathic helix
PARP	Poly-ADP-ribose polymerases
PBS	Phosphate buffered saline
PDH	Pyruvate dehydrogenase
PHD	Plant homeodomain
Pol II	RNA polymerase II
Pr	Propionyl
PRC2	Polycomb repressive complex 2
PSIBLASTpssm	Position-Specific Iterative Basic Local Alignment Search Tool position-specific scoring matrix profile
PTM	Post-translational modification
RNA	Ribonucleic acid

RPD	Reduced potassium dependency
SAM	S-Adenosyl methionine
SAP30	SIN3A Associated Protein 30
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Super Elongation Complex
SID	Sin3 interaction domain
SIN	Switch independent
SIR2	Silent mating-type information regulation 2
Suc	Succinyl
SUMO	Small ubiquitin-like modifier
SUZ12	Suppressor of Zeste 12 homolog
SWI	Switch
t(10;11)(q22;q23)	Reciprocal translocation between region 22 on the long arm of chromosome 10 (a break between exons 8 and 12 of TET1) and region 23 on the long arm of chromosome 11 (a break between intron 6 and exon 11 of MLL)
TAF14	Transcription initiation factor TFIID subunit 14
TDG	Thymine DNA glycosylase
TET1/2	Ten-eleven translocation 1 and 2
TKO	Triple Knockout
TRK1	High affinity potassium transporter
TRK2	Low affinity potassium transporter
tRNA	Transfer ribonucleic acid
TSA	Trichostatin A
UHRF	Ubiquitin-like containing PHD and RING finger domain
WCE	Whole cell extract

Chapter 1. Introduction

1.1. Epigenetics

Epigenetics is a compound word comprised of “genetic” coupled with the Greek prefix ἐπί (epi), which means above/on/in addition. It refers to the layer of transcriptional regulation atop the genetic code. Like DNA, the epigenetic code can be transmitted to the next generation of cells or even organisms. Epigenetics supplements the evolutionary process, as it provides an avenue for rapid adaptation to environmental stressors. Epigenetic adaptation can even occur within a single generation. Just as the marks can be rapidly woven onto DNA and its associated histones, a mechanism exists that permits their brisk removal. An example of this dynamic process can be seen in the adaptation to a nutrient-scarce environment. In reprisal for their resistance in the Second World War, the Dutch were prevented from importing fuel and foodstuff (Roseboom et al., 2001, Banning, 1946). The demographic that was the most detrimentally affected, by the deprivation and unusually cold accompanying winter, were those who were prenatally exposed. Young men affected by the *Hunger Winter* in utero (F1 generation) and during the first half of pregnancy presented with a higher incidence of obesity (Ravelli et al., 1976). The affected children displayed phenotypes associated with insulin resistance; a phenomenon that has historical precedence (Ravelli et al., 1998, Barker, 1995). On an epigenetic level, the deprivation resulted in the differential methylation of the insulin-like growth factor 2 (*IGF2*) gene which moderated transcription factor access to the gene, thereby regulating its expression (Heijmans et al., 2008). *IGF2* stimulates the formation of beta cells, in the fetal pancreas, which produce insulin – the central regulator of glucose metabolism (Li et al., 2015). This differential methylation occurred in the gametes, which permitted the phenotype to be transmitted to the third generation (F2), who were found to have an increased incidence of obesity (Veenendaal et al., 2013). Mouse model studies reveal that an F3 group, that were in no way exposed to the original nutritional insult, still bear some hallmarks of the disease phenotype (Benyshek et al., 2006). It is, however, in this cohort that the changes begin to recede. Three generations in a nutrient rich environment appeared to permit phenotypic normalization (Benyshek et al., 2006). The epigenetic ripples that spread out a few generations from an initial disturbance must eventually dissipate. Epigenetic alterations are needed because the environment changes. When this

environmental stressor abates, the epigenetic adaptations are no longer required and may even become maladaptive. This requires a mechanism that permits the removal of the epigenetic marks.

1.2. Preface

This work will be a two-pronged story of separate, but functionally related, sets of proteins which are central to the removal of epigenetic marks. The first is TET1 and the latter are histone deacetylases 1 and 2 (HDAC1/2). TET1 is implicated in the removal of epigenetic marks in DNA and HDACs are involved in the removal of a subset of corresponding marks in histones. These two processes, like all epigenetic marks, are linked to each other. The most conspicuous example of this relationship exists in metabolism. The citric acid cycle makes use of acetyl-CoA, the precursor molecule used for histone acetylation, and this is eventually processed to produce α -ketoglutarate which is required for TET1's enzymatic activity. There are biochemical processes, like β -oxidation, tangential to the citric acid cycle that can convert acetyl-CoA into other acyl-CoAs which have relevance to histone modification. This metabolic web of epigenetic precursors is illustrated in Figure 1.1.

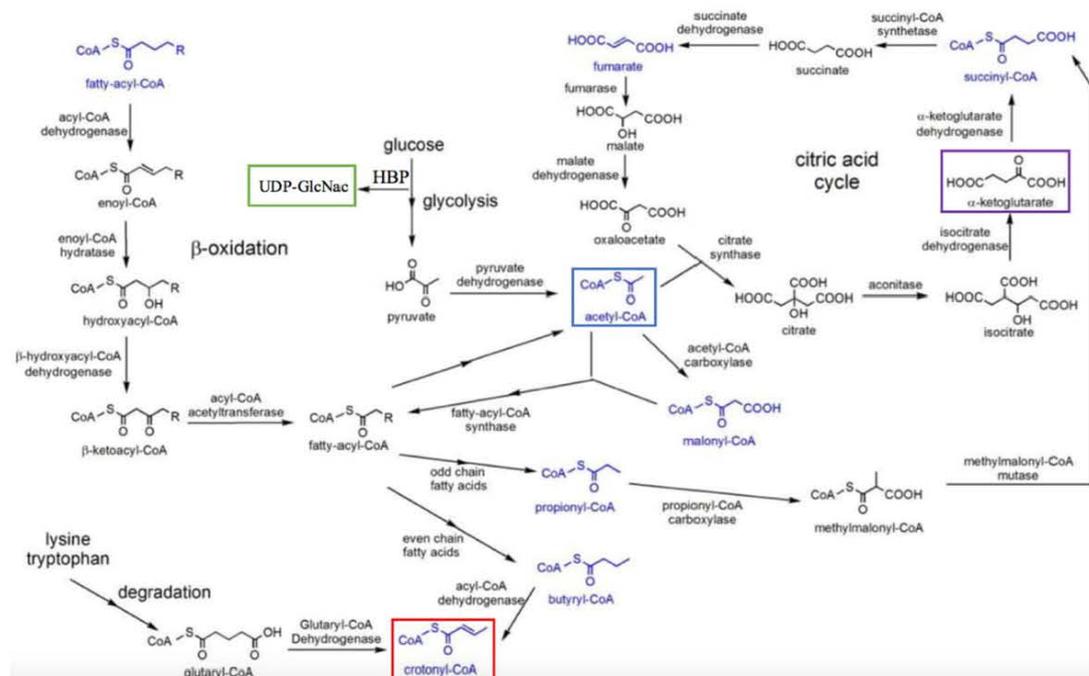


Figure 1.1: Metabolism and the Epigenetic Mark Precursors.

Demonstrates the link between DNA demethylation (α -ketoglutarate), the OGT pathway (UDP-GlcNAc), histone acetylation (acetyl-CoA) and histone crotonylation (crotonyl-CoA). (HBP: hexosamine biosynthetic pathway) (Lin et al., 2012).

While this relationship between TET1 and HDAC1/2 may appear to be somewhat meager from the outset, there is a far more solid linkage between the epigenetic enzymes. This takes material form in the scaffold-protein Sin3a, which is known to physically interact with both HDAC1/2 and TET1. The first part of this investigation will cover the Sin3a:TET1 complex formation and identify the domains, on both proteins, that facilitate this association. The second will cover histone crotonylation and the function of HDAC1/2 in removing this epigenetic mark, a process that has until this research remained elusive. This thesis will cover both these phenomena and will attempt to address their wider biological significance.

1.3. DNA, Histones and Chromatin

There are two main avenues of covalent epigenetic control, both of which rely on the placement of modifying marks onto either DNA or the tails of their entwined histones. The quintessential eukaryotic epigenetic mark is DNA methylation. This covalent mark

can be placed on two of the four DNA bases – adenine and cytosine (Wu et al., 2016). The more canonical cytosine methylation occurs on the fifth, carbon, atom of the residue's pyrimidine ring. This DNA modification is illustrated in Figure 1.2. These modifiable cytosine residues often occur within the genome as CpG dinucleotides, where a cytosine base is immediately followed by guanine and the two nucleotides are attached by phosphodiester bond. The dinucleotides are often enriched in ~1000bp stretches of DNA, termed CpG islands, that occur near transcription start sites/promoters and are resistant to methylation (Deaton and Bird, 2011). DNA methylation of promoters appears to be a repressive mark, whereas gene body methylation is curiously associated with active genes. It should also be noted that a large proportion of CpG islands are found in intergenic regions (Jeziorska et al., 2017). The DNA methylation mark will be covered in greater detail in chapter 1.11.

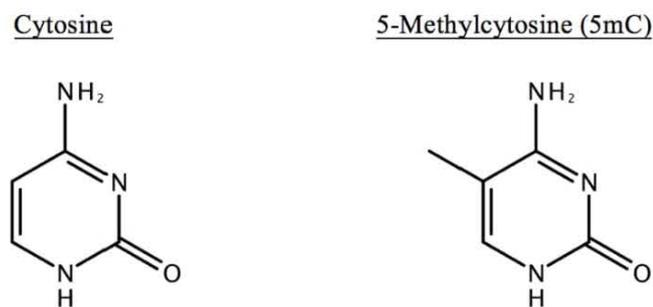


Figure 1.2: *Cytosine (Left) and 5-Methylcytosine (Right).*

The methylation is placed enzymatically on the fifth atom of the aromatic ring.

Transcription can also be negatively regulated through the compaction of DNA. This mechanism is necessary because the length of DNA far exceeds that of the cell by many orders of magnitude. In eukaryotes, this occurs through the function of a group of proteins called histones. These proteins spatially arrange DNA in a manner that makes it not only more compact, but retains enough accessibility to permit the functioning of the cell. The molecule that results from the histone:DNA interaction is called chromatin, which is present in two generalized forms: euchromatin and heterochromatin. Euchromatin is a less compacted arrangement of chromatin, allowing transcription factors access to the DNA, resulting in a more transcriptionally open environment. Contrariwise, heterochromatin is more condensed and is generally functionally inactive.

Constitutive heterochromatin encompasses genetic regions that are poorly expressed in all cells of a given organism – this includes the telomeres, centromeres and repetitive elements (Saksouk et al., 2015). Genes in facultative heterochromatin are also deficient in expression, though these inactive regions are not consistent within an organism. X chromosome inactivation is an example of facultative heterochromatin, as the chromosome that is inactivated varies between cells and can indeed be reactivated when inherited (Rego et al., 2008). The elementary unit of chromatin is called the nucleosome. This structure consists of a core octamer of histones wrapped in 146 base pairs (bp) of DNA (Luger et al., 1997, Zhang and Reinberg, 2001). The octamer consists of two iterations of the canonical histones H2A, H2B, H4 and H3 (Figure 1.3). Multiple octamers are connected to each other by up to ~80bp of DNA (Thomas, 1999). This stretch of DNA contains an additional protein, H1, called a linker histone. A linker histone is present at both the point where DNA enters the octet and where it exits. While it is not considered to be part of the nucleosome, H1 is required for the condensation of chromatin because it helps to maintain the shape of the octet. This condensation results in the 11nm chromatin fiber being wound, into a form that contains six nucleosomes per turn, which yields a 30nm solenoid fiber (Zhang and Reinberg, 2001, Hayes and Hansen, 2001). With the possible exception of H4, multiple iterations exist of each of the canonical histones. These isoforms, called histone variants, are often substituted for their more ubiquitous cognates. Histone variants often serve a specific functional or developmental utility, permitting an additional level of transcriptional control. H2AX is a histone variant that is incorporated into nucleosomes in response to double stranded breaks (Kinner et al., 2008). TSH2B (~H2B.1) is an example of a tissue specific histone variant, and is expressed in the testis (Zalensky et al., 2002). Roughly three quarters of the amino acids in a histone forms the core of the protein and the remaining ~25% forms an evolutionarily conserved, but unstructured, tail (Zheng and Hayes, 2003).

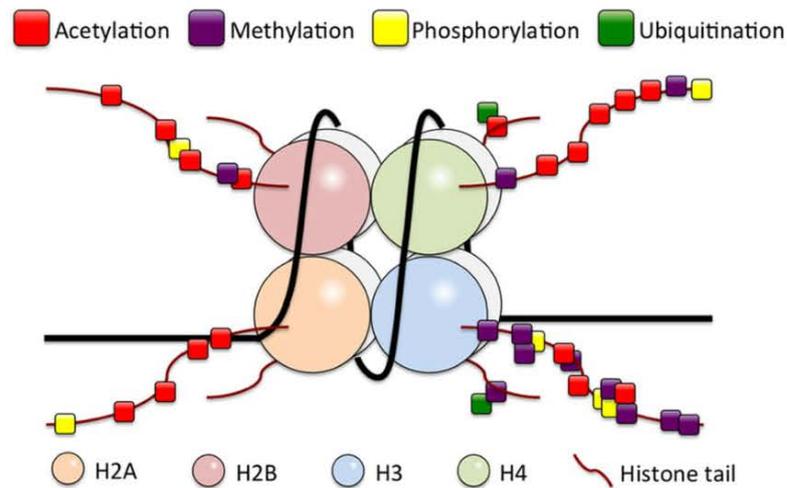


Figure 1.3: Nucleosome Post-translational Modifications.

The nucleosome composed of DNA and the associated histone octet. The histone tails that emerge from the nucleosome host a number of post-translational modifications (Kato et al., 2015).

Proteins can be regulated through the cycle of synthesis and degradation. An ancillary and more responsive layer of regulatory control is achieved through the covalent addition of modifying groups. A constellation of, largely enzyme mediated, covalent post-translational modifications (PTM) have been discovered. Modifications can alter the proteins' electrostatic interactions, as in the case of PARylation, phosphorylation and to a lesser extent – acetylation (Kostrhon et al., 2017). The former along with ubiquitinylation can also add significant bulk to their host protein (Bartolomei et al., 2016, Debelouchina et al., 2017). While the core histone residues can be modified, their tails disproportionately bear these covalent additions. These modifications can influence transcription by altering the properties of the DNA:histone complex. Modifications that enhance the strength of the DNA:histone interaction can promote the formation of the 30nm fiber, an arrangement which results in the DNA being inaccessible to the cell's transcriptional machinery (Schalch et al., 2005). Adding further nuance, most histone PTMs serve as binding sites for effector proteins that can impact gene expression. Due to their ability to alter multiple facets of histone function, the broad transcriptional effects of individual histone PTMs are often difficult to disentangle from their wider regulatory context. Histone methylation is associated with gene activation on some residues and correlated with transcriptional repression when found on others (Berger, 2007). While histone phosphorylation is largely associated with gene expression, it also

appears to facilitate chromatin condensation – suggesting the PTM plays a more intricate regulatory role (Rossetto et al., 2012). While it is commonly thought that histone acetylation is associated with transcriptional activation, the enzymes responsible for removing the mark do appear to be especially enriched at active genes (Kelly and Cowley, 2013). Figure 1.4 describes the effect of these post-translational modifications (PTM) in the most general of terms.

Histone PTMs		
Acetylated lysine (Kac)	H3 (9, 14, 18, 56), H4 (5, 8, 13, 16), H2A, H2B	Activation
Phosphorylated serine/threonine (S/Tph)	H3 (3, 10, 28), H2A, H2B	Activation
Methylated arginine (Rme)	H3 (17, 23), H4 (3)	Activation
Methylated lysine (Kme)	H3 (4, 36, 79), H3 (9, 27), H4 (20)	Activation Repression
Ubiquitylated lysine (Kub)	H2B (123 [§] /120 [¶]), H2A (119 [¶])	Activation Repression
Sumoylated lysine (Ksu)	H2B (6/7), H2A (126)	Repression
Isomerized proline (Pisom)	H3 (30–38)	Activation/ repression

[§]Yeast (*Saccharomyces cerevisiae*).
[¶]Mammals.

Figure 1.4: Histone Modifications, Their Locations and Effect on Gene Expression (Berger, 2007).

1.4. Histone Methylation

Histone methylation conceptually bears some similarities to DNA methylation as well as other histone modifications. S-Adenosyl methionine (SAM) is the donor of methyl groups placed on both DNA and histones (Valente et al., 2014). Histone methylation occurs on arginine and lysine residues. While arginine and lysine residues can be mono- or di-methylated, the latter can also be tri-methylated (Paik et al., 2007). Although the stoichiometry of the modification is important, its effects on transcription are also contextual. The H3R2 methylation is correlated with gene repression, and is antagonistic to the transcriptional activation associated H3K4 methylation (Ringel et al., 2015). Resembling the other major histone modification, acetylation, the ability of methylation to govern the accessibility of DNA to transcription factors largely depends on proteins capable of reading/binding to the epigenetic mark. These histone methylation effector proteins contain a structure called a chromodomain, Tudor domain

or MBT domain (Kim et al., 2006). There are several examples of enzymes containing chromodomains within the MYST family of histone acetyltransferases. The silencing of *ARX*, in beta cell formation, prevents H3K4 methylation; depriving acetyltransferases a binding site which has an overall inhibitory effect on gene expression (Ringel et al., 2015). Histone methylation was once thought to be a unidirectional mechanism, but the discovery of LSD1, a H3K4/H3K9 demethylase, established that the mark was not immutable (Bannister et al., 2002, Rudolph et al., 2013). There are three groups of proteins that give meaning to every epigenetic mark, including histone methylation, these are: the writers, readers and erasers.

1.5. Histone Acetylation

Distinct from histone methylation, acylations are modifications that can only be placed on lysine residues. The most studied histone acylation, and overall histone modification, is histone acetylation. The relationship between this post-translational modification and its influence on transcription was first described in 1964 by Allfrey and colleagues who found that histone acetylation influences chromatin through a similar pathway to histone methylation (Allfrey et al., 1964). Acetyl groups appear to directly promote transcription, in addition to recruiting effector proteins to chromatin. This distinguishes acetylation from histone methylation, and is achieved by masking the positive charge of the histone tail's lysine residues. The phosphate backbone of DNA gives the molecule a negative charge which is attracted to the lysines within the histone tails. The shielding of lysine's positive charge diminishes the strength of the DNA:histone interaction both within and between nucleosomes; the latter interaction taking place by impeding a mechanism called nucleosome tail-bridging (Korolev et al., 2006). The chemistry through which the acetyl group removes the lysine residue's positive charge is elaborated upon in chapter 1.5.1. Even the addition of a solitary acetyl group on a histone can produce a measurable increase in chromatin compaction (Shogren-Knaak et al., 2006). Histone acetylation is tied closely to transcription, permitting its influence on biological processes to be extensive. An example of histone acetylation's impact can be seen in genomic imprinting. In diploid organisms, the majority of genes are present in the form of two functional alleles, where one allele is inherited from each parent.

Imprinted genes are, instead, transmitted in a non-Mendelian manner. In the process of imprinting, the parental origin of the gene is marked, epigenetically, and the imprinted allele is inactivated (Wood and Oakey, 2006). The gametic transmission of these epigenetic marks results in the maintenance of the marks in the offspring's somatic cells. As only one functional allele inherited, a single parent contributes towards the entirety of the gene's function in their offspring. An example of this process can be seen in the differential histone acetylation of H4 lysine residues, which appears to be correlated with the, paternally inherited, functional *IGF2* allele (Grandjean et al., 2001). Inhibiting the removal of the acetyl moieties, through the use of sodium butyrate or Trichostatin A (TSA), de-repressed the silent *IGF2* allele (Pedone et al., 1999). The maternally inherited *H19*'s paternal imprinting appears to be a little more complex. This *IGF2*-linked gene's imprinting relies on synergistic effects of both histone deacetylation and DNA methylation (Pedone et al., 1999).

1.5.1. An Expanded Set of Epigenetic Instructions

Since the discovery of histone acetylation, several other lysine acylations have come to prominence. These include: 2-hydroxyisobutyrylation, β -hydroxybutyrylation (3-hydroxybutyrylation), butyrylation, crotonylation, formylation, glutarylation, malonylation, propionylation and succinylation (Rousseaux and Khochbin, 2015, Chen et al., 2007). These alternative histone acylations are present in organisms spanning from yeast to humans and are illustrated in Figure 1.5 (Zhang et al., 2009, Chen et al., 2007, Li et al., 2017).

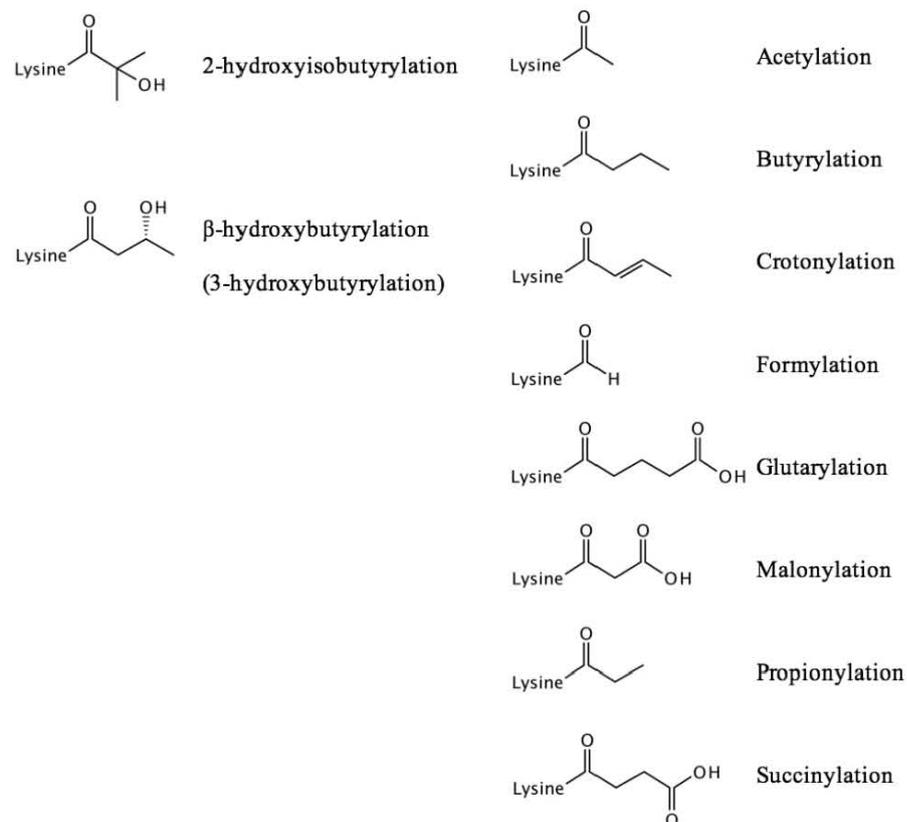


Figure 1.5: A Diagrammatic List of Lysine Acylations.

All acylations have the effect of negating their host lysine residue's charge and the acylations themselves do not need to be negatively charged to have this effect. The formation of the acetyl-lysine amide bond results in the lysine's ϵ -nitrogen, involved in the formation of this bond, losing a proton. This process is depicted in Figure 1.6. The loss of the proton disrupts the coulombic interaction between the DNA and the histones, promoting gene expression by permitting transcription factors greater access to the DNA. These modifications, particularly crotonylation, will be covered more fully in 1.7.5.

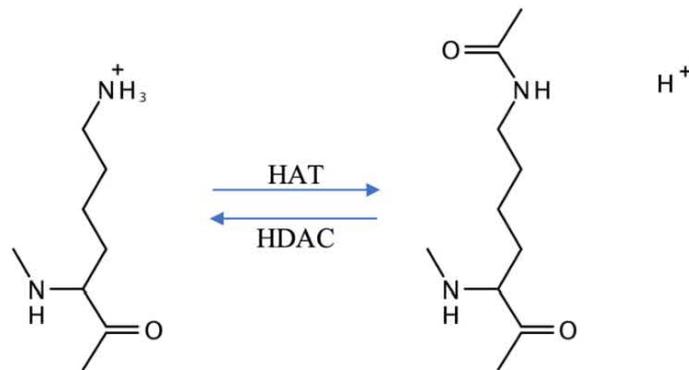


Figure 1.6: Histone Lysine Acylation (Acetylation) Results in the Loss of Lysine's Positive Charge.

1.6. Histone Acetylation Writers

Histone acetyltransferases (HAT) were originally classified into two broad groupings that were defined by their cellular distribution. One nuclear group was found to contain bromodomains which permitted them to read acetylated lysine residues and help to reinforce transcription (Richman et al., 1988). Another were predicted to be active in the cytoplasm where they could act on newly synthesized histones (Richman et al., 1988, Allis et al., 1985). A more appropriate classification divides the enzymes, on the basis of their conserved structural motifs, into three main groupings: EP300/CBP, Gcn5-related N-acetyltransferases (GNAT) and MYST (Neuwald and Landsman, 1997). As was the case with numerous HATs, EP300 and CBP were originally thought to merely be transcriptional co-activators; but were later found to have intrinsic HAT activity (Brownell et al., 1996, Marmorstein, 2001). The EP300/CBP family consist of at least 11 conserved proteins in higher eukaryotes which contain PHD, CoA-binding and ZZ domains (Yuan and Giordano, 2002). The ZZ domains are zinc finger motifs capable of simultaneously binding two zinc ions (Ponting et al., 1996). The GNAT family are more loosely defined and contain five 15-33 amino acid long motifs named A, B, C and D; with acetyl-CoA being held by a structure formed by A, B and D (Modis and Wierenga, 1998, Marmorstein, 2001). The MYST family was named for the group's founding enzymes MOZ, Ybf2 (Sas3), Sas2 and Tip60 (Yuan et al., 2012). Tip60, like the other members of the MYST family, is auto-acetylated within its active site – an event that regulates the enzymes catalytic activity (Yang et al., 2012). The

majority of the HAT lysine substrates are present on H3 and H4 but are distributed throughout linker and even histone variants like H2AX. The core histone lysine residues that are acetylated, include: H2AK(5, 9, 13, 15 and 36), H2BK(5, 12, 15, 20 and 24), H3K(4, 9, 14, 18, 23, 27, 36 and 56) and H4K(5, 8, 12, 16, 20, 31, 44, 77, 79 and 91) (Tan et al., 2011, Sterner and Berger, 2000).

1.7. Histone Acetylation Readers

The reader domains for histone acetylation are the bromodomains, YEATS domains and tandem PHD (plant homeodomain) domains (Sanchez and Zhou, 2009, Zeng et al., 2010). Some of the enzymes that contain a bromodomain are the mediators of histone acetylation themselves, notable examples include the CBP and EP300 histone acetyltransferases. This can lead to a positive feedback loop whereby histone acetylation, particularly of H3 and H4, creates a binding site for bromodomain containing HATs – resulting in further histone acetylation (Delvecchio et al., 2013). The bromodomain and extra-terminal (BET) family of proteins provide a less idiosyncratic example of bromodomain proteins and their functions. The family is comprised of four proteins, which are notable for having two tandem bromodomains (Taniguchi, 2016). Included in their number are BRD2 and BRD4, which primarily bind to a number of acetylated lysine residues on H4 (Filippakopoulos and Knapp, 2012). The two proteins couple histone acetylation to transcription. They do this by recruiting positive transcription elongation factor complex (P-TEFb) which regulates transcription through the phosphorylation of the C-terminal domain of RNA polymerase II (Pol II) (Muller et al., 2011). AF9, a YEATS (Yaf9, ENL, AF9, Taf14, Sas5) domain protein, binds to acetylated H3K9, H3K18 and H3K27. It is also known to contribute to the transcriptionally conducive H3K79 methylation through the recruitment of the DOT1L methyltransferase (Li et al., 2014). The MYST acetyltransferase, MOZ has a tandem PHD domain that allows it to bind to H3K14Ac. This localization permits the HAT to promote acetylation across H3, in a manner that is reminiscent of the CBP/EP300 feedback loop (Qiu et al., 2012).

1.8. Histone Acetylation Erasers: HDACs

The acetylation marks that are placed by HATs are removed by another class of enzymes called histone deacetylases (HDACs). This process increases intra-nucleosome cohesion which represses transcription. The discovery of HDACs was tied to research of a membrane high affinity potassium transporter (TRK1) in *S. cerevisiae* cells. The deletion of TRK1 resulted in cells that were unable to grow in low potassium media (Vidal et al., 1990). Mutations in three genes were found to reduce the potassium dependency (RPD) of the TRK1 cells. A dominant mutation in cells, RPD2, was found to produce a hypermorphic allele of a low affinity potassium transporter (TRK2); RPD1 and RPD3 were found to negatively regulate TRK2 expression (Ko et al., 1990, Vidal et al., 1990). As Vidal and colleagues (1990) hypothesized, the lack of additive effect from combined RPD1 and 3 mutations was indicative of the two proteins being components of the same pathway. RPD1's other name is SIN3, while RPD3 was later revealed to be an orthologue of class I HDACs found in mammals (Taunton et al., 1996, Seto and Yoshida, 2014, Vidal and Gaber, 1991).

Since this discovery, a total of 18 different HDAC enzymes have been identified; these were divided into four classes, based on evolutionary relatedness to their yeast orthologues (Seto and Yoshida, 2014). The general layout of the HDACs and their classifications can be seen in Figure 1.7. The four classes broadly display two types of enzymatic activity. The evolutionarily conserved activity of class I, II and IV HDACs is contingent on zinc binding. These 11 mammalian enzymes are called classical HDACs. Class I HDACs, named after the order of their discovery, include HDACs 1, 2, 3 and 8; these are the closest relatives of RPD3, with the first two sharing the greatest homology. HDAC1 and 2 share 83% identity, and predictably demonstrate significant compensatory effects for each other in knockout experiments (Dovey et al., 2010). Class II HDACs are related to yeast HDA1, this cohort consists of HDACS 4-7, 9 and 10 (Bjerling et al., 2002). Each of these class II enzymes demonstrates at least some degree of cytoplasmic localization (Seto and Yoshida, 2014). The six HDACs in class II are further subdivided into class IIa and IIb. A key tyrosine residue that stabilizes the deacetylation reaction is missing in class IIa HDACS, and this contributes to their

diminished enzymatic activity (Lahm et al., 2007). In addition to the conserved classical deacetylase domain, class IIa enzymes have an extended N-terminus that contains motifs that are important for the enzyme's regulation (Yang and Gregoire, 2005). HDACs 6 and 10 differ from this arrangement, these class IIb proteins each have a putative second catalytic domain that is absent in the other HDACs (Seto and Yoshida, 2014). Class IV's only member is HDAC11 which does not fit neatly into either of the first two HDAC classes. HDAC11 contains amino acid moieties, within its catalytic core, that are shared by both class I and class II HDACs (Gao et al., 2002).

Class III HDACs or 'sirtuins' bear homology to the yeast SIR2. In contrast to the classical HDACs, the deacetylase activity of sirtuins depends on the hydrolysis of an oxidized coenzyme called nicotinamide adenine dinucleotide (NAD) (Smith and Boeke, 1997). There are seven sirtuins (SIRT1–7) in mammals (Michan and Sinclair, 2007). Sirtuins 1 and 2 are known for their lack of substrate specificities, a feature that has led Hsu and colleagues (2016) to refer to them as universal histone deacetylases. Their reliance on NAD⁺ ties these proteins a little more closely to cellular metabolism than other epigenetic enzymes.

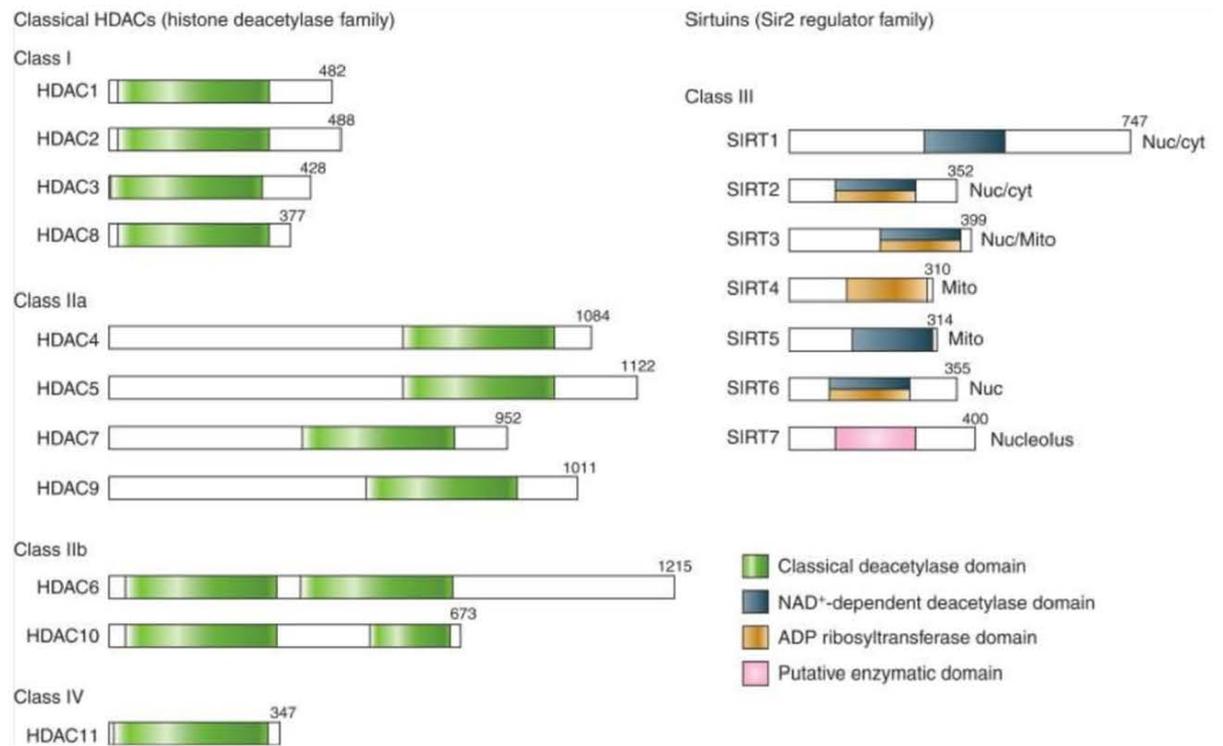


Figure 1.7: The HDAC Family of Enzymes with the Relative Positions of Their Enzymatic Domains (Seto and Yoshida, 2014).

1.8.1. HDAC1/2 Complexes

Class I HDACs are ubiquitous, and demonstrate significant enzymatic activity and nuclear localization (Kelly and Cowley, 2013). The widespread influence of HDAC1/2 largely results from their incorporation in complexes. Broadly speaking there are four major HDAC1/2 complexes – the SIN3, NuRD, CoREST and MIDAC complexes (Kelly and Cowley, 2013, Bantscheff et al., 2011). The former three complexes are the most thoroughly described, and can be seen in Figure 1.8. The Sin3a:HDAC1/2 repressor complex appears to maintain the greatest evolutionary conservation, as each of its core constituents has a yeast orthologue – distinguishing it from the other HDAC1/2 complexes (Kelly and Cowley, 2013).

HDAC1/2 Complexes

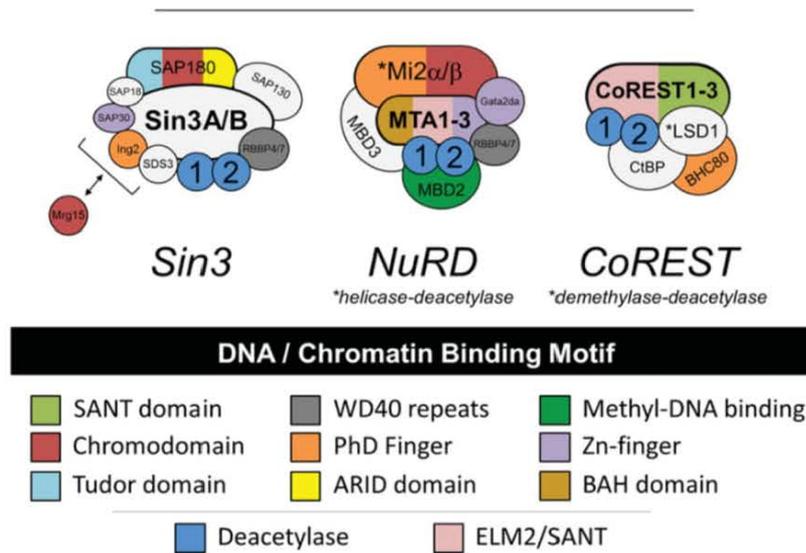


Figure 1.8: The HDAC1/2 Complexes: SIN3 Co-Repressor, NuRD and CoREST Complexes (Kelly and Cowley, 2013).

1.8.2. The Sin3a Co-Repressor Complex

The MAD1-MAX complex negatively regulates gene expression. An investigation into the mechanism behind this activity revealed the Sin3a:HDAC1/2 co-repressor complex (Laherty et al., 1997, Ayer et al., 1995). The complex contains approximately ten major components, all of which either maintain its integrity or help mediate the enzymatic effects of the attached HDAC proteins. RBAP46/RBBP7 and RBAP48/RBBP4 proteins are somewhat unique, as they are the only constituents of the Sin3a corepressor complex that do not directly associate with the scaffold-protein; instead they bind through an association with HDAC1/2, SAP30 and SAP45/SDS3 (Grzenda et al., 2009). Nearly all the other core components bind to the region of Sin3a between PAH3 and the histone deacetylase interaction domain (Figure 1.15). SAP45/SDS3 is needed for the integrity of the SIN3:HDAC complex; mutations in this protein yield phenotypes in yeast that are similar to Rpd1 (SIN3) and Rpd3 (HDAC1 orthologue) mutants (Lechner et al., 2000). SAP30 binds to both PAH3 and HDAC1, and it is through SAP30 that the repressor complex is recruited to a subset of its targets by RBP1 (Lai et al., 2001, Zhang et al., 1998). SAP180 also interacts with SAP30, and may serve to stabilize the complex

(Fleischer et al., 2003). Another Sin3a associated protein called SAP18 likely plays a role as an adaptor that sits between interacting proteins, like the HAIRY1 transcription factor, and the SIN3 repressor complex (Sheeba et al., 2007). SAP130 interacts with corepressors and coactivators which may modulate the repressor complex's activity (Fleischer et al., 2003). ING2 is a protein involved in the recruitment of the complex's histone deacetylase activity to the H3K4me3 mark (Champagne and Kutateladze, 2009). The protein additionally links the Sin3a repressor complex to another PTM pathway – SUMOylation. ING2 has to first receive a small ubiquitin-like modifier (SUMO1) on residue K195 before it can associate with the Sin3a repressor complex (Ythier et al., 2010). Despite these components, the Sin3a repressor complex surprisingly lacks any native ability to directly interact with DNA. This gap is bridged by a diverse host of proteins. This includes the transcription factor p53 which recruits the complex to DNA which results in the silencing of the MXD1 gene (Chun and Jin, 2003). The DNA demethylase, TET1, appears to also recruit Sin3a to DNA (Williams et al., 2011).

1.8.3. NuRD, NODE and CoREST

The nucleosome remodeling and deacetylation (NuRD) complex is closer to the Sin3a repressor complex in terms of components than CoREST. Among these shared elements are: HDAC1/2, RBBP7 and RBBP4 (Zhang et al., 1999). NuRD additionally contains MBD2 or MBD3 in a manner that is mutually exclusive, but necessary for deacetylase activity. The latter homologue, while incapable of binding to methylated DNA, is required for both differentiation and the efficient production of induced pluripotent stem cells (iPSCs) within the context of the NuRD complex (dos Santos et al., 2014). It was initially thought that NuRD was exclusively involved in differentiation because MBD3 was found to suppress the Oct4 pluripotency factor. It was even demonstrated that cells in culture lacking MBD3 could self-renew and were resistant to differentiation even after withdrawal of leukemia inhibitory factor (Kaji et al., 2006, Hu and Wade, 2012). A separate, but closely related complex called Nanog/Oct4-associated deacetylase (NODE), which lacked MBD3, was deemed responsible for the maintenance of pluripotency. While NuRD contains one of the three metastasis associated proteins (MTA1-3), NODE was defined as only incorporating MTA1 (Liang et al., 2008). A

knockdown study revealed that MTA1 suppresses genes involved in differentiation. This MTA1-MBD3 switch was thought to alternate the activity of the NuRD/NODE components between promoting differentiation and maintaining pluripotency. Recently, however, it was found that the NuRD:MBD3 complex alone can maintain both activities in a contextually determined manner and that this complex can enhance the reprogramming of iPSCs (dos Santos et al., 2014).

Corepressor of RE-1 Silencing Transcription Factor (CoREST) contains HDAC1/2 but none of the other components common to both the Sin3a and NuRD complexes (Andres et al., 1999). It contains the H3K4me/me2 demethylase, LSD1 which requires the presence of the HDACs for its repressive activity. The application of TSA, a pan-HDAC (1-11) inhibitor, results in the loss of repression of LSD1 target genes (Shi et al., 2005). The association with CoREST appears to permit LSD1 to function on its nucleosomal targets (Shi et al., 2005). The CoREST complex plays an important role in the establishment and maintenance of pluripotency and in the neuronal pathway (Yang et al., 2011, Abrajano et al., 2009)

1.8.4. Non-Histone Acetylation Targets

It is worth noting that the evolution of the classical HDACs precedes that of the histone proteins (Seto and Yoshida, 2014). As this implies, despite their nomenclature, HDACs have numerous non-histone substrates. One of the first non-histone proteins found to be acetylated was p53 (Glozak et al., 2005). The acetylation of p53's C-terminal domain allows it to bind to DNA in a site-specific manner, but also prevents its MDM2 mediated ubiquitinylation and resultant turnover (Luo et al., 2004, Rodriguez et al., 2000). ING2, present in some iterations of the Sin3a corepressor complex, recruits EP300 to p53. This facilitates the tumor suppressor's acetylation (Pedeux et al., 2005). The increased stability proffered by EP300's acetylation of p53 is reversed by the activity of HDAC1 or even SIRT1 (Luo et al., 2000, Luo et al., 2001). STAT3's DNA binding and transactivation is similarly enhanced by acetylation (Wang et al., 2005). The activity of EP300 in this case is reversed primarily through the activity of HDAC3, though HDAC1 and 2 can do the same with reduced efficiency (Yuan et al., 2005). All

three of the TET proteins are acetylated. TET2's association with DNMT1 protects against aberrant DNA methylation. The formation of this complex depends on the acetylation of TET2's K110 by EP300 (Zhang et al., 2017). Both the acetylation and the TET2:DNMT1 complex are undone by HDAC1/2 (Zhang et al., 2017). Choudhary et al. (2009) demonstrated using HDAC inhibitors and mass spectrometry that 1,750 proteins feature the acetylation PTM. As the aforementioned examples suggest, epigenetic proteins are over-represented in the acetylome. Most major HATs and HDAC1/2 and associated proteins, including MBD2, MBD3, MTA2, RBBP4, RBBP7, RBP1, SAP30 and Sin3a are acetylated at least once. Histone methylation is also regulated by acetylation with MLL proteins, HCFC1 and the Jumonji/ARID domain-containing demethylases all being acetylated. Ubiquitin ligases and deubiquitylases are similarly represented in the acetylome.

1.9. Alternative Acylations: Crotonylation

The alternative acylations can be sorted into three groups as described by Sabari and colleagues (2017). The first are hydrophobic groups which include the butyryl, crotonyl, propionyl and formyl PTMs. These increase the hydrophobicity and bulk of the lysine residue, to different extents, through the addition of a hydrocarbon chain in a manner that is similar to lysine acetylation. Crotonylation is a slight outlier in this group. While similar to butyrylation, it has the distinction of being the only acylation that exists in a single plane due to its inflexible central alkene bond. The second category of acylations decreases the affinity of the DNA:histone interaction even further. These negatively charged acyls, at physiological pH, shift the charge of the lysine residue two places so that it is negative, in a manner that is not dissimilar to phosphorylation (Xie et al., 2012, Tan et al., 2014). The modifying groups that fall into this category are glutaryl, malonyl and succinyl. The last category consists of 2-hydroxyisobutyryl and β -hydroxybutyryl (3-hydroxybutyryl). These polar groups, introduce steric bulk, and through a terminal hydroxide permit the modified lysine residue to hydrogen bond with other molecules (Xiao et al., 2015, Sabari et al., 2017). It is worth briefly mentioning that a fourth group of fatty acylations, such as myristoylation and palmitoylation, are also found on lysine

residues (Choudhary et al., 2014). However, the relevance of this last group of PTMs on epigenetics requires further investigation.

1.9.1. Crotonylation Writers

The discovery of these alternative forms of histone lysine acylations made the absence of a crotonyl-transferase or succinyl-transferase conspicuous. An attempt to isolate the effector of histone crotonylation from HeLa cell extract resulted in the co-purification of histone-acetyltransferase activity (Sabari et al., 2015). The two processes were biochemically indivisible, and were subsequently found to be mediated by the same enzyme – CBP/EP300. EP300 was later established to be a promiscuous acylase and is largely the mediator of all the alternative lysine acylations. The catalytic ability of this molecule is, however, inversely proportional to the length of the acyl's hydrocarbon chain (Kaczmarska et al., 2017). The broad specificity of this acyl-transferase is a result of a hydrophobic pocket present within EP300's active site, which is noticeably absent in many other acyl-transferases (Kaczmarska et al., 2017). The other acyl-transferases are not entirely devoid of this alternate activity. The HAT, PCAF, for example, is involved in lysine propionylation. It is notable that the enzyme comparatively makes use of a much narrower range of Acyl-CoAs (Liu et al., 2009). EP300 appears to be an appropriate broad spectrum acyl-transferase not least because it has been demonstrated, *in vitro*, that it is able to function on all four core histones inside a chromatin template (Lu et al., 2002). It is unknown if lysine formylation results from EP300 activity or if it occurs non-enzymatically, an example of this latter process being 3-phosphoglyceroylation which occurs through the activity of a chemically reactive intermediate molecule (Moellering and Cravatt, 2013). Just as acetyl moieties are transferred, by HATs, from the acetyl-CoA molecule; alternative acylations can also be found in a similar CoA-appended form. These acyl-CoAs, like acetyl-CoA, result as the byproduct of metabolic processes (Figure 1.1). The discovery of alternative acyl-CoAs, and their relative concentrations, exposed the mechanism through which the promiscuous EP300 selects between functioning as an acetyltransferase and alternative acyltransferase. If the ratio of crotonyl-CoA could be increased in comparison to acetyl-CoA, EP300 shifts towards crotonyl-transferase activity (Sabari et al., 2015). This

availability directly ties transcription to the metabolic processes that produce these acyl-CoA molecules. ATP citrate lyase (ACL) and the pyruvate dehydrogenase (PDH) complex activity bolsters acetyl-CoA production, whereas acyl-CoA synthetase 2 (ACSS2) leads to the generation of crotonyl-CoA (Figure 1.9).

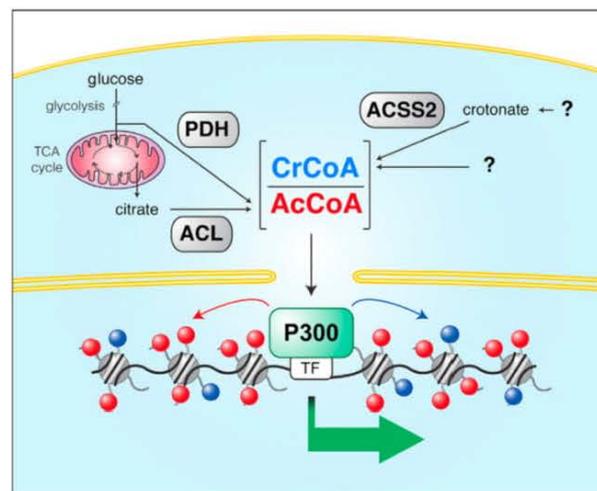


Figure 1.9: The Relative Acyl-CoA Concentrations Dictate the Activity of EP300.

The relative availability, in terms of concentration, of crotonyl-CoA and acetyl-CoA determine whether EP300 behaves as a histone crotonyl or acetyl-transferase (Sabari et al., 2015).

This intracellular ratio can be perturbed through the addition of sodium crotonate in the case of crotonylation. Intriguingly, sodium butyrate was discovered in 1978 to be a potent inhibitor of HDACs, but the same activity was not found for sodium crotonate or acetate (Sealy and Chalkley, 1978, Sabari et al., 2015). Butyrate's promotion of histone acetylation is twofold – it acts as an HDAC inhibitor and in addition provides a metabolic source of acetyl-CoA; each mechanism appears to activate a different set of genes (Donohoe et al., 2012). The belated discovery of these networks of alternative acylations may be due in part to their distribution, in cell culture conditions there are three orders of magnitude more acetyl-CoA than crotonyl-CoA (Sabari et al., 2015). This sits in contrast to their relative availabilities in more complex metabolic systems. In rat liver, for example, the ratio of acetyl-CoA, propionyl-CoA, butyryl-CoA and crotonyl-CoA is 40:23:10:1 (King and Reiss, 1985). In this context, it appears that crotonylation, and certainly propionylation and butyrylation, might have a significant epigenetic impact. There is a dearth of information on lysine formylation in histones,

even in comparison to its more novel acyl cognates. This modification occurs on residues that are frequently methylated or acetylated in a manner that is antagonistic to the placement of these modifications (Fang et al., 2015). Parsing out the effect of this modification is further complicated by the discovery that it is not unique to lysine, but was also found in serine and threonine residues (LeRoy et al., 2009). Outside of formylation, the three groups of lysine acylations appear to have a positive effect on transcription. In the case of crotonylation this appears to be to a much greater extent than even acetylation (Sabari et al., 2015). This apparent similar function, common EP300 acyl-transferase mediator and the seeming partial interchangeability of their distribution on histones (Figure 1.10) led Rousseaux and Khochbin (2015) to propose a model whereby there were two groups of acylations: acetylations and a second, super-group, of the other acylations that they postulated to be interchangeable in function.

H4 K5(Ac/Bu/Pr/Cr/Hib)
 K8(Bio/Ac/Bu/Pr/Cr/Hib/Bhb)
 K12(Bio/Me/Ac/Bu/Pr/Fo/Cr/Suc/Hib/Bhb)
 K16(Ac/Bu/Pr/Hib)
 K20(Me/Ac)
 K31(Bu/Pr/Fo/Suc/Hib)
 K44(Bu/Pr/Hib)
 K59(Hib/Fo)
 K77(Ac/Bu/Pr/Fo/Suc/Hib/Bhb)
 K79(Me/Ac/Bu/Pr/Fo/Hib)
 K91(Ac/Bu/Pr/Fo/Suc/Hib/Bhb)

Figure 1.10: Human H4 Lysine Modifications (Xie et al., 2012, Dai et al., 2014, Camporeale et al., 2004, Wiśniewski et al., 2008, Goudarzi et al., 2016, Arnaudo and Garcia, 2013).

This proposed theory stated that the differences in acylation patterns resulting from this super-group were merely a result of the metabolic availability of the various acyl-CoA molecules (Figure 1.11).

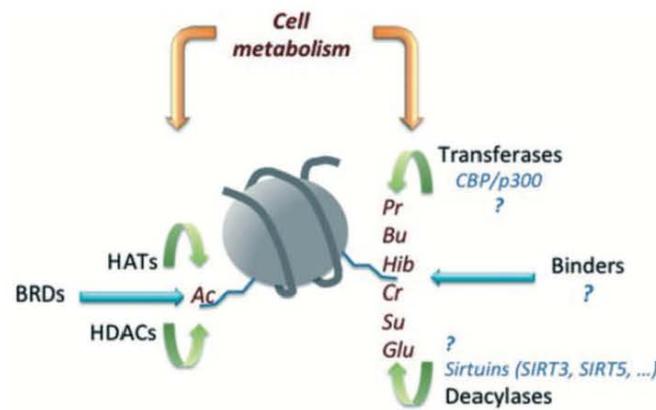


Figure 1.11: The Acetylation and Alternative Acylation Model.

The proposed acetylation vs other acylation mechanism, where alternative acylations are underpinned by metabolic availability (Rousseaux and Khochbin, 2015).

1.9.2. Crotonylation Readers and Function

Propionylation can be read, with reduced efficiency, by the archetypical acetylation reader – the bromodomain (Flynn et al., 2015). This was demonstrably the case regarding BRD4 (Vollmuth and Geyer, 2010). The YEATS domain proteins AF9, TAF14, and YEATS2 were also known to bind to histone acetyl groups, but latterly were shown to have four times more affinity for crotonylated lysines. In addition, the YEATS domain was also found to bind to propionylation and butyrylation (Li et al., 2016). AF9 along with ENL are part of the super elongation complex (SEC). P-TEFb, within the SEC, works with BRD4 and MYC to enable RNA polymerase II to transition from promoter-proximal pausing into productive transcription (Lin et al., 2011). AF9 also promotes transcription through the H3K79 methyltransferase complex it forms with DOT1L (Kuntimaddi et al., 2015). TAF14, within Transcription factor II D, forms the part of the Pol II preinitiation complex that positions the polymerase on the gene's promoter for subsequent transcription (Feigerle and Weil, 2016). YEATS2 is the scaffold subunit of the, GCN5 and PCAF containing, ATAC complex which acetylates H3 (Wang et al., 2008).

The alternative acylations functioned in spermatogenesis, where crotonylation and hydroxyisobutyrylation permitted genes to escape meiotic sex chromosome inactivation

(Sabari et al., 2017). While hydroxyisobutyrylation, butyrylation and acetylation are reduced in this process; crotonylation is undiminished, suggesting regulatory nuance between the alternative acyl moieties. Montellier et al. (2012) proposed that crotonylation, like the H3K27me and H4K20me marks, was resistant to removal; and that it was this that enabled genes to escape inactivation (Huang et al., 2015). The crotonyl mark also plays a role in acute kidney injury response. Crotonylation increases upon this injury and the application of crotonate appears to afford a measure of protection against this insult, though this also resulted in increased expression of SIRT3 (Ruiz-Andres et al., 2016).

1.9.3. Lysine Acylation Erasers

SIRT1-3 are known to have deacetylase activity, but were later found to possess decrotonylation activity; though, only SIRT3 appeared to fulfil this role *in vivo* (Bao et al., 2014). Relatedly, SIRT5 was found to remove the negatively charged histone acylations; an activity that SIRT6 was discovered to have on fatty-acylations (Hirschey and Zhao, 2015, Jiang et al., 2013). The importance of sirtuins in these processes is outlined in Figure 1.12.

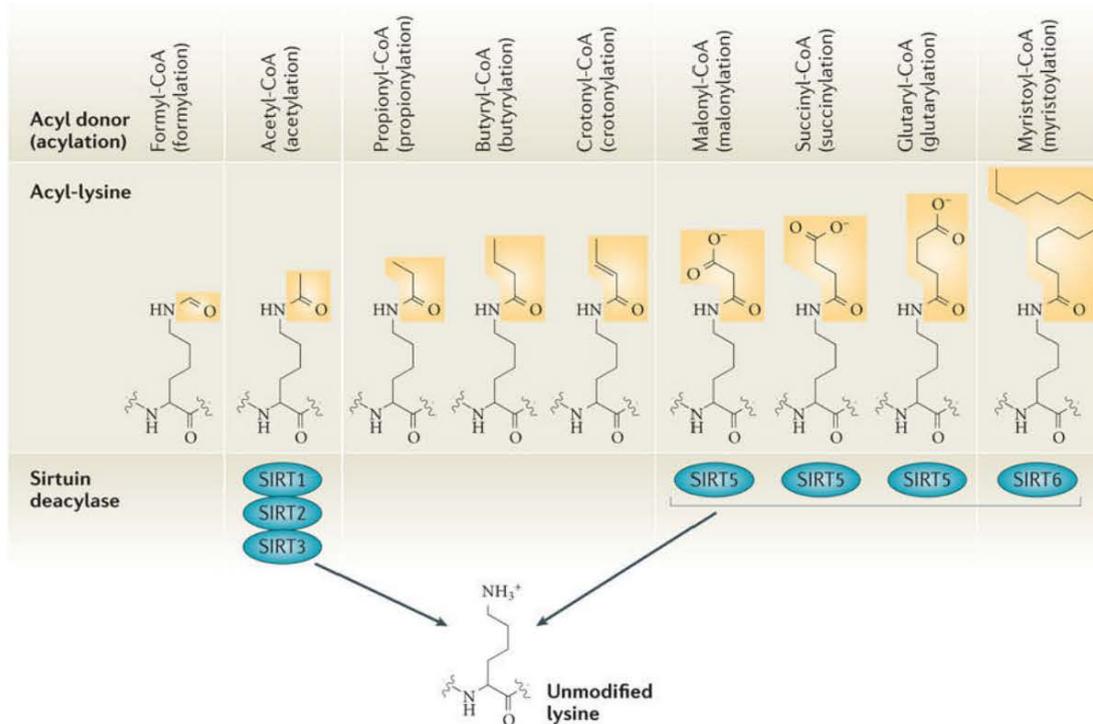


Figure 1.12: A List of Lysine Acylations and Their Erasers (Choudhary et al., 2014).

The nuclear distribution (Figure 1.13) and ubiquity (Figure 1.14) of the classical HDACs, particularly the members of class I, suggests that they may play a role in deacetylation. Indeed, HDAC8 was recently found to remove fatty-acylations, an activity that was briefly thought to be exclusive to SIRT6 (Aramsangtienchai et al., 2016).

Classification	HDAC	Subcellular localization
I	HDAC1	nucleus
	HDAC2	nucleus
	HDAC3	nucleus
	HDAC8	nucleus
IIa	HDAC4	nucleus/cytoplasm
	HDAC5	nucleus/cytoplasm
	HDAC7	nucleus/cytoplasm
	HDAC9	nucleus/cytoplasm
IIb	HDAC6	mostly cytoplasm
	HDAC10	nucleus/cytoplasm
III	SIRT1	nucleus/cytoplasm
	SIRT2	cytoplasm
	SIRT3	nucleus/mitochondria
	SIRT4	mitochondria
	SIRT5	mitochondria
	SIRT6	nucleus
	SIRT7	nucleolus
IV	HDAC11	nucleus/cytoplasm

Figure 1.13: Localization of the HDAC Enzymes (Heidemann, 2015).

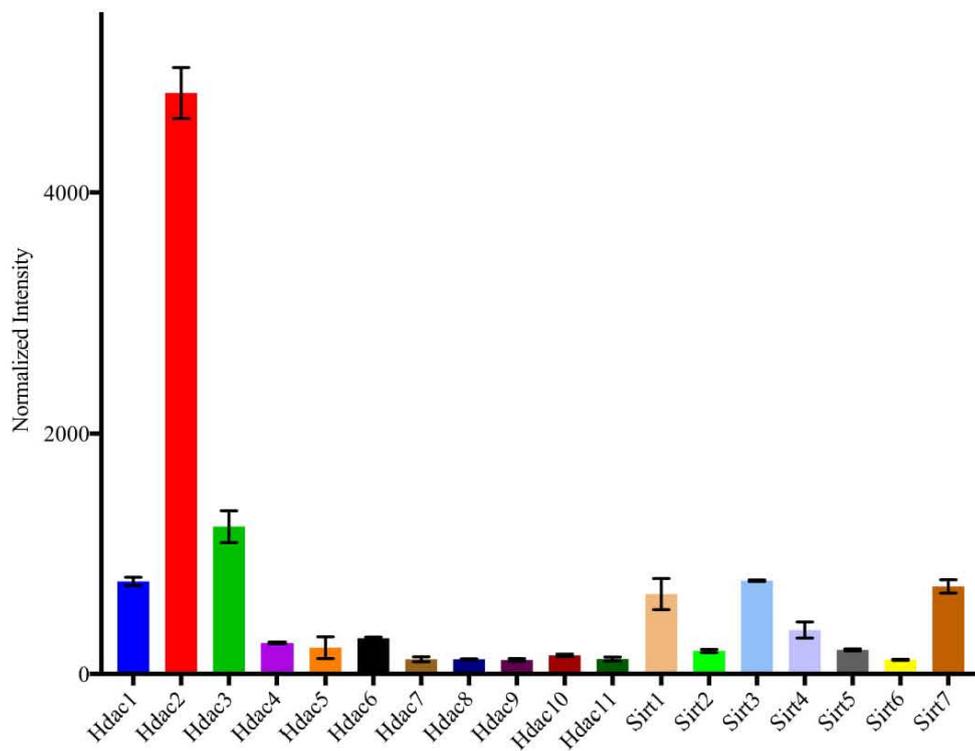


Figure 1.14: Comparative Expression of HDACs in ESCs Obtained from Microarray (Jamaladdin et al., 2014).

1.10. Sin3a and the TET1 Binding Site

Originally discovered in yeast mate-type switching, SIN3 exists as two isoforms in mammals – Sin3a and Sin3b (Sternberg et al., 1986, Ayer et al., 1995). Sin3a is involved in embryonic development, cell cycle progression, differentiation and oncogenesis; its absence results in lethality at ~E3.5, and is similarly fatal in ESCs (Cowley et al., 2005, Kadamb et al., 2015). The Sin3a scaffold-protein ties DNA demethylation to histone deacetylation. The link to DNA demethylation occurs through an interaction with the DNA demethylase TET1. The two proteins have similar binding profiles as TET1 appears to recruit Sin3a (Williams et al., 2011). The formation of the Sin3a co-repressor complex is centered around the region of Sin3a where HDAC1/2 bind, depicted in Figure 1.15. This region is called the HDAC interaction domain (HID). The TET1 interaction domain on Sin3a, however, has not yet been determined.

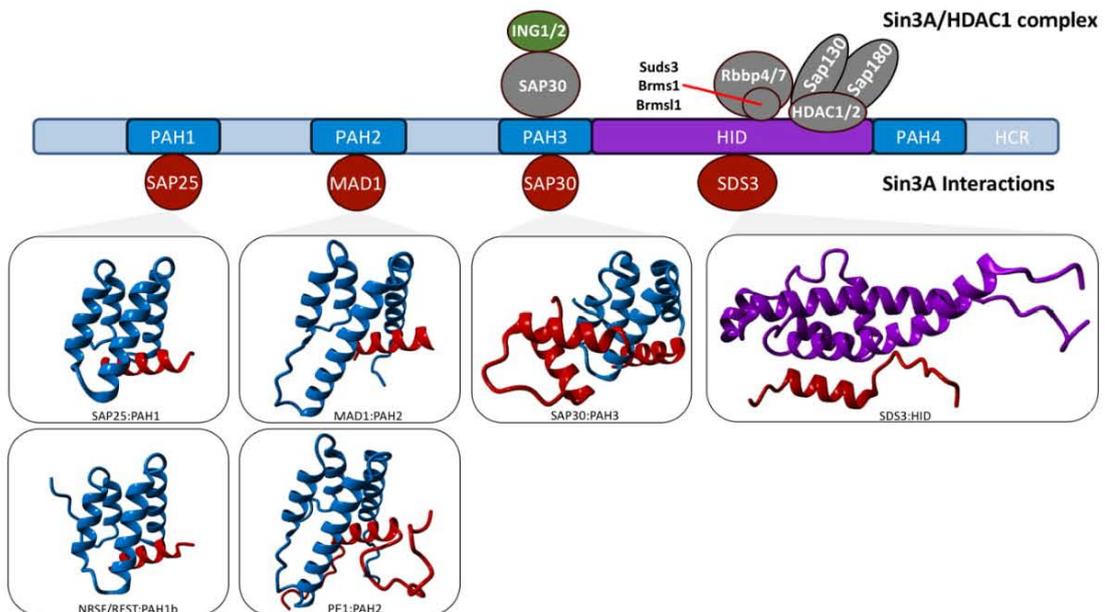


Figure 1.15: A Schematic of SIN3 Serving as a Scaffold-Protein in its Role in the Sin3a HDAC1/2 Repressor Complex (Grzenda et al., 2009).

Illustrations derived from NMR solutions are displayed underneath. From left to right, there are: PAH1 (Sin3a 119-189) and SAP25 (126-186), PAH1 (Sin3b 31-107) and NRSF, PAH2 (Sin3a 295-283) and MAD1/MXD1 (6-21), PAH2 (Sin3a 292-385) and Pfl1 (197-241), and Sin3a 605-729 and SDS3 (202-228) (Sahu et al., 2008, Nomura et al., 2005, Swanson et al., 2004, Kumar et al., 2011, Xie et al., 2011, Clark et al., 2015).

1.10.1. PAH Domains

Sin3a makes the majority of its protein:protein interactions through structures called paired amphipathic helices (PAH), illustrated in Figure 1.16. The structures of PAH1-3 have been solved (Figure 1.15) and they share a high degree of similarity and overall tertiary configuration (Sahu et al., 2008, Zhang et al., 1998, Brubaker et al., 2000). They each consist of four alpha helices that form a hydrophobic cleft (Brubaker et al., 2000). Binding partners attach to this region through an alpha helical structure called a SIN3 interaction domains (SID). The interaction is not promiscuous; instead, the hydrophobic residues in both structures ensure the SID fits within the PAH domain like a key inserted into a lock (Figure 1.17). This ensures that every SID corresponds to a specific PAH domain (Cowley et al., 2004). The hydrophobic residues that have proven critical to binding in PAH1-3 are missing in PAH4; suggesting that it adopts a tertiary structure dissimilar to the other domains (van Ingen et al., 2006).

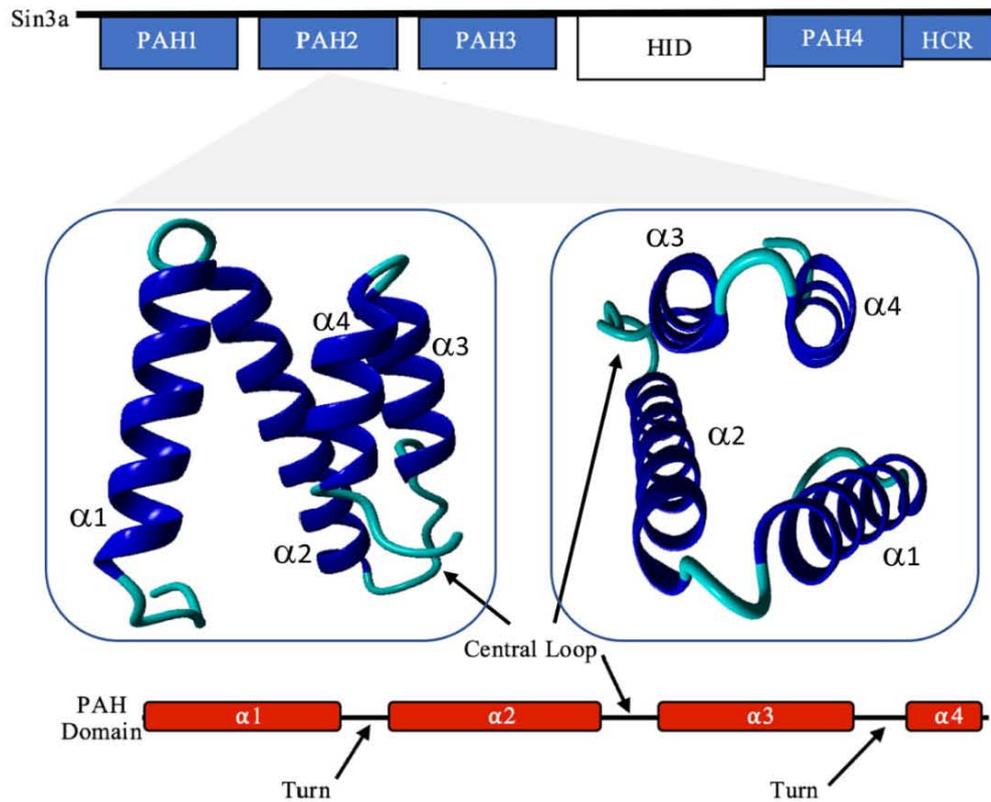


Figure 1.16: Sin3a Has Four PAH Domains, Which Are Each Comprised of Four Alpha Helices.

A schematic diagram of the primary structure of Sin3a as well as the NMR solution for PAH2 viewed from the side (left) and from above (right) (Swanson et al., 2004). A simplified primary structure of one of the first three canonical PAH domains (below). Sin3a consists of PAH1 (119-189), PAH2 (295-383), PAH3 (459-526), HID (545-898), PAH4 (898-965) and the HCR (1001-1137) (Sahu et al., 2008, Swanson et al., 2004, Dhordain et al., 1998, Moehren et al., 2004).

PAH1 interacts with the SIDs of a number of proteins, including, NRSF and SAP25. PAH1-SIDs may be dissimilar in terms of their specific amino acid sequences, but appear to share an eight amino acid motif (Figure 1.17) – $\Phi x \Phi \Phi s x \Phi s$; where ‘ Φ ’ is a bulky hydrophobic residue, ‘x’ is any non-proline residue and ‘s’ is an amino acid with a short side chain (Sahu et al., 2008). Sahu and colleagues (2008) found that the motif is palindromic, in that it can effectively bind PAH1 in either orientation, $\Phi x \Phi \Phi s x \Phi s$ or $s \Phi x s \Phi \Phi x \Phi$ – respectively called type I or type II binding. A problem with this prediction is that the motif is very loosely defined.

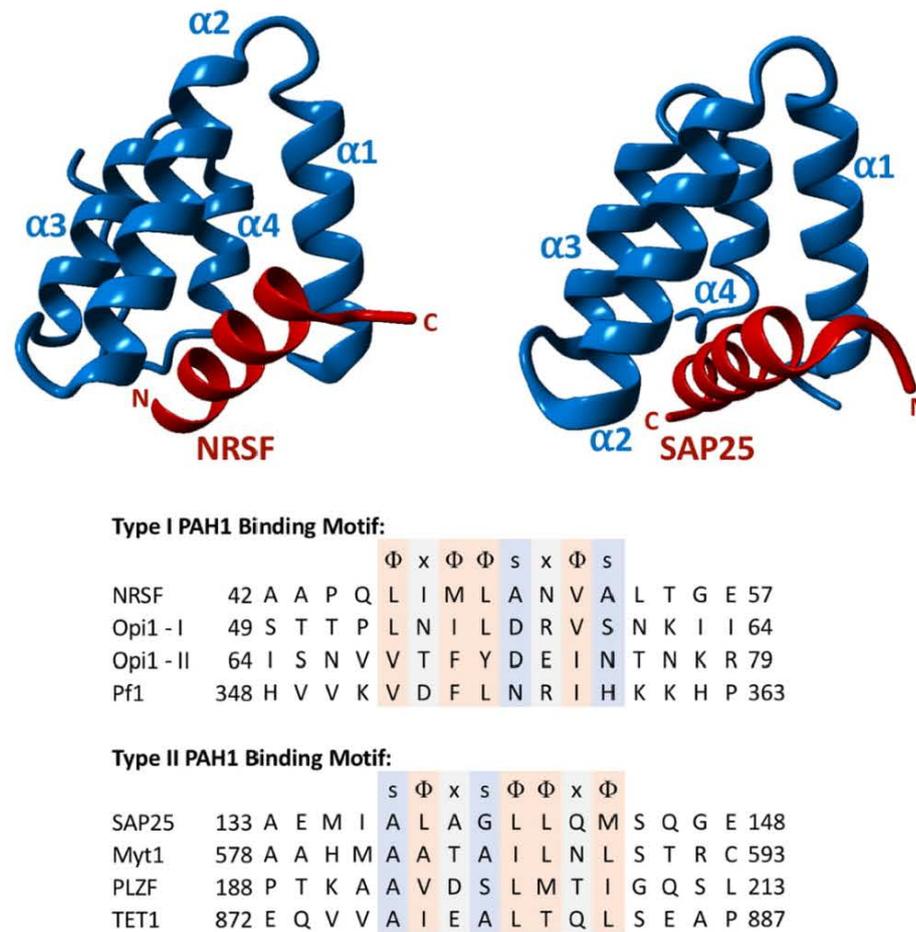


Figure 1.17: Sin3a Binding Motifs.

The image on the top left is of NRSF's SID (red) forming a complex with Sin3b's PAH1 (blue), the image on the top right is Type I binding by SAP2's SID (red) to Sin3a's PAH1 (blue) (Nomura et al., 2005, Sahu et al., 2008). Type I and Type II PAH1 binding motifs are displayed underneath (Sahu et al., 2008).

van Ingen and colleagues (2004) discovered a similar motif was responsible for PAH2 binding. The same group found that Sin3b's K165, above the hydrophobic cleft, could engage in a long range electrostatic interaction with the SID. This brings it close enough for the relatively short range hydrophobic interaction to occur. The lysine residue is conserved in both Sin3a's PAH1 (K134) and PAH2 (K315) domains, and to a lesser extent in PAH3 (R472), depicted in Figure 1.18.

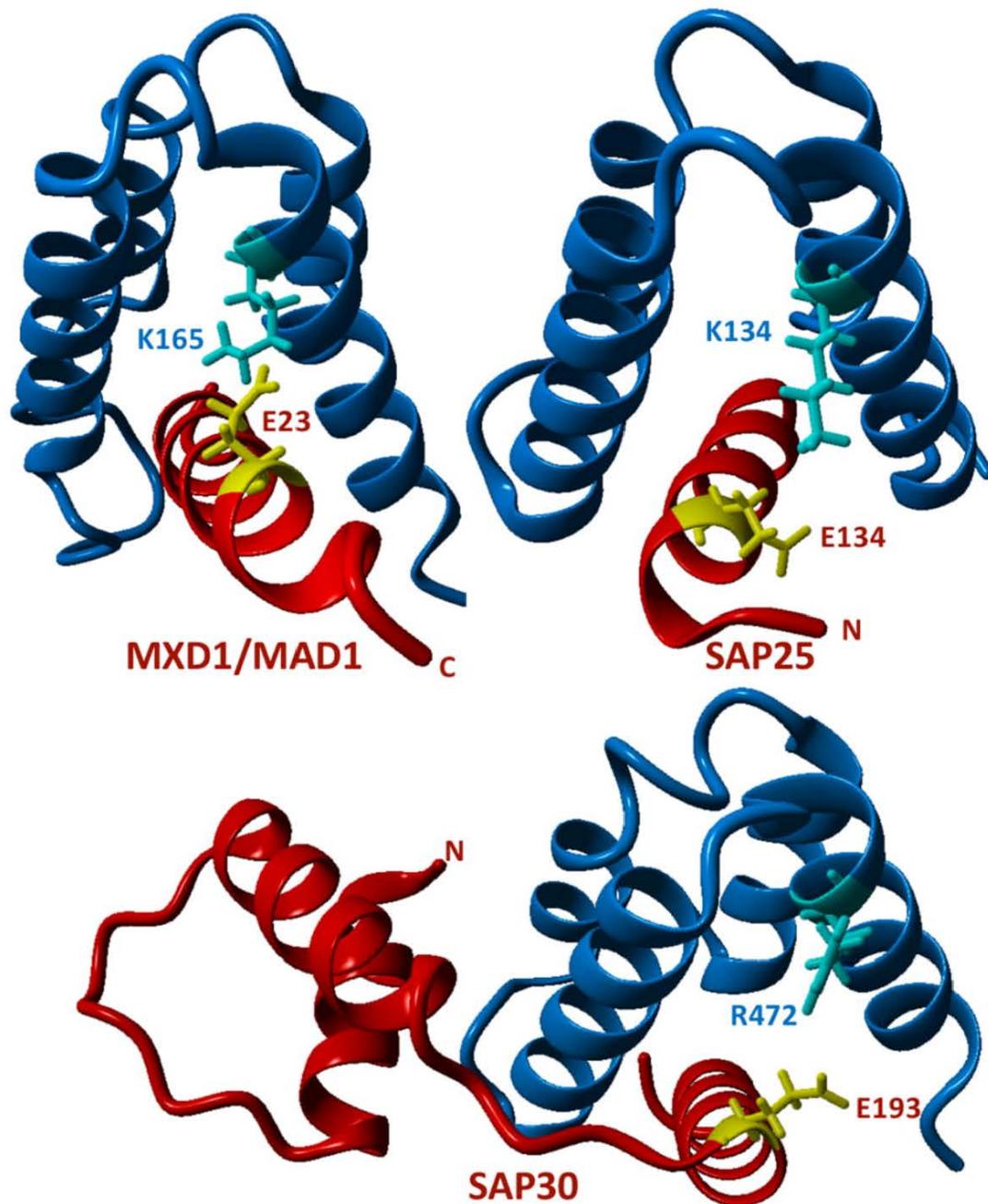


Figure 1.18: Possible Conserved Electrostatic Interaction.

The Conserved Lysine Resides Atop the Hydrophobic Cleft of (top right) PAH1 (K134) and (top left) PAH2 (K315) (Swanson et al., 2004, Sahu et al., 2008). Bottom: PAH3 also appears to have an analogous positive charge (R472) (Xie et al., 2011).

The most completely described PAH3 interaction is with the Sin3a Associated Protein 30 (SAP30), illustrated in Figure 1.19. The protein is part of the Sin3a, and HDAC, repressor complex. SAP30 is thought to serve a stabilizing function in this complex, in part because can directly interact with both SIN3 and HDAC1 (Zhang et al., 1998). The PAH3 hydrophobic pocket is shallower than those found in the more canonical PAH 1 and 2 domains. As a result, an additional hydrophobic surface, formed by the second and third alpha helix, of PAH3 must be bound for stable interaction (Xie et al., 2011). The canonical portion of SAP30's SID can be found towards its C-terminus. This alpha helix interacts with PAH3's hydrophobic pocket. Towards the middle of the SAP30 SID are two additional alpha helices which show a high degree of evolutionary conservation. These stabilize the interaction by wrapping around the side of PAH3 and engaging the PAH domain's hydrophobic residues (Xie et al., 2011).

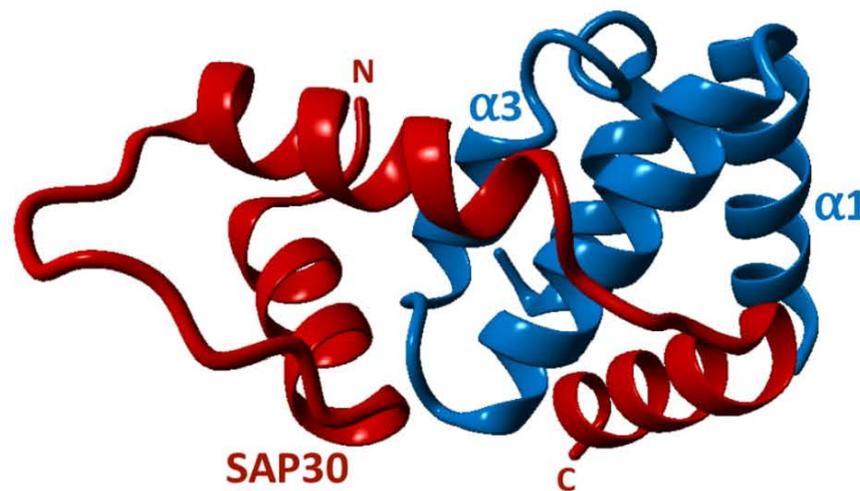


Figure 1.19: SAP30's Interaction with Sin3a's PAH3.

SAP30 SID (aa130–220) in red binding to PAH3 which is illustrated in blue (Xie et al., 2011).

1.10.2. PAH4 and the HCR

O-Linked N-Acetylglucosamine (GlcNAc) Transferase (OGT) is one of the few protein found to interact with PAH4 of Sin3a (Yang et al., 2002). As PAH4 is likely not a functional domain, the interaction may be indirect. The HCR sits on the extreme end of Sin3a immediately after PAH4, though like PAH4 it is poorly characterized. This C-

terminal proximal, 134 amino acid, stretch shares more than 80% amino acid identity with the corresponding region in Sin3b (Moehren et al., 2004). It is from this level of conservation that the highly conserved region derives its name. Alien/CSN2 is the only protein that is known to interact with this region. This interaction appears to occur in a manner that not only requires the HCR, but the HID as well (Moehren et al., 2004).

1.11. DNA Methylation

The histone modifications, regulated by proteins like Sin3a, illustrate only a portion of the epigenetic regulatory landscape. DNA modification plays an equally important role in epigenetic regulation. It can be placed on CpG dinucleotides and when present on a promoter is correlated with being transcriptionally repressive. This permits the long-term silencing of genomic loci that are potentially harmful or not needed in a particular cell type. An example of the former is in the control of transposable elements, which was initially thought to be the exclusive purpose of the epigenetic mark (Yoder et al., 1997). Much of the mammalian genome consists of transposable elements, which can ‘jump’ between different locations within the genome using an enzyme called transposase. Through spontaneous deamination which converts methylated cytosine into thymine, it was thought that DNA methylation promoted genomic stability by permanently degrading transposable elements (Yoder et al., 1997). Large scale DNA methylation occurs twice in development. The first occurrence is after embryo implantation and the second, at approximately e12, within the embryo’s primordial germ cells (Chen and Riggs, 2011). While it is true that a large portion of the wave’s targets are transposable elements, the epigenetic modification is also closely linked to cellular regulation (Molaro et al., 2014). Comparably, X-chromosome inactivation is largely maintained by DNA methylation (Singer-Sam and Riggs, 1993). DNA methylation within gametes is one of the drivers of genomic imprinting. This is the case regarding the paternally inherited *IGF2* gene. DNA methylation of *IGF2* contributed to the obesity observed in the offspring of males prenatally exposed to the *Hunger Winter*. The epigenetic mark is also implicated in numerous diseases. The dysregulation of DNA methylation is seen in many cancers, which often exhibit global hypomethylation,

resulting in genomic instability, coupled with hypermethylation of regions involved in tumor suppression (Kulis and Esteller, 2010, Ehrlich, 2002).

1.12. DNA Methylation Writers

Cytosine methylation is transmitted by a group of enzymes called DNA methyltransferases (DNMTs). The three major DNMTs are: DNMT1, DNMT3A and DNMT3B. DNMT2 and DNMT3L are also involved in cytosine methylation, but their roles are less conventional. The cytoplasmic DNMT2 is unusually specific in its activity, it methylates the tRNA corresponding to aspartic acid on cytosine 38 in the structure's anticodon loop (Goll et al., 2006). DNMT3L, while lacking innate enzymatic activity, stimulates the methyltransferase activity of DNMT3A and DNMT3B by interacting with their enzymatic domains (Gowher et al., 2005). DNMT3L, using its N-terminal cysteine-rich domain, can additionally bind to H3K4 and recruit its paralogues; though, only if the lysine residue is unmethylated (Ooi et al., 2007). DNMT3A and DNMT3B are responsible for de novo DNA methylation (Figure 1.20) – the methylation of cytosine residues that previously lacked the mark (Okano et al., 1999). De novo methylation is important in development and sets the stage for subsequent maintenance methylation. DNMT1 performs this maintenance role (Figure 1.20). Maintenance methylation occurs because DNA replication is semi-conservative. One strand of the original double helix is used to form each daughter strand (Meselson and Stahl, 1958). The methylation content is effectively halved by replication, as the newly synthesized strand lacks the epigenetic mark. DNMT1 ensures that DNA methylation patterns persist instead of being diluted out by each cycle of cell division (Stein et al., 1982). The roles of the DNMT enzymes are not as rigid as once thought. The DNMT3s appear to be necessary for the stable inheritance of DNA methylation marks, as cells with functional DNMT1, but lacking DNMT3A and DNMT3B, progressively lose their methylation (Chen et al., 2003). Similarly, a hemimethylation assay revealed that DNMT3A or DNMT3B were able to restore the methylation of L1 sequences, following transient DNA methyltransferase inhibition, while the maintenance DNMT1 was incapable of doing the same (Liang et al., 2002). Conversely, DNMT1 appears to demonstrate de novo DNA methylation activity, following DNMT3A's initiation of the

process, when the two enzymes are incubated together (Fatemi et al., 2002). DNMT1 therefore does contribute to de novo methylation, just as DNMT3A and DNMT3B provide a degree of maintenance methylation (Jin et al., 2011).

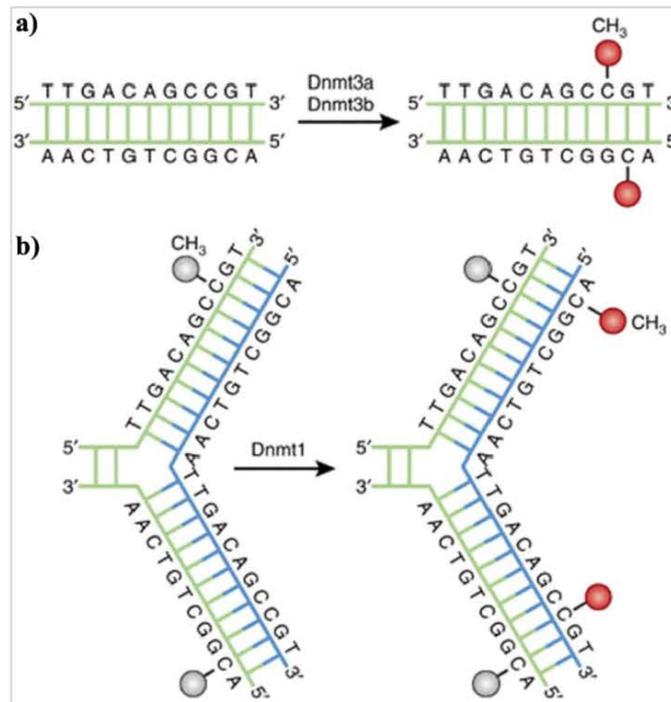


Figure 1.20: Canonical DNMT DNA Methylation.

a) De novo methylation by DNMT3a/b. b) DNMT1 maintaining methylation patterns on replicated DNA (Moore et al., 2013).

1.13. DNA Methylation Readers

The effects of DNA methylation are largely mediated by molecules that can bind to, or read, the epigenetic mark. This reading is the first step that permits the DNA methylation to have, downstream, consequence. There are three groups of reader motifs for cytosine methylation, these are: the zinc-finger domain, the ubiquitin-like with PHD and RING finger domains (UHRF) and the methyl-CpG-binding domain (MBD) (Moore et al., 2013). It can be argued that the first, the zinc-finger domain, is not entirely specific to methylated DNA. Kaiso, a zinc-finger domain protein, can also bind to non-methylated sequences with high-affinity (Daniel et al., 2002). MeCP2 was the

first protein discovered to have an MBD domain (Meehan et al., 1989). It is through these readers that DNA methylation often serves as a bridge to other epigenetic modifications. In the example of ARX silencing, MeCP2 binds to methylated regions of the *ARX* gene and recruits the PRMT6 histone methyltransferase which then methylates H3R2 (Dhawan et al., 2011).

1.14. DNA Methylation Erasers: the TET Proteins

The discovery of DNA demethylation echoed the discovery of the same process in histones; for a time it was thought that DNA methylation was an immutable modification (Bannister et al., 2002). The discovery of TET1, reminiscent of LSD1, led to the realization that DNA methylation could be reversed (Shi et al., 2004, Ito et al., 2011). While the majority of cellular demethylation is almost certainly passive, the activity of the three DNMTs is also counterbalanced by the TET enzymes, which permit a more targeted demethylation. The first of these proteins was discovered in a group of patients suffering from acute myeloid leukemia (AML), as part of a fusion protein with the mixed-lineage leukemia (MLL) histone methyltransferase (Hess, 2004). This t(10;11)(q22;q23) translocation was the driver of malignancy in these patients (Ono et al., 2002). The protein corresponding to 10q22 was named ten-eleven translocation 1, as it was the first protein in its family to be discovered. The three TET proteins resulted from two successive duplications of a single DNA demethylase in an ancestral jawed vertebrate (Iyer et al., 2009). The proteins share a similar layout, with three major domains (Figure 1.21). From the N-terminus, the first is a CXXC zinc finger domain, which is required for CpG binding. This is absent in TET2, which instead relies on the CXXC domain of another protein called IDAX which also negatively regulates the levels of TET2 (Ko et al., 2013). The next motif contained in the TET enzymes is a cysteine rich region which likely serves a role in DNA recognition, a theory informed by an understanding of the AlkB protein which has an analogous section (Iyer et al., 2009). The cysteine rich region is followed by the double-stranded β -helix (DSBH) fold of TET's 2OG-Fe(II) dioxygenase domain. The DSBH is the catalytic core of the enzyme which conducts the oxidation. It adds a hydroxyl group to 5-methylcytosine (5mC) residues. This 5-hydroxymethylcytosine (5hmC) mark can be oxidized further

into 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). These latter two DNA modifications are also the result of the TET enzymatic activity (Ito et al., 2011). The transcription factor SALL4A is thought to have a role in the further oxidation of 5hmC. SALL4A has been found to bind to 5hmC and through the stabilization of TET2, promotes 5hmC's oxidation into 5fC, eventually producing 5caC (Xiong et al., 2016). Remindful of DNA and histone methyltransferases' shared reliance on SAM; the enzymatic activity of both TET DNA demethylases and, Jumonji domain-containing, histone demethylases require the cofactors Fe^{2+} and α -ketoglutarate (Culhane and Cole, 2007).

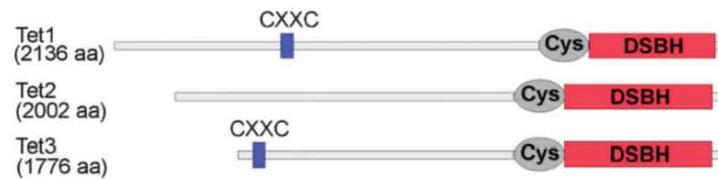


Figure 1.21: The Primary Sequence of the TET Enzymes.

Depicting the relative positions of the conserved CXXC, Cys-rich region and DSBH domains (Kriukiene et al., 2012).

While there are three different TET enzymes, they are not all concurrently expressed (Iyer et al., 2009). TET3 is most active shortly after fertilization where it facilitates a wave of paternal genome demethylation (Gu et al., 2011). The high levels present in the oocyte and the zygote rapidly decreases from the two-cell stage onwards (Iqbal et al., 2011). Following this, there is a wave of de-novo methylation that takes place after implantation (Santos et al., 2002). The inner cell mass of the blastocysts, correspondingly, sees TET1 and TET2 expression increase (Dawlaty et al., 2011). Expectedly, in embryonic stem cells (ESC) it is TET1 and TET2, but not TET3 that are expressed (Ito et al., 2010). The comparative expression levels of the three TET enzymes is illustrated in Figure 1.22.

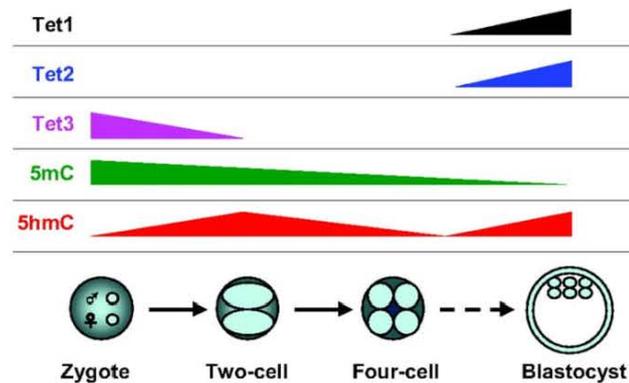


Figure 1.22: TET Expression During Development and the Corresponding DNA Methylation/Hydroxymethylation Levels (Tan and Shi, 2012).

TET1 is the dominantly expressed demethylase in ESCs, as TET2 only makes up less than 17% of the TET enzyme pool in these cells (Koh et al., 2011). There are indications, that each of these two TET proteins can compensate for the loss of the other. The knockout of either gene can produce viable mice, but the concurrent loss of both genes results in a high degree of perinatal lethality (Dawlaty et al., 2011, Li et al., 2011, Dawlaty et al., 2013). Despite their organizational similarity, enzymatic activity and shared functions, the purview of the two proteins does diverge slightly. Enzyme knockdowns reveal that TET1 appears to employ its enzymatic function primarily towards transcription start sites and promoter regions, conversely TET2 seems to oxidize methylcytosine within gene bodies (Huang et al., 2014). TET1 and TET2 were found to have partially opposing functions in ESCs. TET1 was discovered to interact with zinc finger protein 281 which acts to reduce TET2 expression, resulting in a state of pluripotency that is not dissimilar to post-implantation epiblast cells, whereas TET2 expression yields a state of pluripotency that is comparable to what is found in the inner cell mass of the blastocyst (Fidalgo et al., 2016).

1.14.1. DNA Demethylation

There are two proposed models of DNA demethylation which feature the iterative oxidation of methylcytosine. Unlike 5hmC, both 5fC and 5caC can be excised by thymine DNA glycosylase (TDG) and replaced with an unmodified cytosine due to the activity of base excision repair (BER) (Maiti and Drohat, 2011). This theory has several

limitations; the first is that an over-reliance on TDG and BER would compromise genomic stability if the mechanism were used on a global scale (Wu and Zhang, 2011, Pastor et al., 2013). The second is that the knockdown of TDG does not cause levels of 5caC to rise to those of the 5fC precursor as would be expected (Wu and Zhang, 2011). This either implies that TDG is not the only enzyme that processes 5caC, or that the model is otherwise flawed. Ito and colleagues (2011), who discovered 5fC and 5caC, suggested a demethylation model that was also based on iterative oxidation. Their theory sought to address the perceived flaws in the iterative oxidation TDG/BER mechanism. This version of DNA demethylation obviated the need for both TDG and BER, by using an as yet undiscovered decarboxylase. This decarboxylase, performs its eponymous function and converts 5caC to C (Ito et al., 2011). These two demethylation, by iterative oxidation, reactions are illustrated in Figure 1.23.

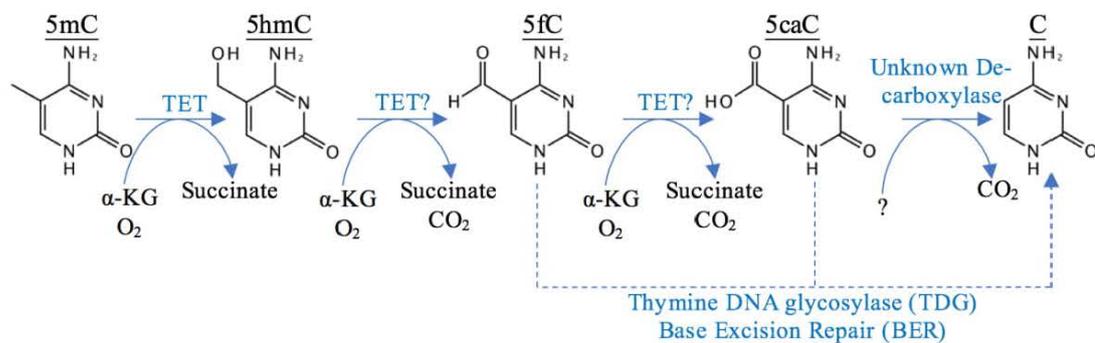


Figure 1.23: TET Mediated DNA Demethylation by Iterative Oxidation.

A demethylation process with greater support is the passive model, whereby the presence of the 5hmC mark may reduce the activity of DNMTs (Hashimoto et al., 2012). Cellular division would at this point serve to dilute out the methylation mark passively. Detracting from this theory, a stably expressed plasmid containing hydroxymethylcytosine was not seen to undergo demethylation through successive rounds of replication (Kubosaki et al., 2012). It is additionally possible that the mere presence of TET1 at sites on chromatin sterically inhibits DNMT function. This is suggested by the simple correlation that that the promoters of genes that are not bound by TET1 tend to be methylated, and that TET1 bound promoters display the same methylation following TET1 knockdown (Wu et al., 2011b). This elegantly mirrors the

mechanism through which TET1, and consequently its enzymatic activity, are impeded by MBD2 and MECP2 (Ludwig et al., 2017). The two proteins bind to DNA, though preferentially to the methylated variety, and physically prevent TET1 from associating with the same regions (Ludwig et al., 2017, Hansen et al., 2010). MBD2 can additionally oppose DNA demethylation through the recruitment of histone deacetylase (HDAC) proteins which promote DNA compaction (Baubec et al., 2013).

1.15. Crosstalk Between DNA and Histone Epigenetic Marks and their Writers

Many of the epigenetic activities do not function in isolation, but rather work in unison to reinforce an epigenetic outcome. In the example of beta cell formation, the silencing of the *ARX* gene is initiated by the DNA methyltransferases. DNMT1 and DNMT3A methylate regions of the *ARX* gene; the methylated DNA serves as a binding site for MeCP2, which is then able to recruit the PRMT6 histone methyltransferase (Dhawan et al., 2011, Papizan et al., 2011). PRMT6 dimethylates H3R2 and this prevents the MLL-family methyltransferase complex from methylating H3K4 (Ringel et al., 2015, Guccione et al., 2007). As H3K4me3 is a binding site for the, chromodomain containing, MYST family of acetyltransferases, this further reduces local histone acetylation. The absence of H3K4 methylation additionally permits DNMT3L to recruit DNMT3A and DNMT3B to this site, reinforcing DNA methylation levels (Ooi et al., 2007). Histone acetylation is also depressed by HDAC1 activity, which is recruited to the DNA by DNMT3A (Fuks et al., 2001, Papizan et al., 2011). This, HDAC induced, loss of H3K9 and H3K14 acetylation, is compounded by the fact that the latter mark serves as a chromodomain binding site – preventing the recruitment of the CBP/EP300 family of histone acetyltransferases (Dhawan et al., 2011, Dhalluin et al., 1999).

Just as DNA methylation and histone methylation are linked through MeCP2 and DNMT3L, TDG may additionally connect DNA demethylation to histone acylation, as it is known to form a complex with the acyl transferase EP300 (Tini et al., 2002). In the same vein, there are multiple lines of evidence that hint towards the recruitment of EP300 by TET1 activity/inactivity (Tini et al., 2002, Gelman et al., 1999, Arany et al., 1996, Kwon et al., 2012). Song et al. (2013) propose that through the generation of 5fC,

TET1 may coordinate the association of EP300 with poised enhancers. This is based on the observation that the number of EP300 sites increased by nearly 2.5-fold in TDG^{-/-} ESCs, and that the presence of EP300 at these poised genes was correlated with a commensurate accumulation of the 5fC mark.

1.16. TET1's Biological Relevance

The discovery of the TET enzymes revealed DNA methylation to be a dynamic process. Following this, a great body of work suggested that TET1 played a large role in embryonic stem cells and mammalian development. This idea was supported by findings that TET1 plays an important role in the production of iPSCs. TET1 was discovered to be able to replace OCT4 as one of the four Yamanaka factors in OCT4, KLF4, SOX2, and C-MYC (OKSM) adult cell reprogramming (Gao et al., 2013). Adding to the protein's import, Ito and colleagues found that TET1 played a role in functioning of NANOG, a transcription factor and marker of stemness, and consequently stem cell maintenance (Ito et al., 2010). The findings of this last study were disputed by Dawlaty et al. (2011) and Williams and colleagues (2011), but supported by Saunders and co-workers (2017). While being otherwise phenotypically normal, a significant portion of the *TET1* KO mice produced in the Jaenisch lab had observable developmental delays (Dawlaty et al., 2011). This partial penetrance was also observed by Khoueiry et al. (2017), though they revealed that it was due to the use of inbred mouse models. Non-inbred mice more clearly demonstrated the emergence of lethal embryonic defects during one of the final stages of gastrulation, resulting from a loss of TET1. TET1 plays a similarly important role in the progression that a stem cell makes towards differentiating into an adult cell. This is hinted at in the function of DNMT1 which is often referred to as the 'maintenance DNA methyltransferase'; TET1 may similarly be a 'maintenance DNA demethylase'. This is given weight by the observation that in differentiated cells, TET1 prevents the spread of ectopic methylation from occurring at unmethylated CpG islands (Jin et al., 2014). It was further supported by the discovery that the absence of both TET1 and TET2, during differentiation, caused hypermethylation in large > 3.5kb regions of the genome that typically contained

low average methylation. These regions were appropriately named methylation canyons (Jeong et al., 2014, Wiehle et al., 2015).

While it is involved in preventing aberrant DNA methylation, TET1 appears to function as a net transcriptional repressor. A ChIP experiment was used to determine the genes that were direct targets of TET1 (Williams et al., 2011). Williams and colleagues (2011) found that a shRNA knockdown of TET1 resulted in more of these target genes being upregulated (495) instead of downregulated (195) in ESCs. Given TET1's enzymatic role, it is surprising that the DNA demethylase directly repressed ~2.5 times more genes than it directly upregulated, as illustrated in Figure 1.24 (Williams et al., 2011). TET1 is an enzyme with dual opposing activities and it likely mediates these contextually, reminiscent of MBD3's function within NuRD, discussed in chapter 1.8.3.

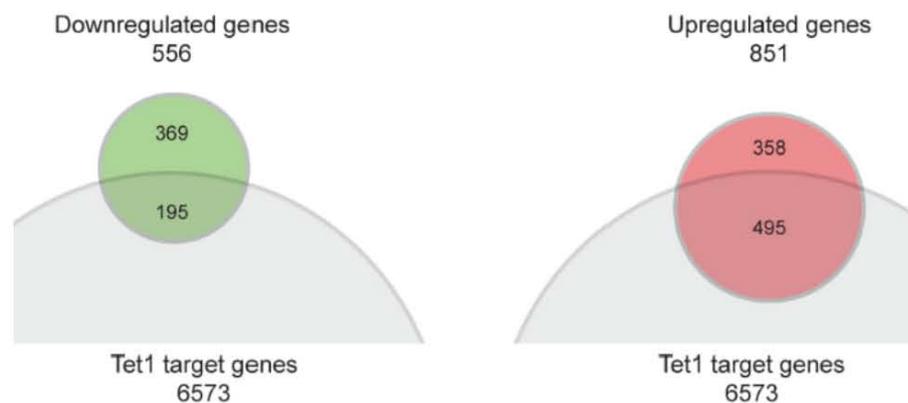


Figure 1.24: TET1 is a Net Transcriptional Repressor.

TET1 shRNA knockdown using a microarray presented with ChIP data regarding TET1 binding sites (Williams et al., 2011).

TET1's influence on gene expression transcends its enzymatic activity. Downregulation of TET1 in ESCs and in cells with a triple knockout (TKO) of *DNMT1*, *DNMT3A* and *DNMT3B* were compared; and it was found that even in the absence of its substrate, TET1 still maintains its repressive effects on gene expression (Williams et al., 2011). Intriguingly, this repression may be the result of a separate mechanism whereby TET1 influences complexes involved in histone modification. Like OGT, TET1 can be found in two discrete pools. The first distribution of TET1 coincides with Suppressor of Zeste

12 homolog (Suz12) and H3K27me3. This mark is deposited by the Polycomb repressive complex 2 (PRC2) of which Suz12 is a constituent member (Williams et al., 2011, Neri et al., 2013). The second grouping of TET1 co-localizes with Sin3a, the protein at the center of a histone deacetylase complex, around transcription start sites in a manner that does not coincide with either Suz12 or H3K27me3 (Neri et al., 2013).

1.16.1. An Expanded Repressor Complex

While both TET1 and OGT appear to be involved with elements of the PRC2 complex, these interactions are unlikely to be synchronous (Vella et al., 2013, Schmitges et al., 2011, Wu et al., 2011b). Despite this, an ESC megadalton complex between TET1 and OGT does exist. Though instead of with PRC2, this occurs through Sin3a as demonstrated through a series of co-immunoprecipitations (Vella et al., 2013). Vella et al. (2013) found two populations of OGT complexes; one with host cell factor C1 (HCFC1) and the other with TET1 and Sin3a. This latter complex was found to interact with TET2, albeit weakly. HCFC1 is part of the SET1/COMPASS complex that trimethylates H3K4, a process partly regulated by TET2 (Deplus et al., 2013). The proposed TET1:OGT:Sin3a binding has the potential to take several different conformations as each of the three proteins is capable of directly interacting with the others, illustrated in Figure 1.25. Though in three dimensions it is possible that this could result in a single configuration.

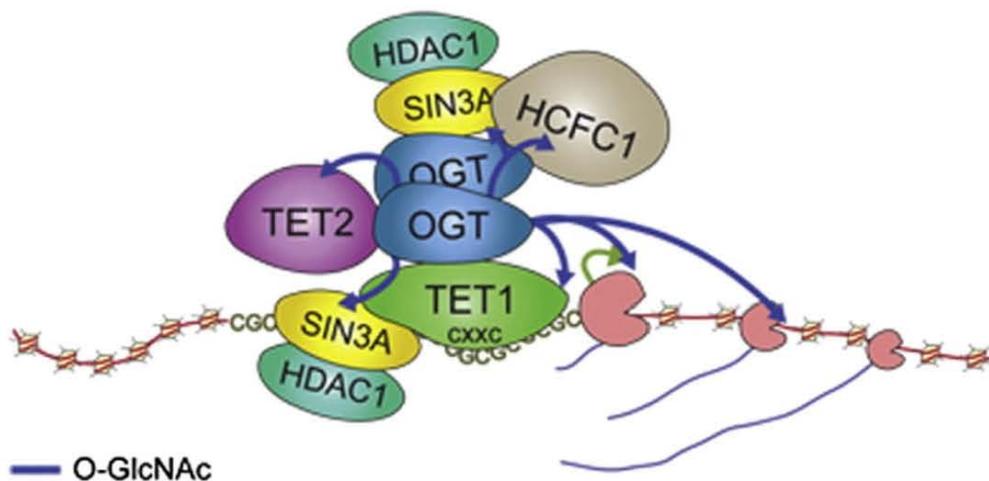


Figure 1.25: The Possible Configurations of OGT Complexes (Vella et al., 2013).

Dr. Adrian Bracken's group corroborated this result, and found that the stability of this complex required the presence of Fam60a (Streubel et al., 2017). This interaction between TET1, OGT and Sin3a includes many of the components of the canonical HDAC repressor complex, including HDAC1 (Streubel et al., 2017). This single complex directly weaves together an epigenetic triumvirate of DNA demethylation, O-GlcNAcylation and histone deacetylation (Figure 1.26).

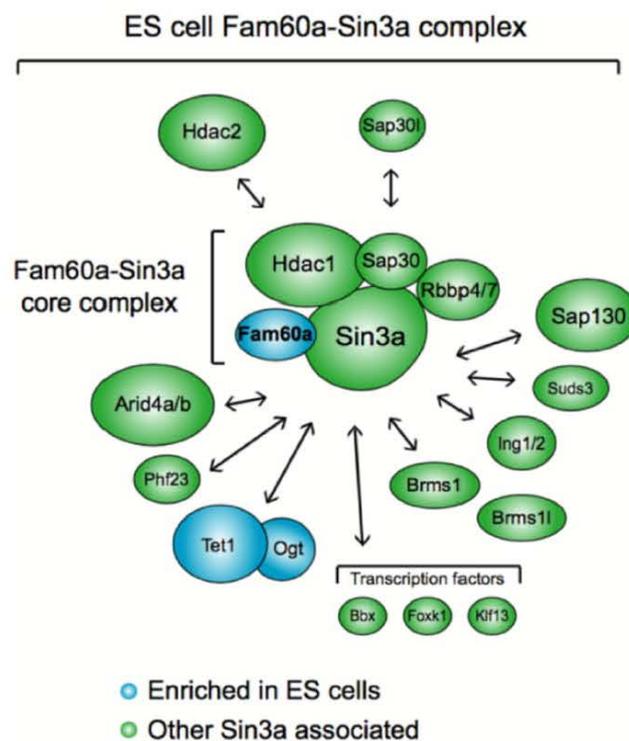


Figure 1.26: The Fam60:TET1:OGT:Sin3a Complex (Streubel et al., 2017).

1.17. Aims

At face value the Sin3a co-repressor complex protein and the DNA demethylase enzyme, TET1, have roles in gene expression that appear opposed. On closer inspection, TET1 maintains dualistic epigenetic functions. While it demethylates DNA, the enzyme functions as a net transcriptional repressor (Williams et al., 2011). This latter activity is independent of TET1's enzymatic activity; but instead, is thought to derive from TET1's interaction with Sin3a. While it is known that TET1 recruits Sin3a to a significant number of its targets and that Sin3a plays a transcriptionally repressive

role in the co-repressor complex, these two phenomena have yet to be meaningfully linked (Williams et al., 2011, Vidal et al., 1990). This thesis will delve into the broader meaning of the TET1:Sin3a interaction by demonstrating that TET1's gene repression derives from its ability to bind to Sin3a. Unlike the TET1:OGT interaction, which has been mapped to 45 residues on TET1's C-terminal end, the physical association between TET1 and Sin3a has not been adequately described (Hrit et al., 2017). Addressing this issue, by determining the minimal regions of Sin3a and TET1 that are responsible for the TET1:Sin3a interaction, will be the primary focus of this project. This thesis has the additional aim of investigating lysine de-crotonylation. While the mark appears resistant to removal, as SIRT1-3 demonstrate, it is reversible (Montellier et al., 2012, Bao et al., 2014). Although it remains to be determined if the TET1:Sin3a complex plays a role in decrotonylation; the ubiquity and localization of the repressor complex components, HDAC1/2, suggest that they may act in this capacity (Jamaladdin et al., 2014, Heidemann, 2015). This project seeks to illustrate the role of, the classical class I HDACs, HDAC1/2 in lysine decrotonylation.

Chapter 2. Materials and Methods

2.1. Agarose Gel-Electrophoresis

A 1.5% agarose gel was made using 1.5g of agarose powder mixed with 100mL of 1×TAE, microwaved until molten (2min 40sec). When $< 60^{\circ}\text{C}$, 3 μL of Ethidium Bromide (EtBr) was added and mixed into the solution. Samples were run at 100V for half an hour. A UV trans-illuminator was then used to visualize the EtBr that intercalated into the nucleic acid samples.

2.2. Cell Culture

M15+LIF media was made, for E14 and E14 derived ESCs, from a mixture of 83% Knock-Out Dulbecco's Modified Eagle's Medium (DMEM) [ThermoFisher, Cat: 10829018], 15% Fetal Bovine Serum (FBS), 1% Penicillin/Streptomycin/L-glutamine [ThermoFisher, Cat: 10378016], 5mM β -mercaptoethanol [Sigma-Aldrich, Cat: M6250] and 1000U/ml Leukemia Inhibitory Factor (LIF).

M10 media was made, for HEK293 cells, using 89% DMEM [ThermoFisher, Cat: 11965092], 10% Fetal Bovine Serum and 1% Penicillin/Streptomycin/L-glutamine.

Cells were grown on 6, 10 and 15cm cell culture plates as needed. These were incubated with 0.1% gelatin PBS solution spread across the working surface of the plate at room temperature, for a minimum of 15 minutes. Following this, the gelatin was aspirated off and was replaced with warm media. The media was M15+LIF for ESCs and M10 for HEK293T cells. The appropriate number of cells were added to the media and the mixture was gently triturated to spread the cells evenly. When the cells were evenly spread, the plates were stored at 37°C with 5% CO_2 in a mammalian cell incubator. The media on the plates containing the ESCs was aspirated off and replaced every day, while the M10 media for the HEK cells was removed every other day or sooner if the phenol red indicator within the DMEM signified increased acidity.

2.2.1. Cryostorage

Cryostorage media was made of 60% M10/M15+LIF, 40% FBS and 10% Dimethyl sulfoxide (DMSO) [Sigma-Aldrich, Cat: D2650-100mL].

The cells underwent the passaging protocol detailed in chapter 2.2.2. When they were detached and washed, the cells were re-suspended in cryostorage media and aliquoted into cryostorage vials [Corning, Cat: 430488] such that $\sim 2 \times 10^6$ cells were distributed per vial. -80°C was reached using a freezer and a Mr. Frosty™ Freezing Container [ThermoFisher, Cat: 5100-0001], filled with room temperature isopropanol. After three hours, the vials were transferred into a pre-cooled polycarbonate cryostorage box [MedSupply Partners, Cat: CY-PB2100]. The box was then placed into an aluminum rack which was immersed in liquid nitrogen within a cryogenic storage system.

2.2.1.1. Revival from Cryostorage

The vial with the desired cells was retrieved from the polycarbonate cryostorage box which was stored inside liquid nitrogen. When everything was replaced the vial was warmed by hand until the contents were liquid; they were then diluted in warm M10/M15+LIF. The cells were centrifuged at 250rcf and the media was removed, to diminish the DMSO content. The cells were then plated on a gelatinized 10cm plate with 15mL of media, and were permitted to recuperate in the incubator for 24 hours.

2.2.2. Passaging

Cells were split when $\sim 75\%$ confluent. Plates were washed with an appropriate volume of Dulbecco's phosphate-buffered saline (DPBS) [ThermoFisher, Cat: 14190169], following media removal. Cell dissociation buffer, applied to the cells for $\sim 3\text{min}$, was a mixture of 75% DPBS and 25% of Accutase Cell Dissociation Reagent [ThermoFisher, Cat: A1110501]. An equal volume of M10/M15+LIF was added to terminate the trypsin/Accutase activity. Cells were collected in a 15mL falcon tube and centrifuged for 5min at 250rcf. They were re-suspended in 2mL of media and the desired number

were plated on a gelatinized plate containing the appropriate media. The plate was then stored at 37°C with 5% CO₂ in a mammalian cell incubator.

2.3. Cloning

All Infusion cloning was conducted by Dr. Xiaowen Yang and Dipti Vashi of the University of Leicester's PROTEX Service. The pertinent primers and gene blocks are listed in Table 7.2.

2.4. Co-Immunoprecipitation Using IgG antibody and Sepharose Beads

IP buffer contains 250mM NaCl [Fisher Scientific, Cat: BP358-1], 10% Glycerol [Fisher Scientific, Cat: BP229-1], 0.5% IEGPAL [Sigma-Aldrich, Cat: 18896-50mL] and 50mM Tris [Fisher Scientific, Cat: BP152-5] at pH 7.5. 250mM of salt was used for investigating interactions that are not ionic, but hydrophobic. 150mM would be more appropriate for the former.

500µL of the buffer was needed for each of the samples being tested. Protease Inhibitor Cocktail Set III [Calbiochem, Cat: 535140-1SML] was added to this at 1:100. An additional 50mL of IP buffer without the Protease inhibitor was required for the wash steps.

Cells were grown on 10cm plates [Corning, Cat:430167-500EA] and then lipofected. These were grown for 1-2 days to allow for protein expression. 15cm plates [Thermo Scientific, Cat: 157150] were used to culture PiggyBac/endogenously produced proteins. The media was changed as needed until the harvest. Prior to this, the bottle of Protein G Sepharose Beads [GE Healthcare Life Sciences, Cat: 17-0618-01] was washed with 20% ethanol if new.

Cells were washed with 5-10mL of phosphate buffered saline (PBS) [Fisher Scientific, Cat: BP399-20] and harvested with a cell lifter [Fisher Scientific, Cat: 08-100-240] in 1mL PBS, permitting its transfer to a 1.5mL Eppendorf tube [Corning-Axygen, Cat:

MCT-150-C]. The cells were pelleted at 250rcf for 3min using an Eppendorf centrifuge [Eppendorf, Model: 5415D] and the PBS was aspirated. 430 μ L of IP buffer with protease inhibitor was added, and the cells were sheared mechanically by pipette trituration. This lysate was then incubated at 4°C for 20min on a 360 Degree Vertical Eppendorf tube Rotator [Grant Bio, Model: PTR-30]. The lysate tubes were then centrifuged at 25,000rcf for 25 minutes, at 4°C, using a microcentrifuge [Eppendorf, Model: 5417R].

80 μ L of sepharose bead slurry was immunized per sample, half for the IP and the remainder for the negative control. The beads were then washed thrice with 1mL IP buffer without protease inhibitor. 500 μ L IP buffer was added and the beads were divided between two Eppendorf tubes. The IP tube received 1 μ g of relevant IgG antibody/sample, an example was anti-mouse Flag [Sigma-Aldrich, Cat: F1804] unless otherwise indicated by the manufacturer. The negative control tube was incubated with an equivalent amount of normal mouse IgG [Santa Cruz, Cat: SC-2025]. This was also placed on the 360 Degree Vertical Eppendorf tube Rotator at 4°C for 20min. The beads were subsequently washed 3 times with IP buffer and were then aliquoted into the required number of Eppendorfs (a negative and IP tube per sample).

Protein concentrations in the samples were estimated and normalized, with IP buffer, using a Bradford assay. The amount of buffer to be added to the lysate sample was calculated using the following formula:

$$\left(\frac{\text{measured protein concentration of sample}}{\text{desired concentration}} - 1 \right) \times 430\mu\text{L}$$

25 μ L of each sample was retained for the input (positive control). 200 μ L was used for each the IP and negative control. The bead/lysate admixture was incubated at 4°C for 6h/overnight on a rotator. Samples were washed thrice in IP buffer and the supernatant was aspirated. A consistent amount of 4 \times LDS sample buffer [Fisher Scientific, Cat: NP0007] with 0.375M DTT [Melford, Cat: MB1015] was added to all the tubes. These were then heated to 95°C for 5min and the beads were compacted by > 5,000rcf centrifugation. The samples were then used in a western blot.

2.5. Western Blot

2.5.1. Whole Cell Extract Preparation

The whole cell extract (WCE), or compacted bead co-IP, is used for the western blot. If it is not already prepared, WCE can be prepared as follows:

Co-IP buffer was made (with 1 μ L/mL of TSA for histone extraction). Protease inhibitor cocktail set III [Calbiochem, Cat:535140-1SML] was added to the buffer in a 10 μ L/mL ratio. 400-500 μ L of the buffer was used to re-suspend to the cell pellet derived from a 10cm plate. Cells were sheared by pipette trituration, and the lysate was incubated at 4°C for 20min on a rotator. The supernatant was extracted via centrifugation at 4°C and 25,000rcf for 25 minutes. The concentration of protein whole cell extract was normalized using a Bradford assay. The pellet was retained for histone extraction.

Histone extraction required the maceration of the pellet with a pipette tip and its subsequent overnight 4°C rotational incubation in 100 μ L of 0.2M (0.4N) H₂SO₄.

2.5.2. SDS-PAGE

All samples were incubated at 95°C for 5min in 20 μ L of LDS buffer with DTT, after which they were centrifuged at > 5,000rcf for 3min. 10 and 20 sample PAGEs were run in XCell SureLock™ Mini-Cell tank [Life Technologies, Cat: EI0001] and XCell4 SureLock™ Midi-Cell Electrophoresis System [Fisher Scientific, Cat: WR0100], respectively. These used 10 well NuPAGE® Novex® 4-12% Bis-Tris Gels [Fisher Scientific, Cat: NP0321BOX] and 20 well NuPAGE® Novex® 4-12% [Fisher Scientific, Cat: WG1402BX10], respectively. Samples were loaded (10 μ L input, 20 μ L negative and 20 μ L IP) along with ~7 μ L PageRuler™ Plus Prestained Protein Ladder [Fisher Scientific, Cat: 26619].

Gels were run at 100V for 2h, in MES SDS running buffer [Fisher Scientific, Cat: NP0002-02]. 1mL of NuPAGE Antioxidant [Fisher Scientific, Cat: NP0005] was added between the gel plates, into the MES buffer. MOPS buffer might be better for high molecular weight proteins.

2.5.3. Gel to Membrane Transfer (Blotting)

Transfer buffer consisted of TG buffer [BioRad, Cat: 161-0771] with 10% methanol and 0.125% SDS [Fisher Scientific, Cat: AM9820]. The buffer was not recycled due to pH changes. The transfer to nitrocellulose membrane [GE Healthcare Life Sciences, Cat: 10600003] or PVDF was conducted at 700mA for 3 hours, or for low molecular weight proteins (< 100KDa) – 100V for 1 hour, in a Criterion™ Blotter [BioRad, Cat: 170-4070] placed in an ice bath.

2.5.4. Western Blot Immunizing and Imaging

The membrane was blocked, for one hour, with Odyssey® Blocking Buffer [LI-COR, Cat: 927-40003]. Primary antibody was applied after being diluted 1:1000 in blocking buffer (unless otherwise stated – see Table 7.9). The membrane was washed thrice in 50mL of PBS [Fisher Scientific, Cat: BP399-20], with 0.1% Tween 20 [Sigma-Aldrich, Cat: P1379-1L]. The secondary IRDye® 800CW goat anti-mouse [LICOR, Cat: 926-32210] and IRDye® 680LT donkey anti-rabbit [LICOR, Cat: 926-68023] antibodies, were then applied in a 1:10,000 dilution. The membrane was then washed thrice as before, though the last time without Tween 20. An Odyssey® CLx Imaging System [LICOR, Model: 9140] and Image Studio™ Software for Odyssey CLx [LICOR, Cat: 9140-500] was then used for analysis.

2.6. Crystallography Procedures

Protein concentration was estimated using two methods. The Bradford assay (595nm) measured the concentration based primarily on the presence of arginine and, to a lesser extent – histidine, lysine, tyrosine tryptophan and phenylalanine (Compton and Jones,

1985). The UV280nm reading, combined with extinction coefficients, measured protein concentration on the basis of tryptophan, tyrosine and phenylalanine (in decreasing order) content. Crystallography trials used TTP Labtech's Mosquito LCP robot to deposit the His-Sin3a PAH1+GB1-His-TET1 878-911 protein mixture across the different crystallization screens.

Dr. Nicola Portolano helped set up the crystallization trays and, unassisted, screened the different conditions for crystals.

2.7. *E. Coli* Transformation and Protein Production

25 μ L of competent *E. Coli* cells were transformed with 100-500ng of DNA plasmid via heat shock (42°C for 30s). Transformed cells were rescued in media for an hour and then selected for by ampicillin (100mg/mL) resistance. At an OD600nm of 0.6 the *E. Coli* were induced into protein production through the addition of 0.2mM Isopropyl β -D-1-thiogalactopyranoside (IPTG). They were incubated at 37°C for 3 hours, or 20°C for 10 hours for NMR, on a shaking incubator. The cells were pelleted and stored at -80°C until needed.

2.8. Gal4 VP16 Reporter Assay

The cell lysis buffer was 6.25mM Tris at pH 7.8, 10mM DTT, 10mM EDTA, 50% v/v Glycerol and 5% v/v Triton X-100.

β -gal stock solution was 60mM Na₂HPO₄ [Fisher Scientific, Cat: S472-500], 40mM NaH₂PO₄ [Fisher Scientific, Cat: 10723621], 10mM KCl [Fisher Scientific, Cat: 10375810] and 1mM MgCl₂ [Sigma-Aldrich, Cat: M8266-1KG].

Maxipreps [Fisher Scientific, Cat: K0491] of the VP16-LexA DNA binding domain, pGL2-Lex-Gal-Luc (LexA response element–Gal4 response element–Luciferase) and β -galactosidase plasmids were made. Minipreps [Qiagen, Cat: 27106] of the Gal4 DNA binding domain (DBD) fusions were produced, these included: MAD 1-35 (MAD-SID),

TET1 878-911, TET1 878-911 I894A, TET1 878-911 L897A, TET1 878-911 T898A and TET1 878-911 L900A. HEK293T cells were grown in 48 well plate [Fisher Scientific, Cat: 150687] to a density of 50,000 cells per well. A mastermix of 0.18 μ g/well of β -galactosidase, 0.1 μ g/well Gal4DBD fusion and 0.23 μ g/well of pGL2-Lex-Gal-Luc plasmids were made in 500 μ L Opti-MEM Reduced Serum Media [Fisher Scientific, Cat: 1985062]. These were transfected, by Transit X2 [Mirus Bio, Cat: MIR 6006], into the HEK293 cells (in triplicate). After 2 days the cells were washed with PBS and then 140 μ L of lysis buffer/well was added and this is incubated for 2 hours on a shaker at room temperature. This plate was then stored at -70°C.

2.8.1. β -gal Activity Measurement

Substrate A and B, from a Biovision luciferase reporter kit [BioVision, Cat: K801-200], were reconstituted as per the manufacturer's instructions. 12mg of macerated o-nitrophenyl- β -galactopyranoside (ONPG) substrate and 21 μ L of β -mercaptoethanol (BME) were added to 6mL of β -gal stock solution. 80 μ L of lysate/well was used for the β -gal normalization in a transparent 96 well plate. Three wells were reserved for only lysis buffer alone. 100 μ L of the β -gal stock solution, with ONPG and BME, was added to all wells. The plate was placed on a 37°C shaking incubator for 2 min. The plate was then placed in a PerkinElmer Victor X5 multi-label reader with dispenser option, and the 405nm emissions were read.

2.8.2. Luciferase Activity Measurement

An opaque 15mL tube filled with 100 μ L/well of solution B was piped into the dispenser portion of the PerkinElmer Victor X5. 20 μ L of lysate was added to the wells of an opaque white 96 well plate, with three being reserved for lysis buffer alone. Solution B was dispensed into the plate and the luminescence reading was taken. The reaction that was measured is described in Figure 2.1.

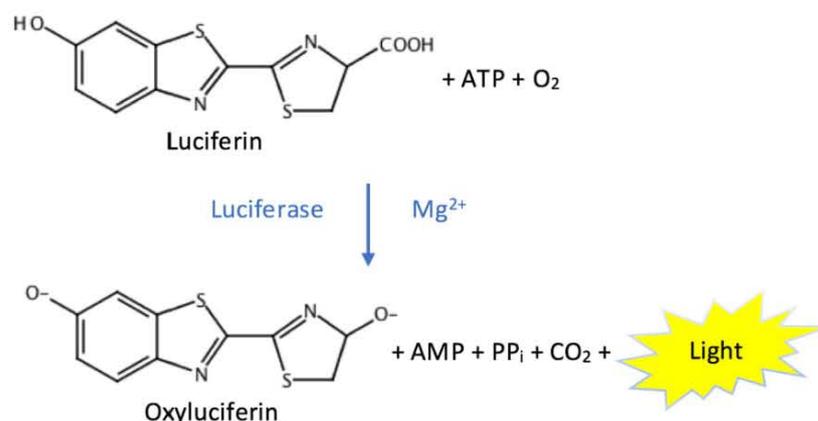


Figure 2.1: Measuring Luminescence Indirectly Permits the Quantification of Relative Luciferase Activity.

2.9. GST-Pulldown and IVTT

NETN buffer was 0.5% IEGPAL, 20mM Tris at pH 8 and 150mM NaCl.

The GST-fusion protein bait was produced as per 2.7. The pellet was re-suspended in (~30mL) NETN buffer and the cells were sheared by pipette trituration. GST beads were washed with NETN thrice and then incubated with the supernatant lysate on a rotator at 4°C for 60min.

The *In Vitro* Transcription and Translation (IVTT) TnT[®] Coupled Reticulocyte Lysate System [Promega, Cat: L5020] used: (5μL rabbit reticulocyte lysate [kit], 1μL amino acid mixture lacking methionine [kit], 1μL SP6 or T7 polymerase/reaction [kit], 1μL/reaction S³⁵ methionine, 1μg template and sterile water to 50μL)/reaction. This was incubated at 30°C for 90 minutes. 1 μL of each IVTT sample was kept for the input.

Beads were washed thrice and then re-suspend in 1-5mL NETN and incubated with 5μL of the appropriate S³⁵ IVTT reaction, for two hours at 4°C, on a rotator. The tubes were washed thrice with NETN. The samples were then boiled in LDS sample buffer with DTT for three minutes. The IVTT inputs and samples were then run on an SDS-PAGE. The results were previewed using instant blue. The gel was placed on thick blotting paper and covered with Clingfilm and dried [BioRad, Cat:1651746] for ≥ 4 hours. A,

fluorescent light, blanked storage phosphor screen was then placed atop the dried gel. The duration (hours to days) of this changed depending on the age of the S³⁵ Methionine used. A Typhoon Imager set to Storage Phosphor, Phosphor Mode: Best Sensitivity, Orientation – set to automatic and Pixel Size: 200 microns (Optimally > 800, in the future) acquired the image. Image Quant software was used to orient and crop the image.

2.10. HDAC Deacylation Assay

HDAC buffer was 50mM NaCl [Fisher Scientific, Cat: BP358-1] and 50mM Tris [Fisher Scientific, Cat: BP152-5] at pH 7.5.

HDAC lysis buffer was 50mM NaCl, 50mM Tris at pH 7.5. 5% Glycerol [Fisher Scientific, Cat: BP229-1], 0.3% Triton [Sigma-Aldrich, Cat: X100-500ML] and a fresh protease inhibitor tablet [Sigma-Aldrich, Cat: 05892791001].

Developer solution was 500µL of Trypsin [Sigma-Aldrich, Cat: T0303-1G] and 1.5µL of 0.02mg/mL TSA.

Fresh cell pellets or those previously flash frozen were homogenized, via pipette trituration, in 300µL of HDAC lysis buffer and then incubated, at 4°C for 20min, on a rotator. The supernatant was extracted by centrifugation at 4°C and 25,000rcf for 20min. A Bradford assay estimated total protein concentration. Histone Deacetylase Substrate (BOC-Lys(Ac)-AMC) [Sigma-Aldrich, Cat: SCP0168] was thawed and prepared for use, as was an analogous lysine-crotonyl substrate (BOC-Lys(Cr)-AMC). This latter substrate was synthesized by Dr. Andrew Jamieson and Dr. Naomi Roberts, formerly of the University of Leicester's Chemistry department. The crotonyl substrate consisted of a fluorogenic molecule (AMC) bound to a crotonylated lysine residue, which was possible due to an attached BOC group (Isidro-Llobet et al., 2009). A diagram of the synthesized molecule is below:

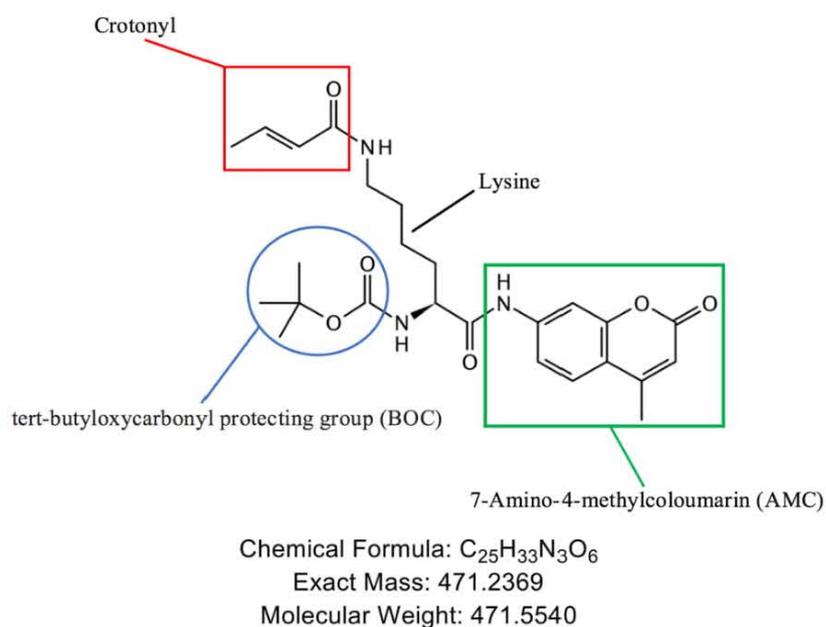


Figure 2.2: BOC-Cr-Lys-AMC Substrate.

Produced by the Chemistry Department's Dr Jamieson and Dr Roberts.

Both the BOC substrates were dissolved in DMSO [Sigma-Aldrich, Cat: D2650-100ML] and were made to a working concentration of 50 μ M using HDAC lysis buffer (10 μ L/well was needed). The required amount of lysate (typically with 0 μ g, 3 μ g, 10 μ g and 30 μ g of total protein) was added, in triplicate, to the black flat bottom 96 well polystyrene microplate [Corning, Cat: 3694]. The volume of each well was brought to 40 μ L using the HDAC buffer. 10 μ L of the required 50 μ M substrate was added to the appropriate wells. The plate was centrifuged for 8-10s (< 1000rcf) [Eppendorf, Model: 5810], and then incubated in the dark at 37 $^{\circ}$ C in an incubation shaker [INFORS HT, Model: Ecotron]. Acetyl substrate incubation was 1 hour, whereas the crotonyl substrate was incubated for 18h. 50 μ L of the developer solution was added per well, concluding the reaction. The plate was then centrifuged for 8-10s. A VICTOR X5 Light Plate Reader machine [PerkinElmer, Cat: 2030-0050] set to the HDAC assay setting quantified the result.

2.11. Rescue of *TET1/2* knockout cells with PiggyBAC TET1

TET1/2 DKO cells, created using CRISPR which employed the guide RNA described in “*One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering*”, were obtained from Professor Dame Amanda Fisher’s Lab at the MRC in the Imperial College London (Wang et al., 2013).

2.11.1. Generating the TET1 Rescue Cells Through PiggyBAC Integration of Flag-TET1 Constructs in *TET1/2* DKO Cells

TET1 was reintroduced to *TET1/2* DKO cells using the PiggyBac transposon system described in Figure 2.3. PiggyBAC vectors of wildtype *TET1* (A), *TET1*^{L897A/L900A} (C) and *TET1*^{ΔCXXC} (deletion of the 584-624 CXXC domain) (E) were generated by Prof. Shaun Cowley and Dr. Neil Bates.

2.5µg of TET1 PiggyBAC/G418 resistance plasmid, containing flanking inverted terminal repeat (ITR) sequences, were transfected into 5×10^5 *TET1/2* DKO cells/well in a 6 well plate using OPTIMEM and Lipofectamine 2000 [ThermoFisher, Cat: 11668019], as per the manufacturer’s instructions. The next day the cells were detached by trypsin and 10,000 cells were plated on a fresh 10cm plate. G418 [ThermoFisher, Cat: 10131027] resistant colonies were then picked, following ten days of selection. The cells were picked into a 96 well round bottom plate containing 50µL TrypLE [ThermoFisher, Cat: 12605010] per well, and were then transferred into a gelatinized flat bottom 96 well plate containing 100µL of M15+LIF. The cells were then passaged onto progressively larger plates until the desired number were available and a portion of them were cryopreserved.

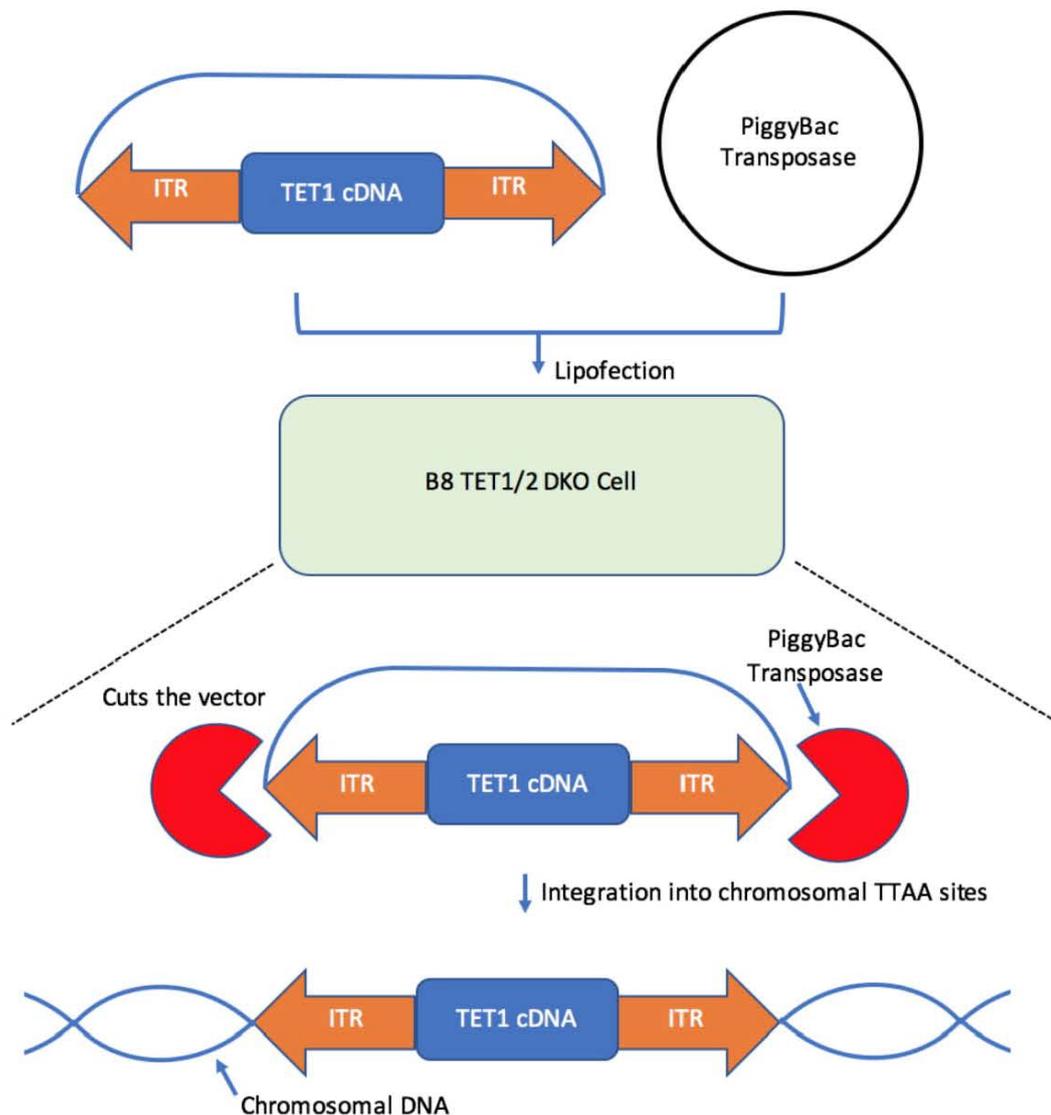


Figure 2.3: A PiggyBac System Was Used to Rescue the B8 Cells with TET1.

TET1 was placed in an artificial transposon that was stably but randomly integrated into the cell's genome via transposase.

2.12. Protein Purification for Gel-Filtration and NMR

Proteins were generated as per 2.7 though proteins for NMR, including N^{15} labelled Sin3a PAH1, were grown in BL21 Star (DE3) *E. Coli* [ThermoFisher, Cat: C601003]. N^{15} labelled proteins were produced using M9 minimal media containing (^{15}N) NH_4Cl . Protease Inhibitor Cocktail Set III [Merck, Cat: 539134] was employed during protein extraction. Lysate was run through an HPLC's nickel column, for isolation, and then

through a Q column, for purification and concentration, if possible (it was not for His-Sin3a PAH1).

An example of His-PAH1 purification follows (Figure 2.4, Figure 2.5, Figure 2.6, and Figure 3.20):

Lysate was centrifuged and the, yellow and clear, supernatant was run through a column of sepharose beads conjugated to Ni^{2+} . The nickel bound the His-tag histidines' imidazole side chains, separating the tagged proteins from the lysate. A gradient of imidazole in solution was used to competitively inhibit this interaction, permitting the tagged protein's elution. The gradient (Figure 2.4) was produced by mixing an increasing ratio of His Column Buffer B (20mM Tris at pH 8, 0.5M NaCl, 0.5M Imidazole, adjusted to pH 8 with ~6mL Concentrated HCl) to His Column Buffer A (20mM Tris at pH 8, 0.5M NaCl, 20mM Imidazole, adjusted to pH 8 with ~0.2mL Concentrated HCl).

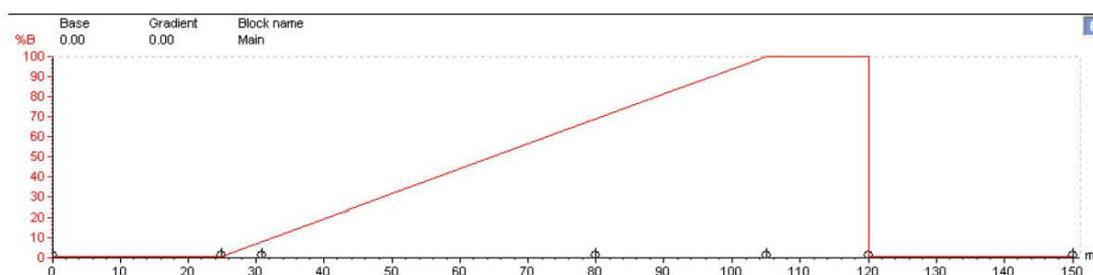


Figure 2.4: Imidazole Gradient of the Histidine Column Used to Extract Protein Fragments for NMR and Gel Filtration.

The increasing imidazole concentration elutes the histidine tagged protein of interest.

Protein concentrations were measured by UV absorbance. Elution could be detected using this (Figure 2.5), or by a sudden drop in conductivity (Atchley and Nichols, 1925). Both these methods identify fractions A4-A7 as containing the protein of interest.

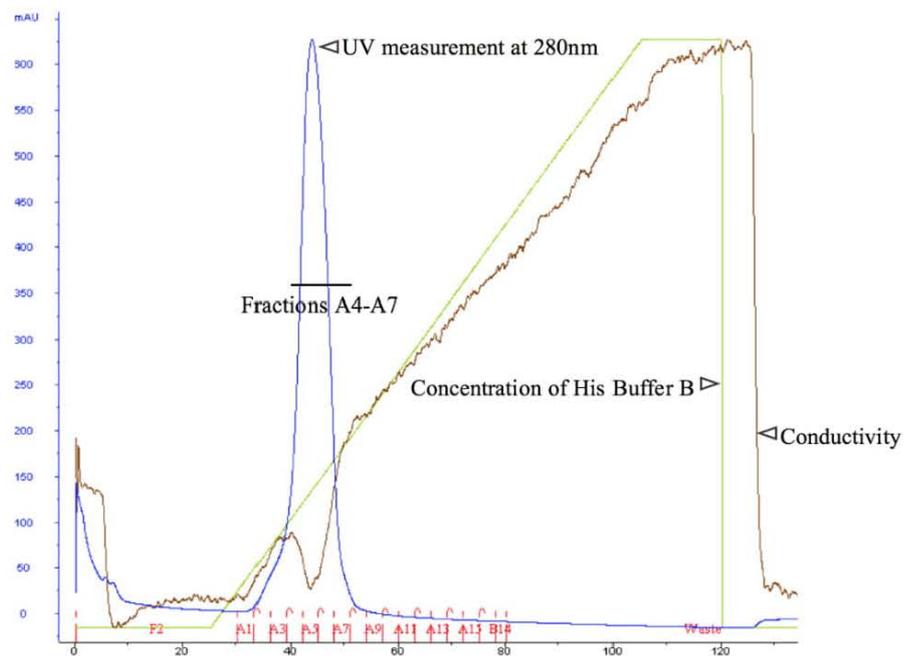


Figure 2.5: His-Sin3a 115-212 Elution in Fractions A4-A7.

His Buffer B contains a high concentration of imidazole, its measurement is therefore analogous to one of imidazole. The UV and conductivity measurements are also indicated, these measure the elution point of the proteins.

All relevant fractions were collected and run on SDS-PAGE (Figure 3.18). Fractions containing significant portions of the appropriate protein were combined (like fractions A5 and A6 of PAH1). All tags contained a TEV cleavage site, permitting their removal. The tags and the TEV itself could be isolated from the protein, via nickel column, as they were both histidine tagged.

A Q column was used to purify and concentrate proteins isolated by the nickel column. This occurred by charge, which required an increasing salt concentration gradient. This gradient (Figure 2.6) was provided by the ratio of Q Column Buffer B (20mM Tris at pH 8 and 1M NaCl) to Q Column Buffer A (20mM Tris at pH 8 and 60mM NaCl).

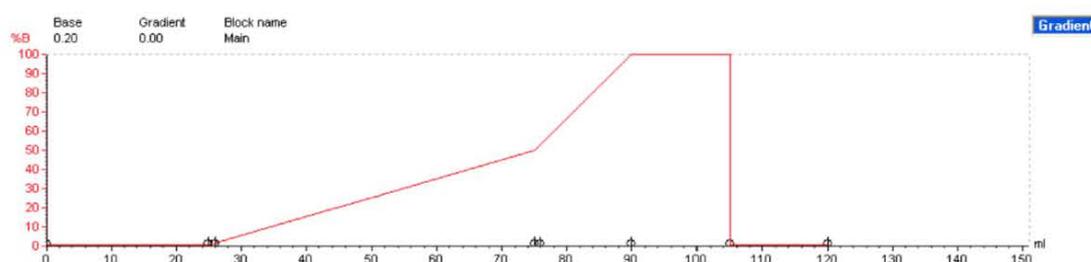


Figure 2.6: NaCl Gradient, in a Q Column, Purifies Proteins on the Basis of Charge.

This step also permitted the concentration of the proteins, that permitted the use of NMR.

16% PAGE gels were used for the gel-filtration analysis, due to the similar sizes of His-GB1-TET1 878-911 and His-PAH1. Gel filtration buffer was 50mM Tris pH8.0, 150mM NaCl and 1mM DTT. NMR proteins required dialyzing into 20mM Tris pH 7.4, 50mM NaCl and 1mM DTT. Due to cost, Dr. Neil Bates input the purified proteins into both the NMR and gel filtration apparatuses; the gel filtration peaks and NMR spectra are the products of this effort. The NMR model of the PAH1:TET1-SID interaction, presented in Figure 3.47, was produced by Prof. Geerten Vuister.

2.13. Polymerase Chain Reaction (PCR)

The thermal cycler was used to amplify defined segments of DNA from a template. A typical setting on the cycler, where X is the annealing temperature, was: 95°C 2min, (95°C 30s, X°C 30s, 72°C 30s) × 35, 72°C 5min, 12°C ∞s (~10 minutes in practice).

Reactions typically contained: 10×Standard Taq Reaction Buffer (contains MgCl₂), 200μM of all four dNTPs, 0.2μM forward primer, 0.2μM reverse primer, 0.5μg DNA template, 1U/40μL Taq DNA Polymerase and sterile nuclease-free water to the desired volume (25-50μL).

2.14. Quantitative Reverse Transcription PCR (qRT-PCR)

2.14.1. RNA Extraction

Cells derived from a confluent 6cm plate were pelleted, snap frozen in liquid nitrogen and then stored at -80 °C. RNA was extracted from frozen pelleted cells, that were first thawed, using an RNeasy® Mini Kit [Qiagen, Cat: 74106], as per the manufacturer's instructions.

2.14.2. Reverse Transcription

4µL of qScript [VWR, Cat: 733-117] was added to 0.5µg of total RNA and 20µL of nuclease free autoclaved water. The mixture was lightly vortexed and centrifuged. Following this, the mixture was added to the thermal cycler, which was run at: 5min at 25°C, 30 min at 42°C, 5 min at 85°C, ∞ at 4°C (~10min).

Nuclease free autoclaved water was then added (typically 75µL) to the synthesized cDNA to attain a volume of 100µL. 2µL of this was used for each qRT-PCR reaction; as this was done in triplicate, every qRT-PCR run consumed ~6µL of diluted cDNA.

2.14.3. Gel Extraction and PCR Cleanup and Dilution Series

qRT-PCR started with gel extraction [Qiagen, Cat: 28704] followed by 'PCR cleanup' [Qiagen, Cat: 28006] to produce a dilution series to establish $\geq 90\%$ primer efficiency. Both kits were used as per the manufacturer's instructions. The series of 10 dilutions required 10µL of Minelute® eluate, diluted in 90µL of nuclease-free autoclaved water (Borman, 1987).

2.14.4. Quantitative Reverse Transcription PCR Reaction

Each well on a 96-well plate had 2µL RT-DNA added to it. To this 8µL of Sensifast solution was added, which consisted of: 5µL Sensifast [Bioline, Cat: BIO-86005],

0.5 μ L (2mM) forward + reverse primer (Table 7.5), 2.5 μ L nuclease free autoclaved water. The assay was done in triplicate.

All qRT-PCR measurements were normalized to GAPDH. GAPDH will be used to illustrate the primer efficiency process (Figure 2.7):

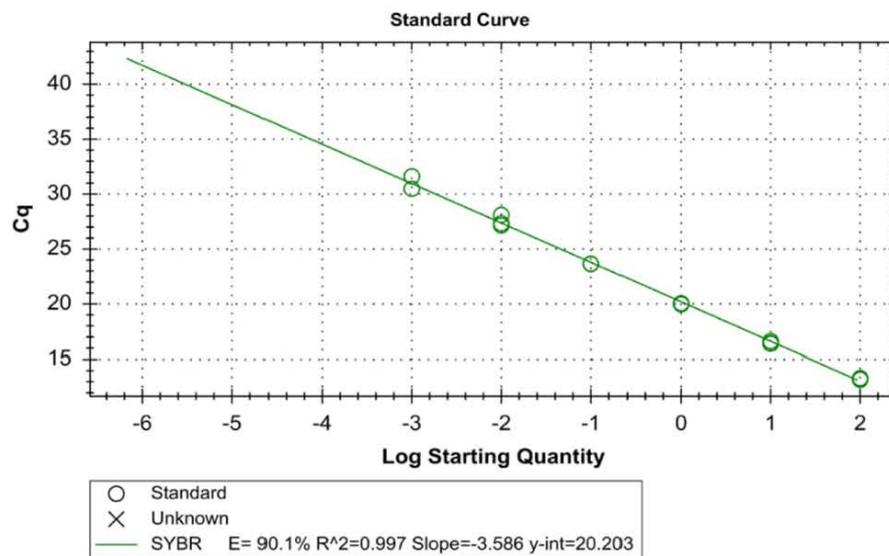


Figure 2.7: Primer Efficiency Curve of GAPDH.

Obtaining a primer efficiency of $\geq 90\%$ was required for accurate qRT-PCR.

The melt curve precluded the presence of multiple amplification sites (Figure 2.8).

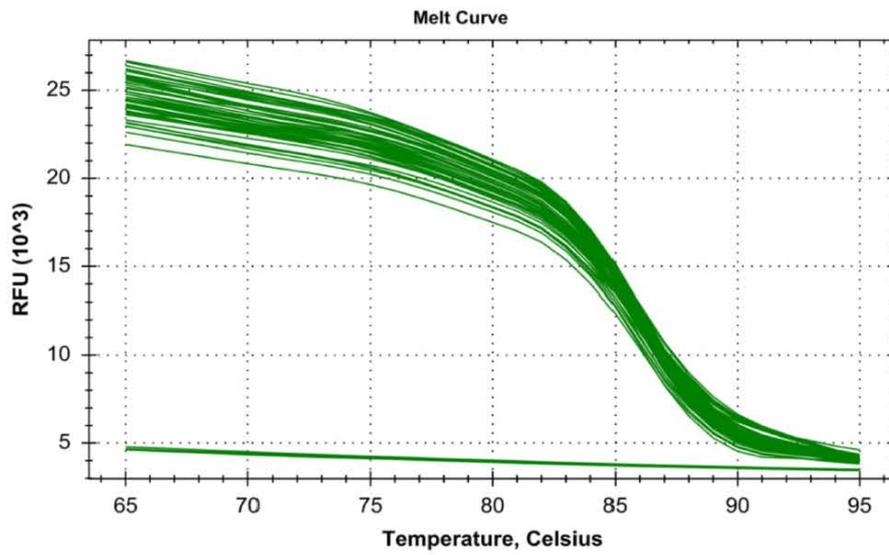


Figure 2.8: qRT-PCR Melt Curve of GAPDH.
The uninterrupted decrease in signal indicates the presence of a single amplicon.

Similarly, the melt peak (Figure 2.9) established that only one amplification product was produced, and that each of the samples only produced this one amplicon.

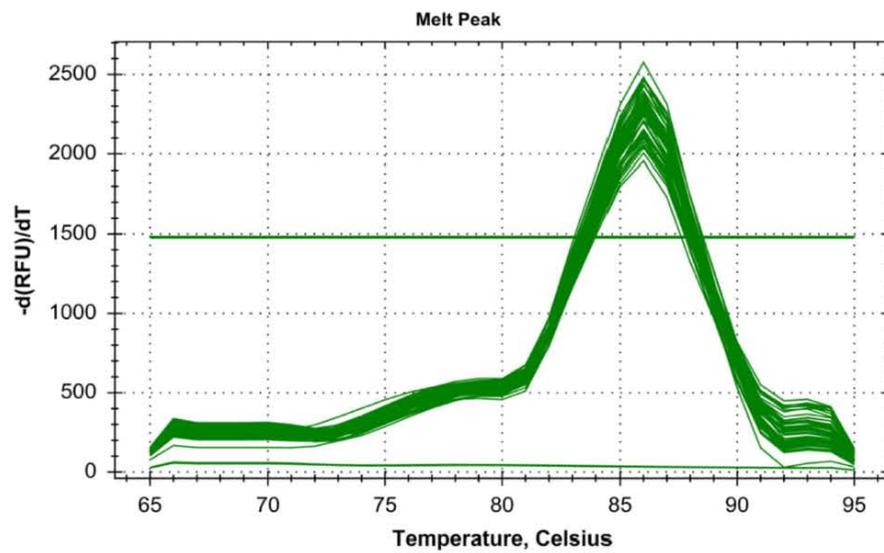


Figure 2.9: qRT-PCR Melt Peak of GAPDH.
The single peak demonstrates that only one product is produced by each PCR reaction, and that they are all producing the same amplicon.

2.14.5. qRT-PCR Calculation of Fold Change

The qRT-PCR data was measured in cycle threshold units (Ct), which were the number of PCR cycles before the signal exceeded a predefined threshold. The change in Ct values were obtained for both the sample (ΔCt Experimental) and the control GAPDH (ΔCt Control). The subtraction of ΔCt Control from ΔCt Experimental provided the $\Delta\Delta\text{Ct}$. $2^{-\Delta\Delta\text{Ct}}$ provided the fold change in expression between the control group (E14 ESCs) and the B8 *TET1/2* DKO or PiggyBAC rescue cells.

Chapter 3. Results: Sin3a-TET1

3.1. A Conserved Sin3a Interaction Domain

The first objective was to describe the TET1 Sin3a interaction domain. Identifying this domain was aided by the findings of two separate research teams. The Helin group, in 2011, used immunoprecipitation in tandem with mass spectrometry to determine that TET1 binds to Sin3a (Williams et al., 2011). This same technique also suggested that TET2 does not directly interact with Sin3a. The following year, Deplus et al. (2013) confirmed these findings and additionally demonstrated that TET3 also binds to Sin3a. The TET family of enzymes are thought to have been formed, over evolutionary time, through two consecutive iterations of gene duplication (Akahori et al., 2015, Iyer et al., 2009). This suggests that it is more likely that the structure, in TET1 and TET3, responsible for binding to Sin3a has been evolutionarily conserved. The alternative to this theory is that a structure performing an identical function evolved convergently within the same gene family, a prospect that is somewhat implausible. The primary sequences of the TET1 and TET3 proteins were therefore compared using the online sequence alignment tool, Multalin (Figure 3.1). A short 12 amino acid region was found in the two proteins that was conspicuously absent in TET2 (Corpet, 1988).

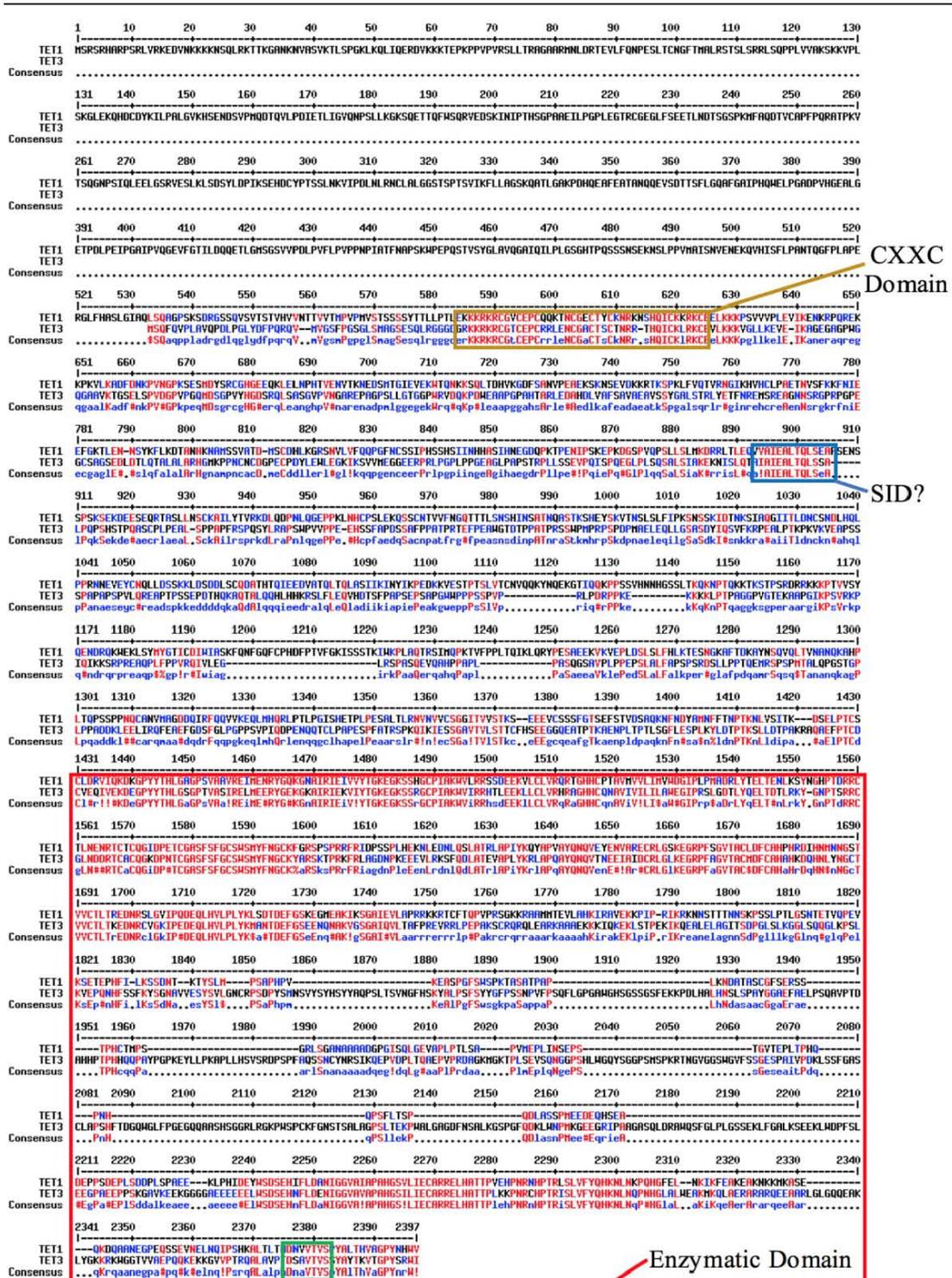


Figure 3.1: Sequence Alignment of TET1 and TET3.

The online sequence alignment tool, Multalin, was used to align TET1 and TET3. A small, potentially helical, sequence was found in TET1 (892-903) and 3 (259-271) but not TET2. This appeared to be a putative Sin3a interaction domain (SID). DNA binding region (CXXC), OGT binding region and the enzymatic domain are also indicated.

Indeed, the region of interest appears to be conserved quite tightly across evolutionary time (Table 7.1, in the appendix). Predictably, the sequence is highly conserved between primates; but surprisingly, this is largely the case across mammals, birds, reptiles and even fish - Figure 3.2.

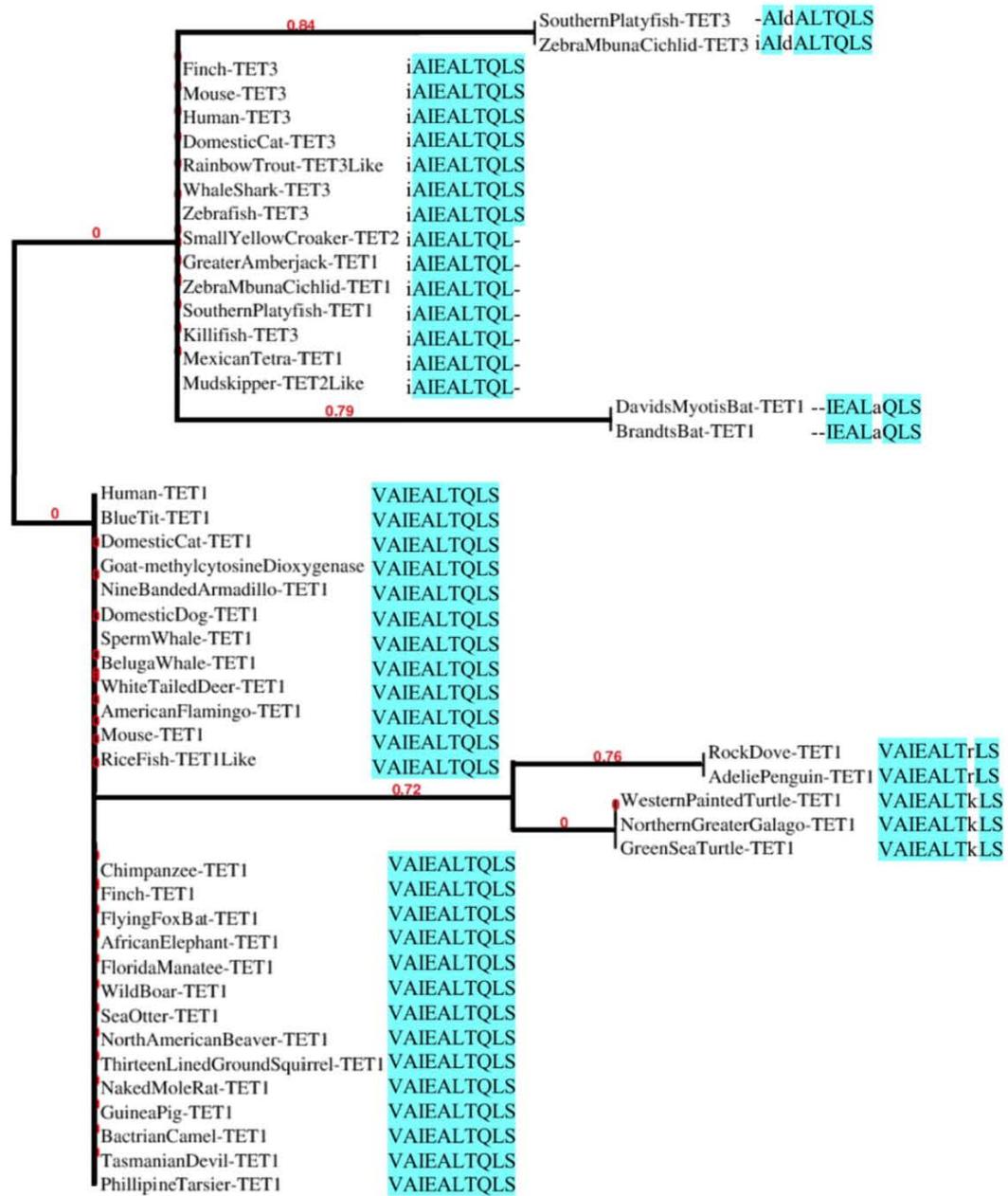


Figure 3.2: Phylogeny of the Putative TET1-SID (892-901) Sequence.

A protein BLAST followed by a phylogeny.fr plot of TET1 892-903 (Dereeper et al., 2008).

3.1.1. An α -Helical Conserved Domain

This nascent hypothesis, was further supported by another online analytical tool called Joint Network prediction of secondary structure (Jpred) which revealed the region of

interest was likely an alpha helix – a prominent Sin3a interaction domain motif (Figure 3.3). This region predicted by Jpred (TET1 888-900), to be an alpha helix, coincides with the Multalin identified region of TET1 892-903. It bears mentioning that Jpred is a collection of predictive tools, and the 888-900 prediction was the result of Jnet PSIBLASTpssm (Position-Specific Iterative Basic Local Alignment Search Tool position-specific scoring matrix profile). Jnethmm (hidden Markov model) profile predicted the alpha helix consisted of residues 888-898. These multiple analytical tools are combined to provide an estimate as to the ‘reliability of prediction accuracy’ for each amino acid (Drozdetskiy et al., 2015). TET1 888-900 has an average accuracy of 7.54 out of 9. It should be noted that the prediction accuracy for the last three amino acids is negligible. A helix (H1) appears in the Jpssm prediction between residues 876 and 883, which precedes the putative SID (H2). Multalin demonstrates that these residues are not particularly well conserved between the TET homologues. The Jpred average prediction accuracy for this region is correspondingly low – 2.22. This would translate to 24.69%, while 888-900’s average reliability score would be 83.78%. Both Jpred and Multalin concur, in suggesting that this 876-883 helix is unlikely to be significant in the context of this thesis. While both tools agree that the putative SID is significant; there is divergence on the definition of the region’s boundaries. T898 of the **VAIEALTQLS** (892-903) sequence, revealed by Multalin, is the last residue Jpred technique predicts with high accuracy. Jpred also hints that there may additionally be a few amino acids preceding this sequence that may take part in the formation of a helix; though neither Q890 nor V891 are conserved in TET3.

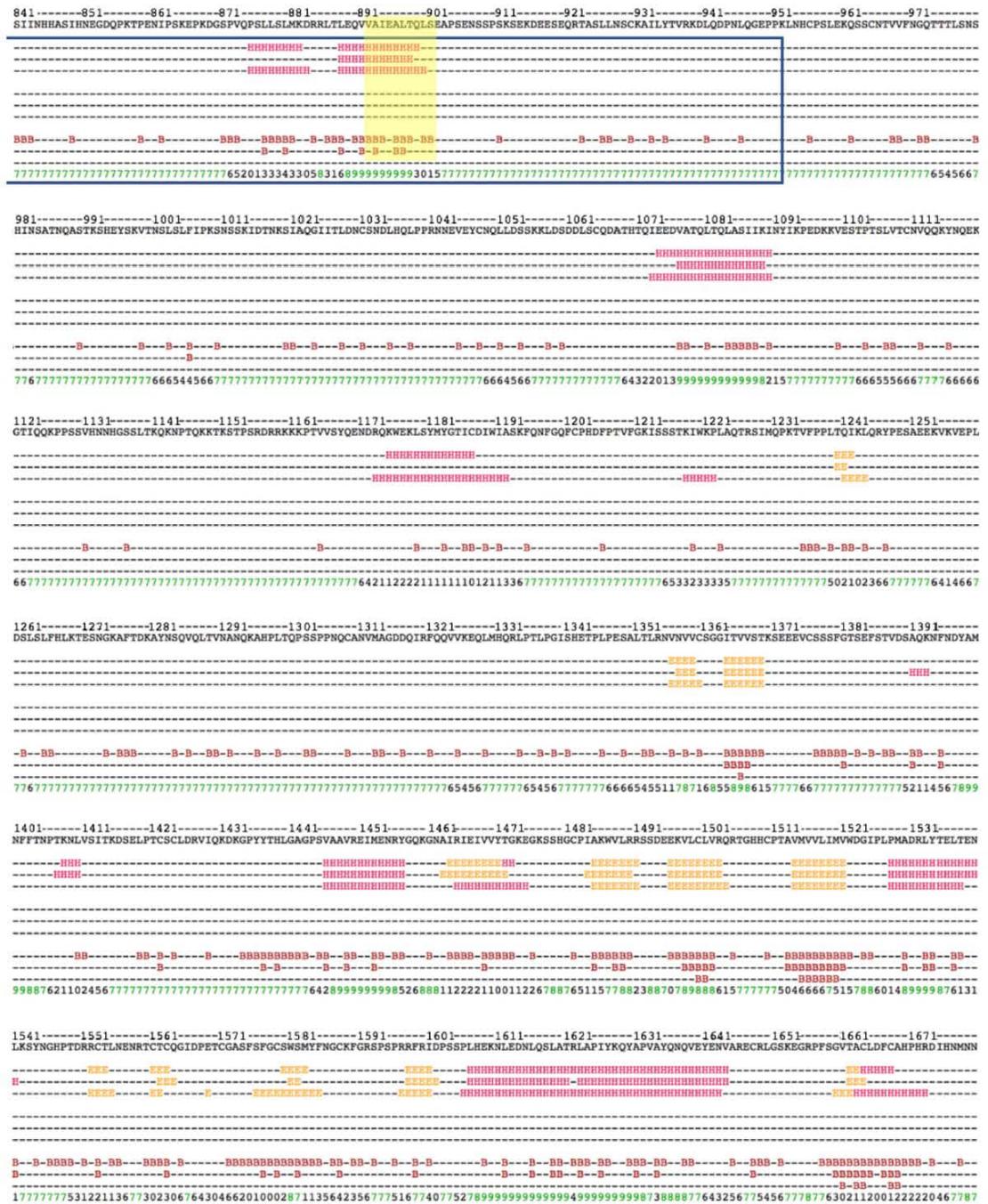




Figure 3.3: Jpred Secondary Structure Prediction of TET1.
 The TET1 892-903 region is highlighted and the 746-951 and 1800-2136 regions placed in boxes. H indicates a putative helical structure, number at the bottom is the reliability of prediction accuracy.

3.1.2. The α -Helix Appears to be a SID

The TET1 putative SID appears to be an amphipathic helix. This is consistent with known SID domains including SAP25 133-148, NRSF 42-57 and Pf1 207-225, illustrated in Figure 1.17 (Sahu et al., 2008, Nomura et al., 2005, Kumar et al., 2011). A helical wheel prediction of TET1 891-901 shows a fairly convincing helical structure (Figure 3.4). There were, however, a few things still to be determined. It is unknown if

A896 sits on the hydrophobic or hydrophilic side of the amphipathic helix, as it is only weakly hydrophobic and therefore may have a limited impact on binding (Nilsson et al., 2003). There was also the curious presence of, the hydrophilic amino acid residue, S901 on the hydrophobic face of the amphipathic helix. It could be an indication that S901 lies on the end of the helix and therefore denotes the boundary of the TET1-SID, this would also fit with the region of conservation with TET3 (see Figure 3.1 above). It would also be in accordance with the Jpred prediction which implies the SID might end somewhat prior to the end of this region of conservation.

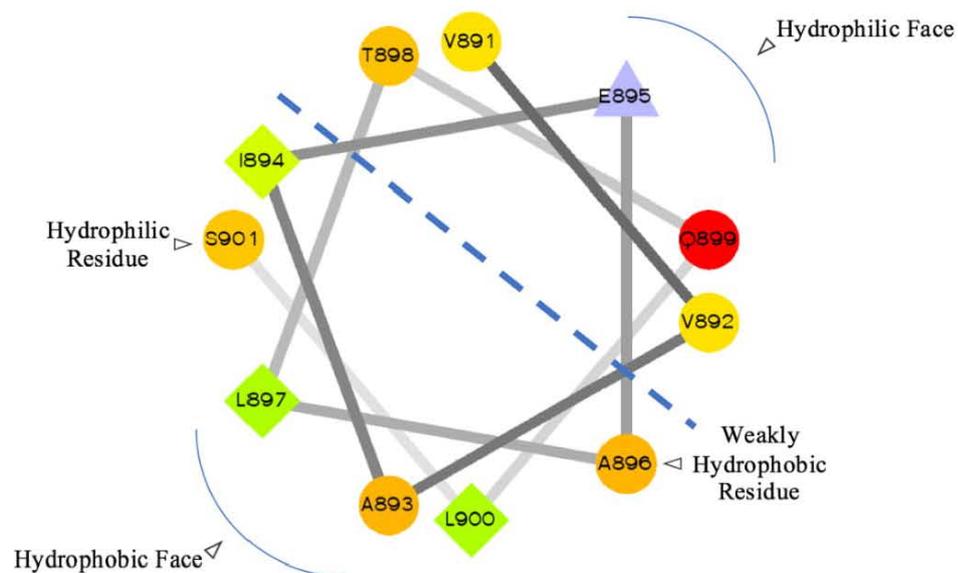


Figure 3.4: Putative SID Appears to be an Amphipathic Helix.

Helical wheel prediction of TET1 891-901 (courtesy of the Prof. Raphael Zidovetzki lab). Central line representing the division between the hydrophobic and hydrophilic faces of the amphipathic helix.

3.1.3. The SID is Correlated with Binding

The TET1 SID is correlated with Sin3a association – only fragments of TET1 that contain the putative SID were found capable of binding to Sin3a via co-IP. Three fragments of TET1 were produced to test for the presence of an active SID, including: TET1 746-951 (~32kDa), 746-1255 (~49kDa) and 1800-2136 (38kDa). The first two contained the putative SID while the TET1 1800-2136 fragment did not. The TET1 746-951 and TET1 1800-2136 regions can be seen in Figure 3.3. These three regions

were used along with Myc-Sin3a (~150kDa) in a co-immunoprecipitation (co-IP) experiment using 293T cells (Figure 3.5). The experiment revealed, as predicted, that both SID containing fragments (TET1 746-951 and 746-1255) bound to Sin3a while the C-terminal fragment (TET1 1800-2136) was unable to do the same. This is seen when lanes 9 and 12 are compared with lane 15 in the lower panes of Figure 3.5.

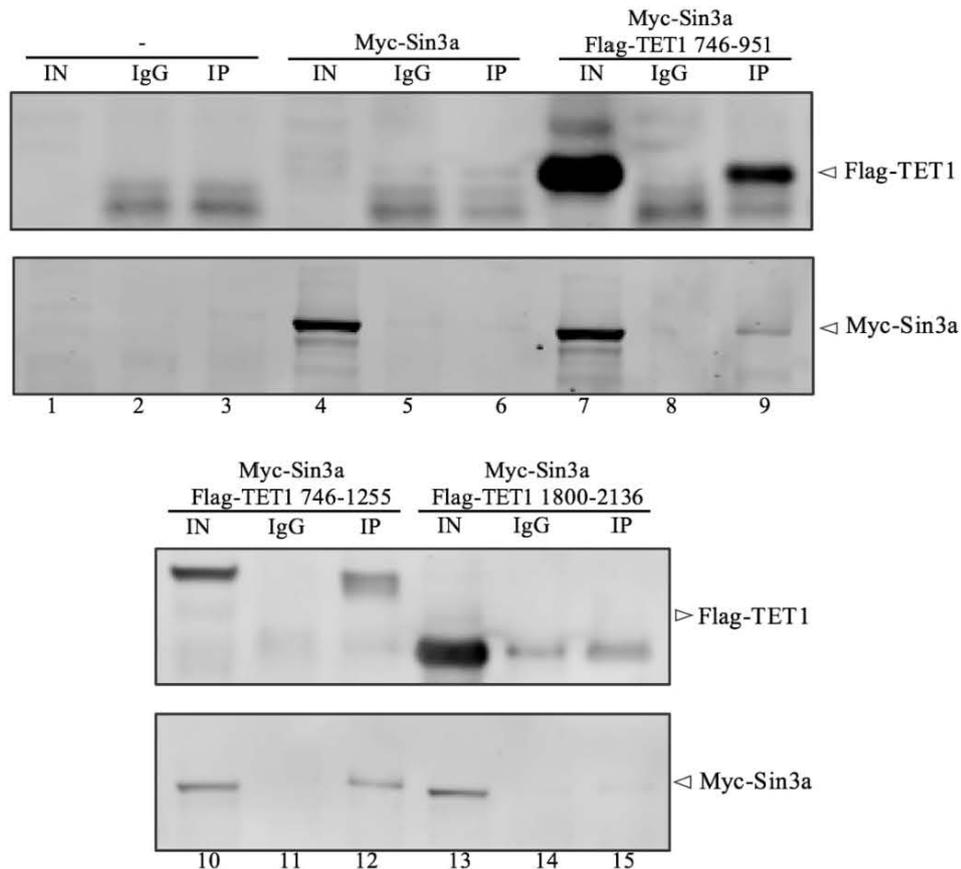


Figure 3.5: SID Containing Fragments Interact with Sin3a.

A co-IP analysis – Flag-TET1 (746-951, 746-1255 and 1800-2136) fragments were used to pull down Myc-Sin3a in 293T cells. In = input, IgG = negative control, IP = immunoprecipitation.

3.1.4. The TET1-SID is Necessary for Sin3a Binding

Co-IP, using a deletion construct, revealed the presence of the TET1 SID to be necessary for the interaction with both endogenous and transfected Sin3a. While the preceding co-IP (Figure 3.5) supports the hypothesis that the putative SID is the TET1

region responsible for binding Sin3a, the TET1 746-951 fragment is too large to provide a definitive answer. Three additional TET1 fragments were produced to fill this gap. The first, 746-887 (~21kDa), terminates just before the putative SID; while the second, 746-909 (~24kDa), includes this region. Additionally, a TET1 746-951 fragment missing (Δ) amino acids 891-902 was also produced. As expected, the 746-887 and 746-951 Δ 891-902 constructs were unable to bind to HALO-hSin3a (~173kDa), demonstrated by the lack of bands in lanes 9 and 12 in the middle panel of Figure 3.6. Conversely, 746-909 did interact with HALO-hSin3a as seen in lane 15 of the same panel. The Multalin and Jpred predictions regarding the 876-883 region are validated by the inability of TET1 746-887 to bind to HALO-hSin3a in this co-IP. The TET1 746-951 Δ 891-902 and 746-909 results reveal that the SID predicted by Multalin is necessary for Sin3a binding. Perhaps even more reassuring is the ability of the 746-951 and 746-909 fragments, to not only bind to HALO-hSin3a, but also to endogenous Sin3a (~145kDa) as well (lanes 6 and 15, bottom panel). The minimal binding region that has so far been established is TET1 746-909.

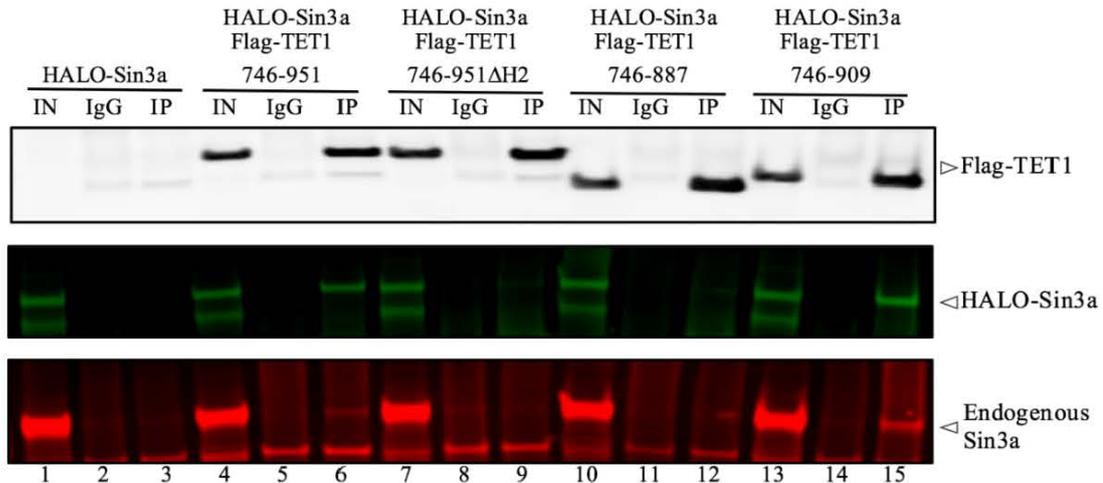


Figure 3.6: Putative SID Required for Sin3a Binding.

Co-Immunoprecipitation analysis of Flag-TET1 fragments with HALO-hSin3a (green) and endogenous Sin3a (red) in 293T cells. TET1 746-951 Δ H2 corresponds to TET1 746-951 deleted (Δ) for amino acids 891-902.

3.1.5. The SID is Sufficient for Sin3a Binding

A GST-pulldown demonstrated the SID itself, in the form of TET1 878-911 was sufficient to interact with Sin3a. Establishing this minimal region required the use of another technique – GST pulldowns in conjunction with S³⁵-Met labelled Sin3a made by *in vitro* transcription and translation (IVTT). The benefits of this system over co-IP are that smaller fragments can be used and reliably detected. In addition, there is no possibility of the heavy and light antibody chains obfuscating the result during imaging. While the 891-902 region is necessary for binding to Sin3a, it could potentially be a piece of a larger SID. An example would be if the first helix was needed in addition to the second for the formation of the complex. Three different GST-TET1 fragments were generated, all centered around the putative SID (891-902). They were designed to have successively larger flanking regions ending close to a proline residue. In addition to GST-TET1 878-911, 868-927 and 848-952; a GST-MAD1 1-35 (MAD-SID) fusion was produced as a positive control as its interaction with Sin3a was well-established (Brubaker et al., 2000). All four constructs bound to Sin3a as seen in Figure 3.7. Helix 2, the putative SID (Figure 3.3) was the only unabridged helix present in the TET1 878-911 fragment, thus confirming it was responsible for Sin3a binding. Moreover, the TET1 878-911 fragment was sufficient to bind to Sin3a. It is perhaps worth noting that the 868-927 fragment appeared to interact with Sin3a slightly better than its sister fragments (878-911 and 848-952).

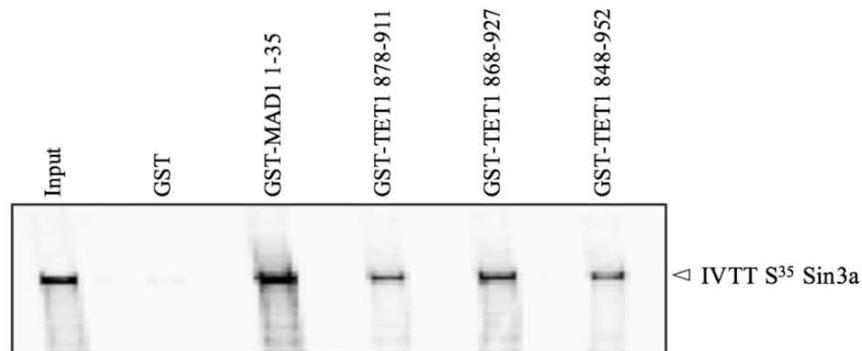


Figure 3.7: Putative SID Sufficient for Sin3a Binding.

GST pull-downs to determine a minimal binding region. GST-TET1 SID and GST-SAP25-SID fragments were used to pull down S³⁵-Met IVTT Sin3a.

3.2. Sin3a's TET1 Interaction Domain

The experiments to discover the TET1 binding region of Sin3a was somewhat more straightforward as compared to the converse. This was partially because there were Sin3a constructs already in the lab from prior research. Equally important was the significant body of published research on the structure of Sin3a. While theoretically any site along the length of the protein could be involved in facilitating the interaction with TET1, there were only five very strong possibilities. These were the four PAH domains as well as the HDAC interaction domain (HID).

Four C-terminal truncations of Sin3a were used to investigate these regions. These were the Myc-Sin3a 1-1015, 1-680, 1-479 and 1-205 constructs that were previously validated in the literature (Laherty et al., 1997). Each of these constructs was created with different Sin3a domains in mind and as such, each successively shorter fragment lacked an additional domain from the C-terminal of the protein (Figure 3.8). For example, the 1-1015 fragment lacked PAH4, while the 1-680 construct lacked PAH4 and additionally the HID. A full-length Sin3a construct lacking the HID and a deletion missing amino acids 524-851, were also used.

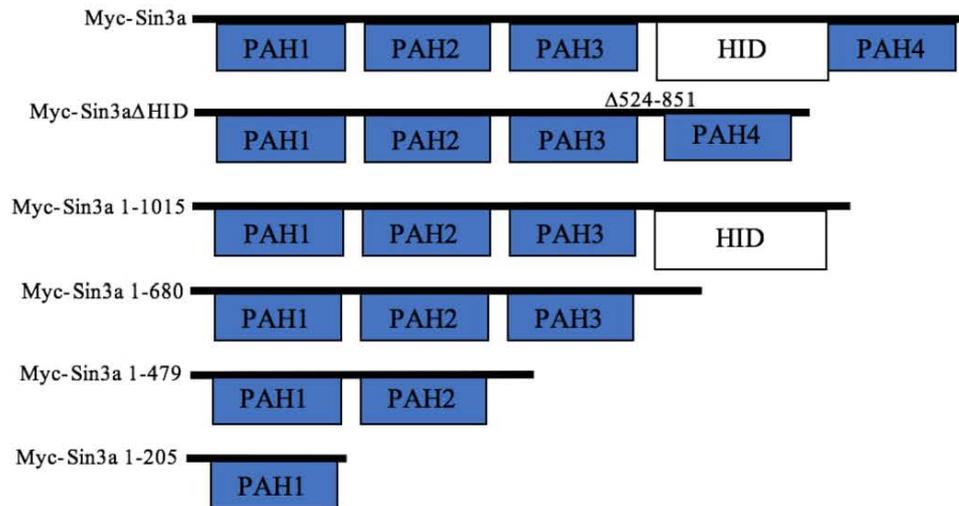


Figure 3.8: Schematic Diagram of the Six Different Myc-tagged Sin3a Constructs Used in Pulldown Experiments (Laherty et al., 1997).

The amino acid boundaries are indicated. Paired amphipathic helix (PAH) domain and HDAC interaction domain (HID) are also shown.

3.2.1. Sin3a's TET1 Binding Region Occurs Within PAH1-3

At least one of the domains in the PAH1-3 region appear to be important for the Sin3a:TET1 interaction. As respectively shown in lanes 9, 12, 15 and 18 in the lower panels of Figure 3.9; full-length Sin3a, Δ HID, 1-1015 and 1-680 were all able to bind to TET1 746-951. PAH1-3 were the only intact domains within the smallest bound fragment – Myc-Sin3a 1-680. This strongly suggested one of these three domains was the region responsible for binding to TET1. The light antibody chain (~25kDa) runs at approximately the same position as Flag-TET1 746-951, which is visible throughout the top panes of Figure 3.9; though this can also be seen in the top panels of Figure 3.5 and Figure 3.10.

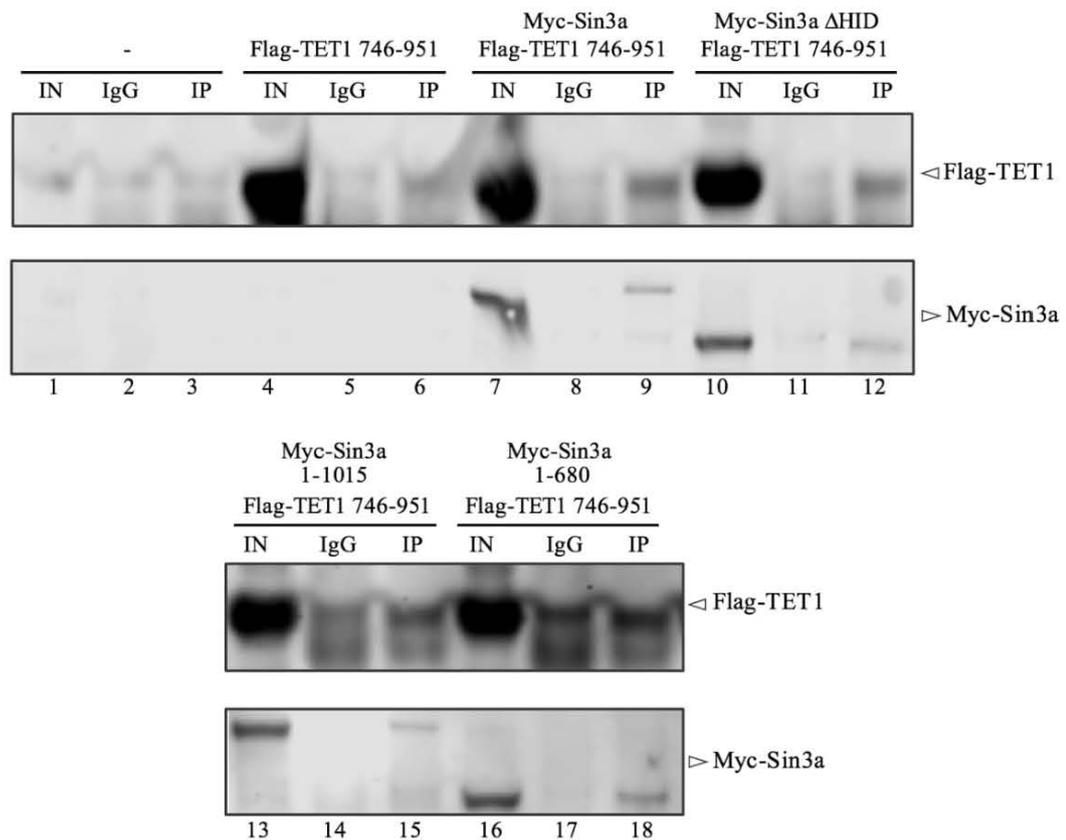


Figure 3.9: Sin3a's TET1 Interaction Domain is Between PAH1-3.

Co-Immunoprecipitation analysis Flag-TET1 746-951 (top panels) was used to pull down C-terminal deletion Myc-Sin3a fragments (bottom panels) in 293T cells. Sin3a Δ HID corresponds to Sin3a deleted (Δ) for amino acids 524-851.

The Sin3a 1-479 and 1-205 fragments did not produce any interpretable result. This is due to the bands (lanes 20 and 23 in Figure 3.10) present in the negative control IgG lanes. Repeating the experiment did not provide further clarity. This in combination with the interference from the heavy and light antibody chains present in all the co-IPs, contributed to the shift towards the use of GST pulldowns coupled with radiolabeled Sin3a synthesized using IVTT. An added benefit of this latter method is the increased detection sensitivity that enabled the use of smaller proteins, as demonstrated in Figure 3.7.

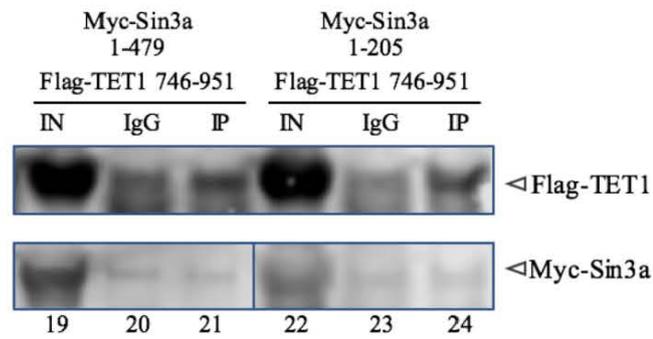


Figure 3.10: Indeterminate Result Regarding Myc Sin3a 1-479 and 1-205.

Co-Immunoprecipitation analysis Flag-TET1 746-951 (top panels) was used to attempt a pull down of C-terminal deletion Myc-Sin3a fragments 1-479 and 1-205 (bottom panels) in 293T cells.

3.2.2. TET1 Weakly Binds Sin3a 1-205 (PAH1), but Strongly Binds Sin3a 1-479 (PAH1-2)

TET1 878-911 was found to weakly bind to a fragment that corresponded to PAH1. This was established by a GST-pulldown that used four Sin3a C-terminal deletion fragments (1-1015, 1-680, 1-479 and 1-205) and GST-TET1 878-911 (TET1-SID) as the bait. In addition to the TET1-SID, the SAP25-SID (126-186) was used as a positive control. Its interaction with PAH1 has been extensively described through numerous methods including NMR (Shiio et al., 2006, Sahu et al., 2008). All four C-terminal deletions, each of which contain an intact PAH1 domain, expectedly bind GST-SAP25 126-186 (SAP25-SID) as seen in Figure 3.11. Superficially, it appeared that GST-TET1-SID showed binding to all fragments except for Sin3a 1-205; but a longer exposure suggested the TET1 construct was also a PAH1 binder, albeit at lower levels than SAP25 (right panel, Figure 3.11).

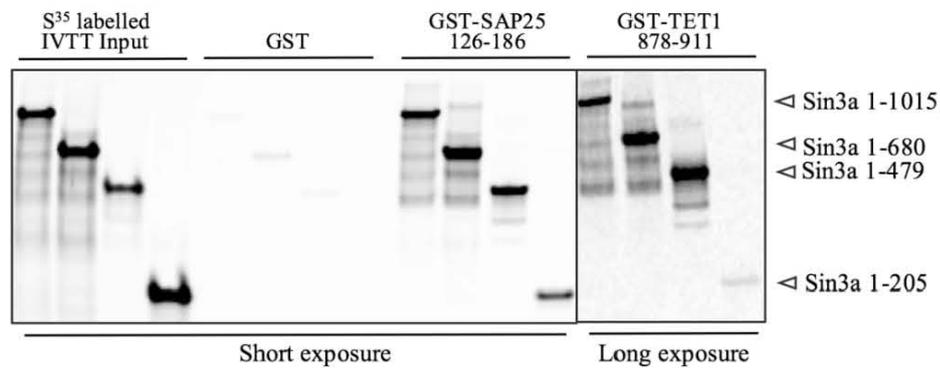


Figure 3.11: GST-TET1 SID Appears to Associate with PAH1.

GST pull-down – GST-SAP25-SID and TET1-SID were used to pull down S^{35} -Met IVTT Sin3a C-terminal deletion constructs. The GST-TET1-SID pull-down set required a greater exposure to be successfully imaged.

3.3. PAH1 is Required for TET1 Binding

PAH1 was found to be required for TET1 binding, this was established using a second set of constructs – Sin3a 1-388 and Sin3a 291-899, which consisted of PAH1+PAH2 and PAH2+PAH3 respectively. These were also previously validated in the literature (Cowley et al., 2004). GST-MAD1 1-35 is seen to bind both Sin3a 1-388 and Sin3a 291-899, in addition to full-length Sin3a in Figure 3.12. This is consistent with it being a PAH2 binder, as both fragments contain this domain. In the same image, GST-SAP25-SID is only able to bind to the Sin3a 1-388 in addition to the full-length protein, in agreement with its PAH1 binding ability. There is a band for GST-SAP25-SID corresponding to Sin3a 291-899, but this is at lower intensity than the GST nonspecific background. Remindful of GST-SAP25-SID, GST-TET1 868-927 only interacts with Sin3a 1-388 (in addition to full-length Sin3a), which would only be the case if it exclusively bound PAH1 (Figure 3.12). This result is in agreement with that of Figure 3.11. The Sin3a 1-205 and 1-388 fragments, used in the experiments detailed in Figure 3.11 and Figure 3.12, have a ~100 amino acid stretch prior to PAH1. It could be argued that this region might be responsible for the TET1 interaction. This issue will be addressed in Figure 3.13 and Figure 3.14.

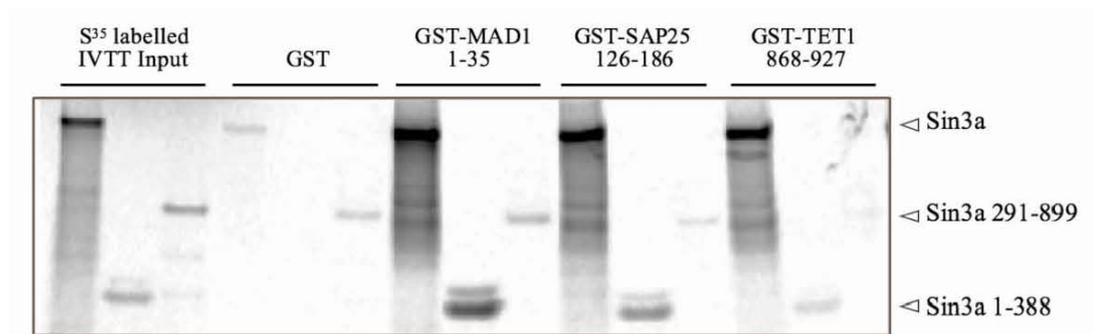


Figure 3.12: GST-TET1 868-927 Binds to PAH1.

GST pull-down – GST-SAP25-SID, SAP25-SID and TET1 868-927 fragments were used to pull down S³⁵-Met IVTT Sin3a as well as Sin3a 291-899 and 1-388 and constructs.

3.4. PAH1 is Necessary for TET1 Binding

An intact PAH1 domain is necessary for TET1 SID binding, and this can be disrupted by a double mutation. The third and final set of, publication verified, Sin3a constructs were used in the experiment illustrated in Figure 3.13 (Cowley et al., 2004). Each of these were full-length versions of Sin3a, but with an inactivating double proline mutation (PP) in each of the PAH domains. Only PAH1PP failed to interact with TET1. This was demonstrated, in the left image in Figure 3.13, by the lack of a band corresponding to PAH1PP, in the bottom GST-TET1-SID pane. In the accompanying graph on the right, it is similarly clear that PAH1PP is the only Sin3a construct unable to bind to GST-TET1 878-911. Together these results demonstrate PAH1 is required for TET1 binding.

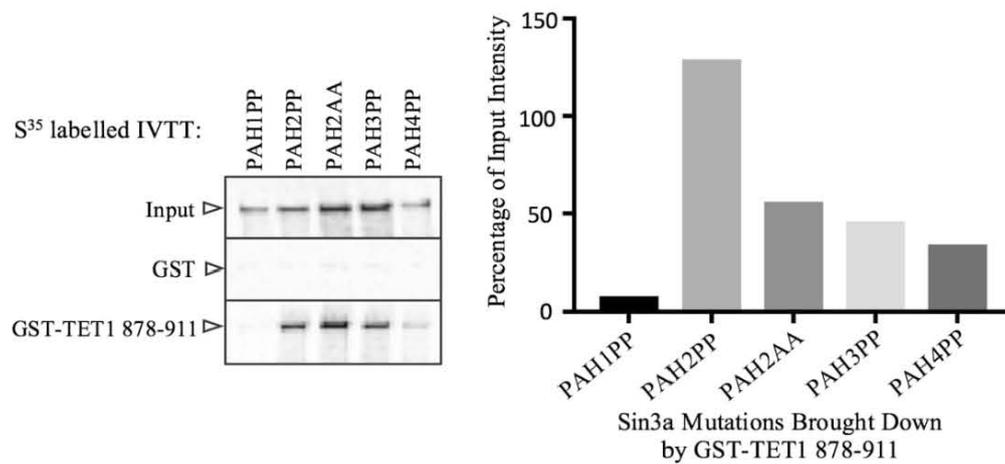


Figure 3.13: PAH1 is Necessary for TET1 Binding.

GST pull-down – GST- TET1-SID was used to pull down S^{35} -Met IVTT full-length Sin3a PAH1PP, PAH2AA, PAH2PP, PAH3PP and PAH4PP constructs. PAH1PP: L127P/ L130P, PAH2AA: L329A/L332A, PAH2PP: L329P/L332P, PAH3PP: L481P/L485P and PAH4PP: L903P/L907P.

3.5. PAH1 is Sufficient for TET1 Binding

PAH1, in isolation, is sufficient for TET1 binding. This was determined by GST-pulldown. GST fused PAH domains were assayed for interaction with S^{35} -Met labelled TET1 746-951. GST-PAH4, however, did not express. Amending the boundaries of the fragment failed to resolve the problem. GST-PAH1 was the only domain able to interact with the TET1-SID 746-951, as seen in Figure 3.14. This establishes that PAH1 is sufficient for TET1 binding. This simultaneously demonstrates that the ~100 amino acids preceding PAH1 in the Sin3a 1-205 and 1-388 fragments, in Figure 3.13 and Figure 3.14, are not involved in complex formation.

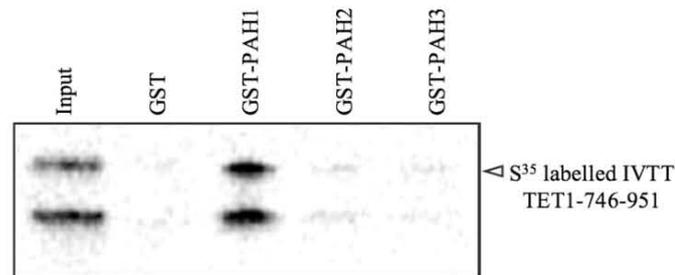


Figure 3.14: PAH1 is Sufficient for TET1 Binding.

GST pull-down –GST- Sin3a PAH1, PAH2 and PAH3 were used to pull down S³⁵-Met IVTT TET1 746-951. GST-PAH1: 115-212, GST-PAH2: 292-388 and GST-PAH3:488-530. GST-PAH4:856-1273, would not sufficiently express.

3.6. A Single PAH1 Interaction Site on TET1

Serial protein fragmentation was used in a GST-pulldown to establish that there is only a single PAH1 interaction site on TET1. With the aid of the Jpred prediction (Figure 3.3), to avoid disrupting any domains, TET1 was sequentially divided into fragments that covered the entire length of the protein. Multiple unsuccessful attempts were made at producing the S³⁵-Met TET1 384-745 fragment. The only TET1 region that was capable of binding to GST-Sin3a 115-212 (PAH1) was S³⁵-Met TET1 746-951, this can be seen in Figure 3.15.

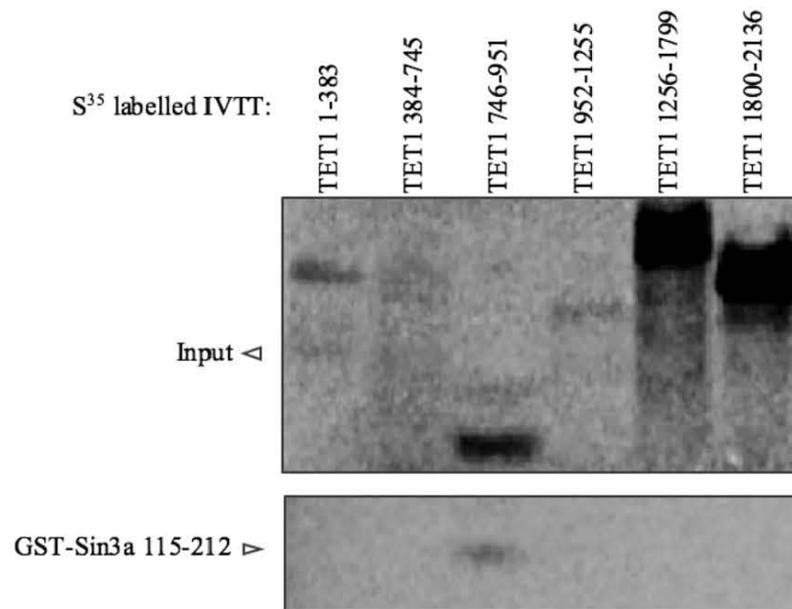


Figure 3.15: Only TET1 746-951 is Capable of Binding to PAH1.

GST pull-down –GST- Sin3a PAH1 was used to pull down S^{35} -Met IVTT TET1 1-383, 384-745, 746-951, 952-1255, 1256-1799 and 1800-2136. GST-PAH1: 115-212. S^{35} -Met IVTT TET1 384-745 would not sufficiently express.

Unlike the C-terminal region of TET1 (1256-1799), which expressed strongly but did not bind, the TET1 384-745 fragment would not express in the experiment depicted in Figure 3.15. Despite this, TET1 384-745 is devoid of secondary structure normally associated with protein-protein interactions (Figure 3.3). This is additionally confirmed by an s2D prediction (Figure 3.16) which produces a model with three possible helical regions in the TET1 384-951 fragment, however, none of these fall within the 384-745 region. Jpred uses evolutionary conservation to predict the structure of a protein sequence, while s2D's predictive ability is derived from a pattern recognition system that was developed based on NMR shift data (Sormanni et al., 2015). It seems unlikely that there is another PAH1 binding site within TET1.

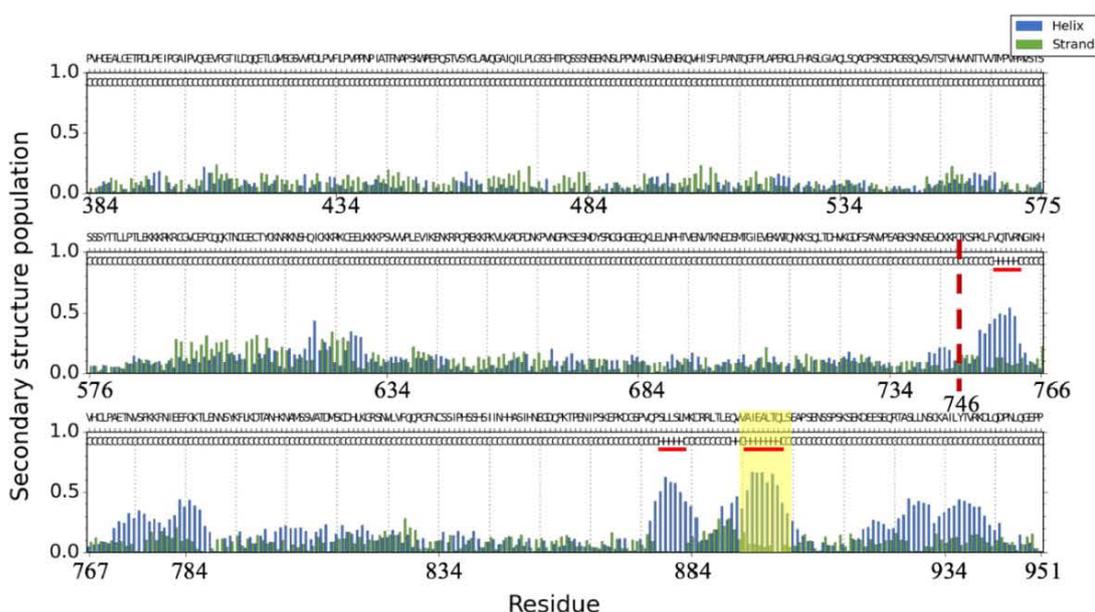


Figure 3.16: TET1 384-745 is Unstructured.

s2D Prediction of Secondary Structure TET1 384-951. The TET1 SID (892-903) sequence is highlighted. Red horizontal line: predicted alpha helical regions. Vertical brown striated line: border between the TET1 384-745 and 746-951 fragments.

3.7. Characterizing the Residues of the TET1-SID

The putative SID, in TET1, was shown to be the critical region for the formation of a complex with Sin3a and the TET1 A896, A893, L900, L897 and I894 residues appear to disproportionately contribute to this interaction. These residues were initially indicated by Multalin (Figure 3.1), and their importance was suggested by the helical wheel projection (Figure 3.4). Based on previous studies of SID/PAH domain interactions, it is the hydrophobic side of the helix that is responsible for binding to the PAH domain. This would be energetically favorable as it would diminish the exposure of the hydrophobic amino acids A896, A893, L900, L897 and I894 to aqueous solution (Figure 3.17 (c)). Several GST-TET1-SID mutations, based around these hydrophobic residues, were produced to establish their effect on binding. The only mutants capable of interacting with S³⁵-Met Sin3a were TET1-SID^{A893D/A896D} (A893D/A896D) and TET1-SID^{I894A} (I894A), as can be seen in Figure 3.17 (a) and (b). TET1-SID^{L897E/L900E}, TET1-SID^{L897A/L900A}, TET1-SID^{I894E} and TET1-SID^{I894Q} did not bind. The A893D/A896D mutant (Figure 3.17 (b)) could bind Sin3a despite the increase in bulk,

hydrophilicity and charge; the binding, however, is barely detectable. L900 and L897 are important residues as the relatively inoffensive leucine to alanine transition appears to abolish binding. This can be seen in both Figure 3.17 (a) and (b). I894A was tolerated, but the polar (Q) and charged (E) substitutions were not. The importance of I894 is difficult to assess, the reduction in signal was not nearly that observed in the case of A893D/A896D, though this latter construct contained two charged mutations. L900, L897 and possibly the I894 residues appear to be the most important in the SID (Figure 3.17 (c)). This mirrors the Multalin overlap between TET1 and TET3 (Figure 3.1), but ends much after the T898 residue predicted to be at the terminus of the SID by Jpred (Figure 3.3).

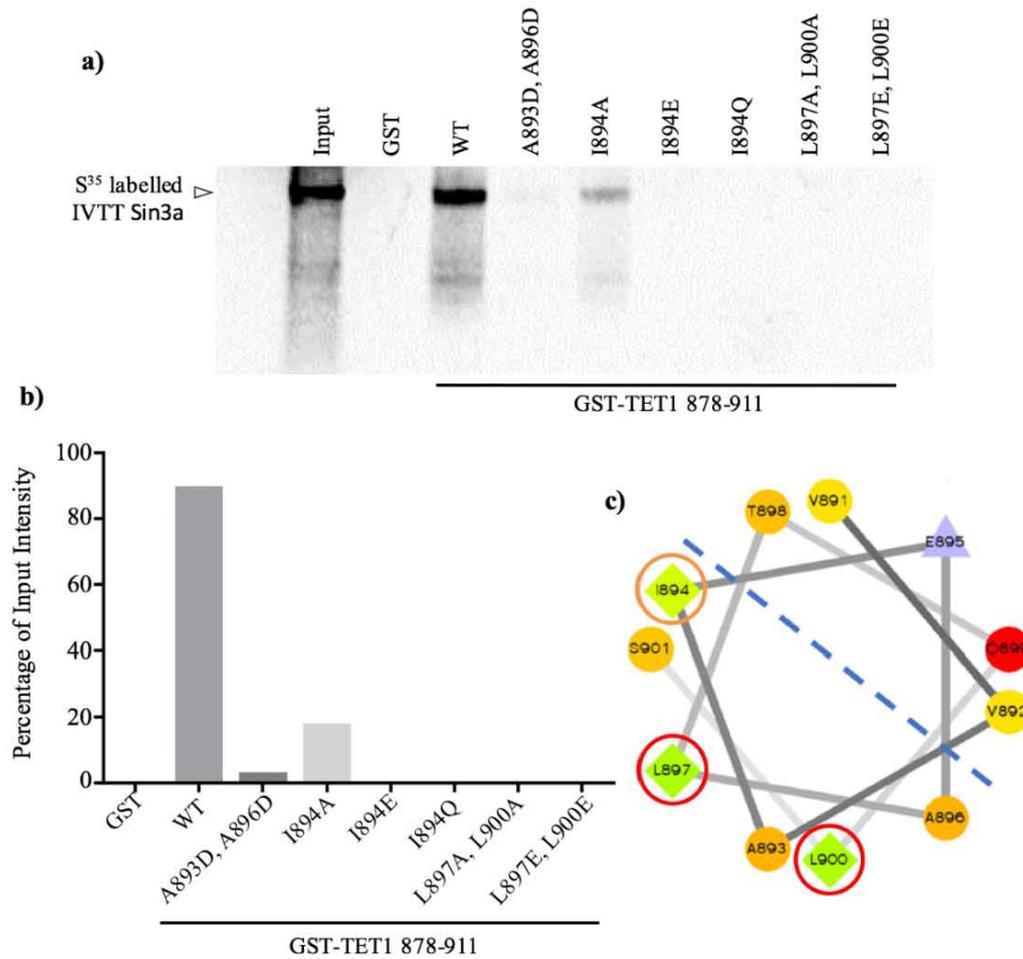


Figure 3.17: L89A, L900 and I894 Are Critical Residues in the TET1 SID.

a) GST pull-down – GST-TET1-SID mutants were used to pull down S35-Met IVTT Sin3a. b) a bar chart representing the relative intensities observed in the GST pull-down, normalized for input signal. c) a helical wheel plotting the TET1-SID with the most critical residues circled.

3.8. Characterizing Sin3a:TET1 Binding

While the GST-pulldowns and the co-IPs led to the discovery of the TET1-SID (Figure 3.7) and the TET1 binding region of Sin3a (Figure 3.14), a technique of greater sensitivity was needed to investigate the importance of single amino acid residues. Gel filtration/size exclusion chromatography and nuclear magnetic resonance (NMR) were used to address this issue. In addition to this, there were several questions that arose during the previous experiments. SAP25's SID appeared to always have a higher signal for Sin3a binding than the TET1 SID, which was unexpected. Demonstrably the case in

Figure 3.11 and Figure 3.12, particularly in the former, as the TET1 fragment required a longer exposure to be imaged. Moreover, the problem of whether Sin3a and TET1 bound in a 1:1 ratio remained open to question. It should be noted that only the His-PAH1 was labelled with N¹⁵, which enabled it to be visible to the NMR machine. The binding partners were therefore only indirectly observed by their effects on His-PAH1.

The use of gel filtration and NMR required the production of purified protein, detailed in 2.12. This process, which starts from IPTG induction of *E. Coli* and ends with HPLC column purification, is illustrated in Figure 3.18.

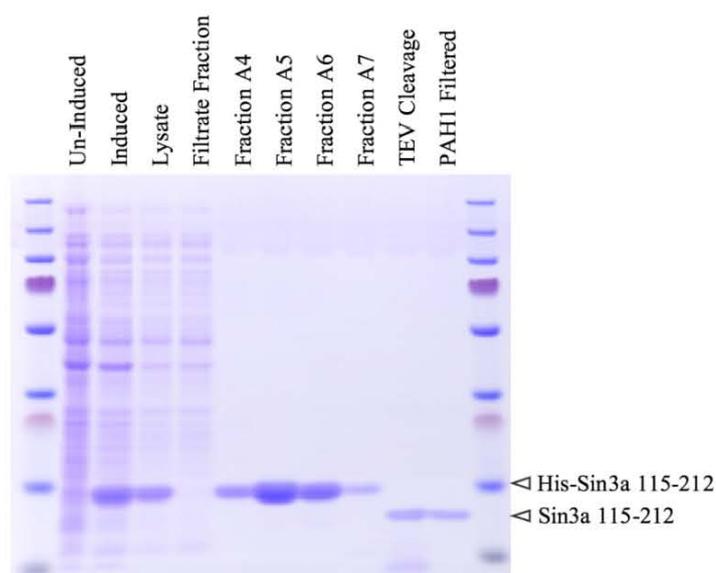


Figure 3.18: SDS-PAGE of His-PAH1 Protein Production and Purification.

Fractions A5 and A6 were high enough in concentration to be combined. Un-Induced: *E. Coli* cells not actively producing the protein of interest (POI). Induced: *E. Coli* induced by IPTG to produce the POI. Lysate: Sonicated induced cells. Filtrate Fraction: HPLC flow through. TEV Cleavage: TEV enzyme removal of His/GB1 tag.

The same process was conducted for His-GB1-SAP25-SID, His-GB1-TET1-SID, His-GB1-TET1 868-927 and His-GB1-TET1 848-952. These are illustrated in Figure 3.19.

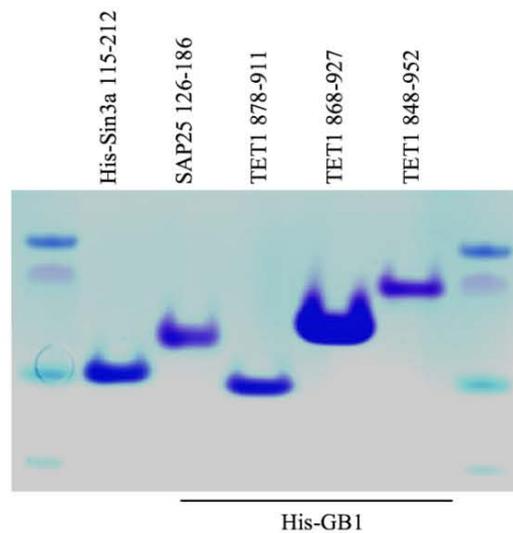


Figure 3.19: SDS-PAGE of the Final Purified Proteins.

This process demonstrates the difficulty in distinguishing between His-GB1-TET1-SID and His-PAH1.

The His-GB1-TET1-SID and His-PAH1 proteins were used together and in isolation to determine the point at which each of these was eluted from the gel-filtration column. The resultant fractions were collected and run on an SDS-PAGE to confirm the identity of the eluted products. Figure 3.20 demonstrates the formation of a TET1-SID:PAH1 complex that travels through the column more quickly than the individual complex components. Gel-filtration corroborates the results of the pulldowns seen in both Figure 3.7 and Figure 3.14. This additionally depicts what a wildtype complex looks like migrating through the column, for comparison with mutant complexes.

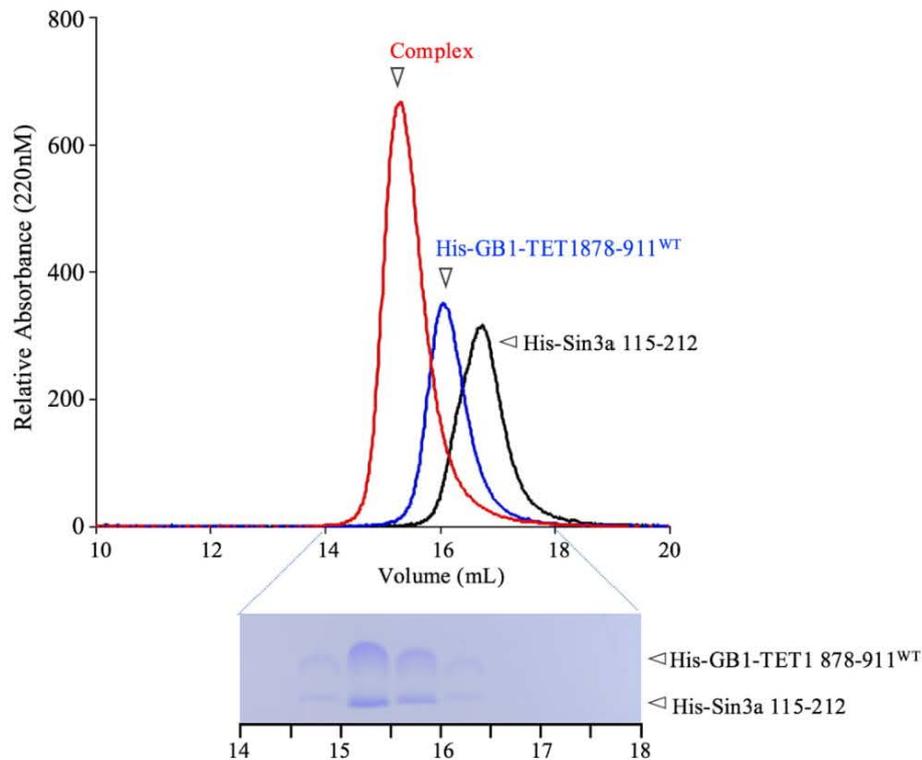


Figure 3.20: Gel-Filtration Demonstrates the His-GB1-TET1-SID and His-PAH1 Interaction.

Gel-Filtration of His-GB1-TET1-SID and His-Sin3a 115-212. This is verified by SDS-PAGE. The SDS-PAGE was conducted using a 16% gel.

The NMR spectra in Figure 3.21 also illustrates the chemical shift between unbound His-PAH1, and His-PAH1 bound to His-GB1-TET1-SID. There is a large conformational change in the PAH1 domain on TET1 binding, demonstrated by the clear majority of His-PAH1 peaks shifting their position. In subsequent experiments the mutant His-GB1-TET1-SID constructs were combined with His-PAH1, and their spectra compared to that of the wildtype complex and, in the case of poor binding, the spectra of the unbound His-PAH1.

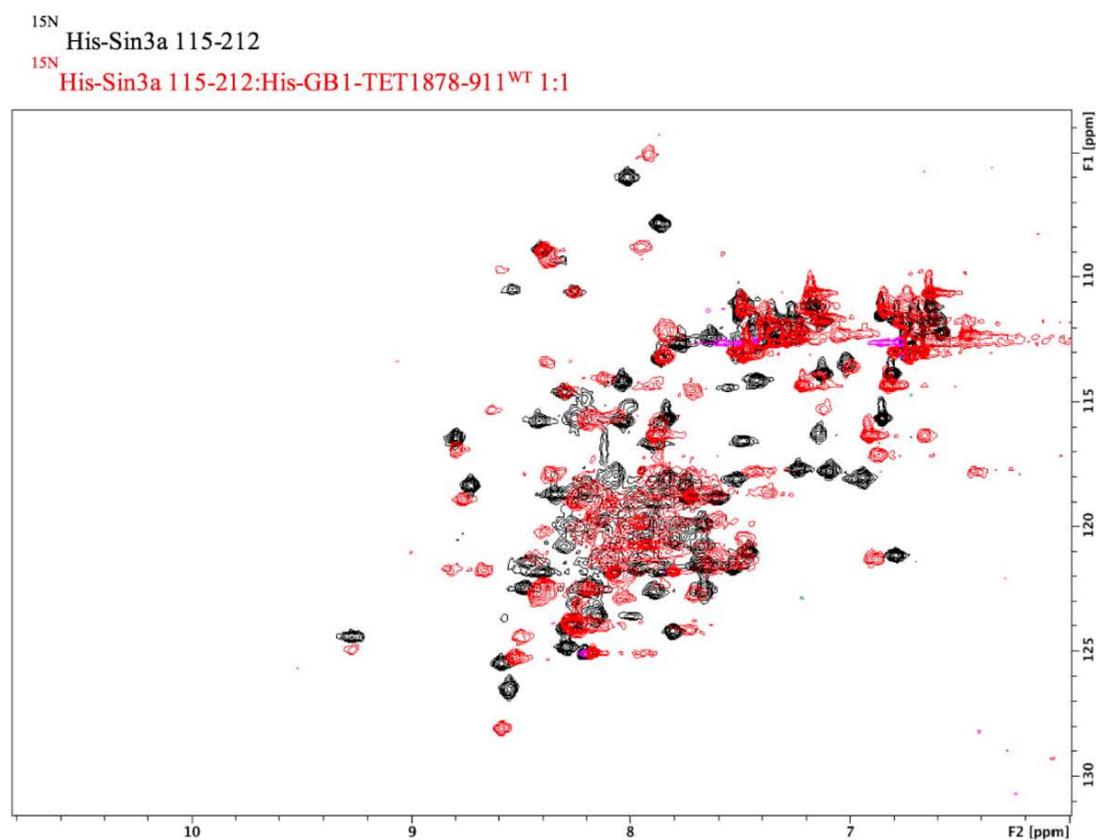


Figure 3.21: NMR Spectra Demonstrates the His-GB1-TET1-SID and His-PAH1 Interaction.

Unbound His-PAH1 (black). His-PAH1 binding unlabeled His-GB1-TET1-SID^{WT} (red).

Three different preparations of His-PAH1 and His-GB1-TET1-SID^{WT} (TET1-SID^{WT}), were made in the ratios 1:1, 1:2 and 1:0.5. Each of these protein samples was run through the gel-filtration column. Whenever there was an excess of either constituent it would come off the column by itself, unattached. This is observed in the second maxima present in the gel-filtration curves of Figure 3.22, that coincide with the complex component's individual gel-filtration curve maxima, as found in Figure 3.20. The SDS-PAGE in Figure 3.22 corroborates this, the interaction between the two proteins in equimolar.

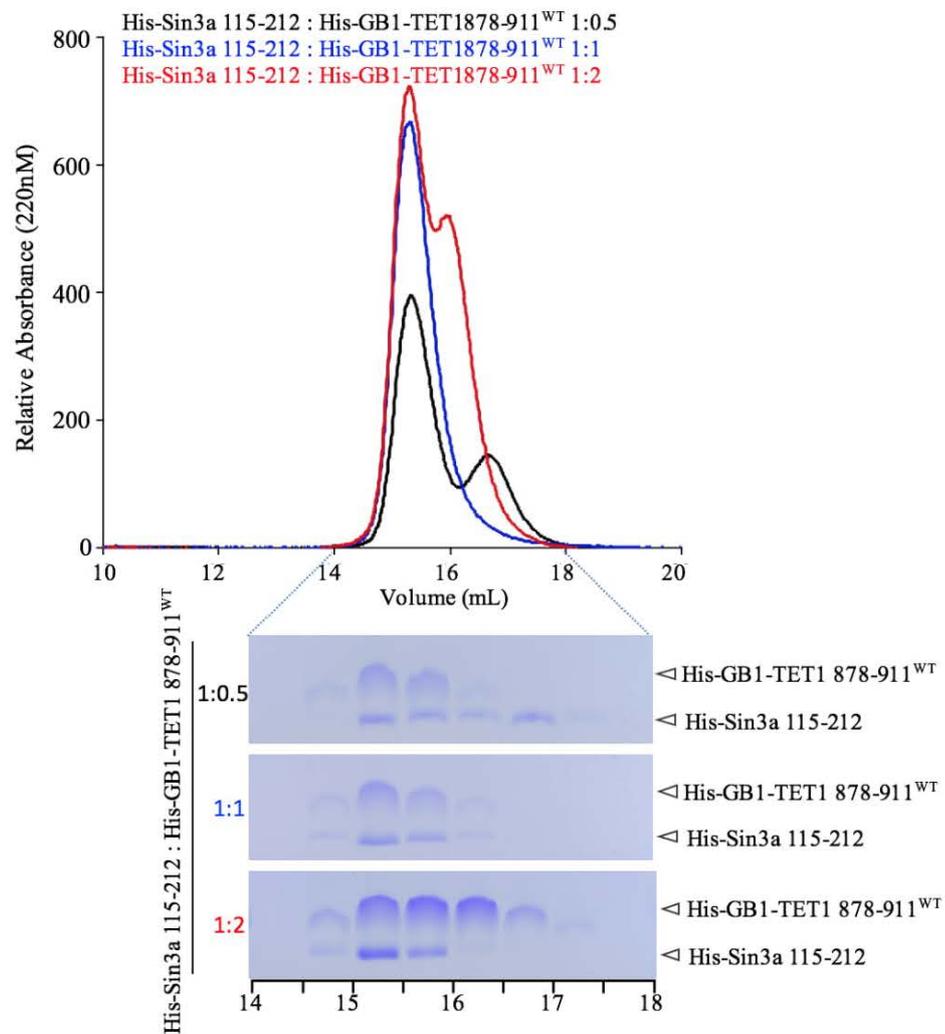


Figure 3.22: TET1 and Sin3a Bind in a 1:1 Ratio.

Gel-filtration and coupled SDS-PAGE. His-PAH1 and His-GB1-TET1^{WT} are examined in 1:1, 1:2 and 1:0.5 ratios. Red: 1:2 complex, Blue: 1:1 complex, Black: 1:0.5 His-PAH1: TET1-SID^{WT} complex. The SDS-PAGE was conducted using a 16% gel.

Multiple experiments suggested that the SAP25-SID bound Sin3a more efficiently than TET1-SID. Figure 3.23 reveals that SAP25-SID binds 9.14 times the amount of Sin3a than the TET1 fragment.

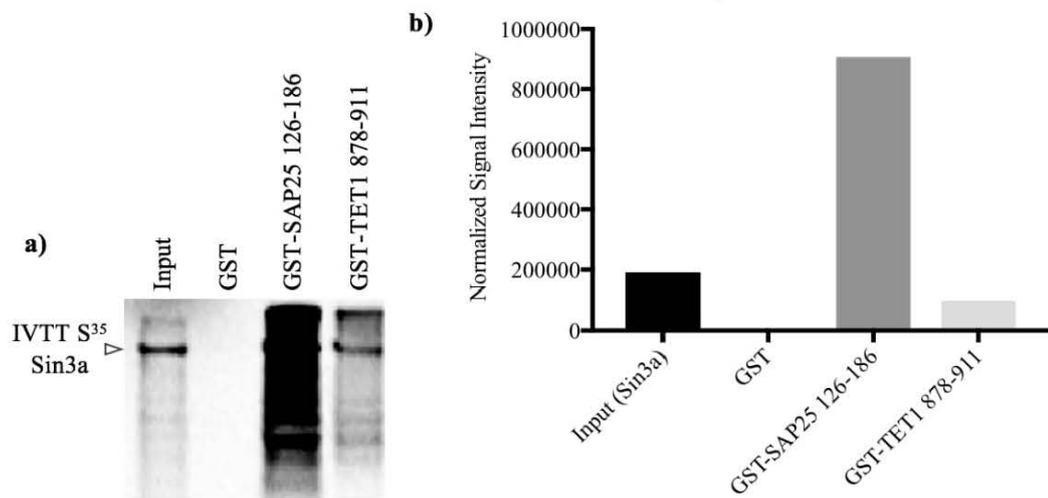


Figure 3.23: SAP25-SID Binds Nearly an Order of Magnitude More Sin3a Than the TET1-SID.

a) GST pull-down – GST-SAP25-SID and GST-TET1-SID were used to pull down S35 IVTT Sin3a. b) a bar chart representing the relative intensities observed in the GST pull-down.

His-GB1-SAP25-SID and His-PAH1 were run through the gel-filtration column both together and in isolation to determine the point each of these eluted from the column. SDS-PAGE was used to confirm the presence of the proteins in the appropriate fractions. Both the gel filtration and the SDS-PAGE can be seen in Figure 3.24. When compared to Figure 3.20, the gel-filtration peaks for the His-GB1-SAP25-SID complex and the TET1-SID^{WT} complex appear similar.

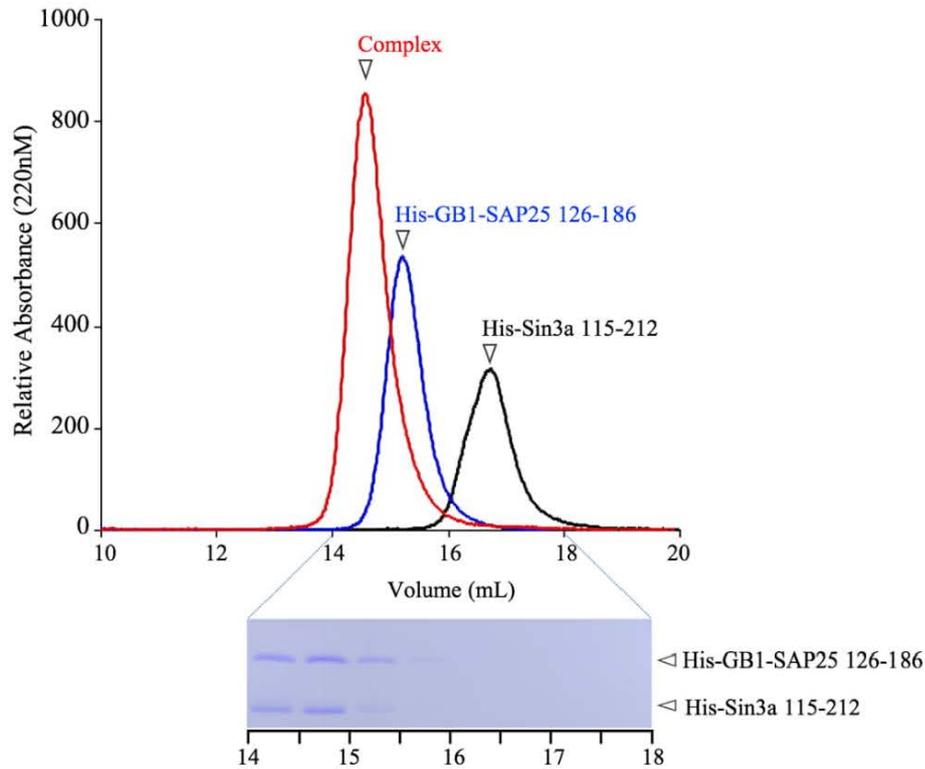


Figure 3.24: Gel-Filtration Demonstrates the His-GB1-SAP25-SID and His-PAH1 Interaction.

Gel-Filtration His-GB1-SAP25-SID and His-PAH1. Red: His-PAH1:His-GB1-SAP25-SID complex, Blue: His-GB1-SAP25-SID, Black: His-PAH1. This is verified by SDS-PAGE. The SDS-PAGE was conducted using a 16% gel.

An equimolar solution was prepared of His-PAH1, TET1-SID^{WT} and His-GB1-SAP25-SID. This was run through the gel filtration column. Just as in Figure 3.22, the gel-filtration elution curve in Figure 3.25 (red) reveals an extra peak of the protein that is present in excess. This extra peak belongs to TET1-SID^{WT}. SDS-PAGE reveals that the entirety of TET1-SID^{WT} is unbound. It appears that SAP25 entirely outcompetes TET1 for PAH1, explaining the results seen in Figure 3.23.

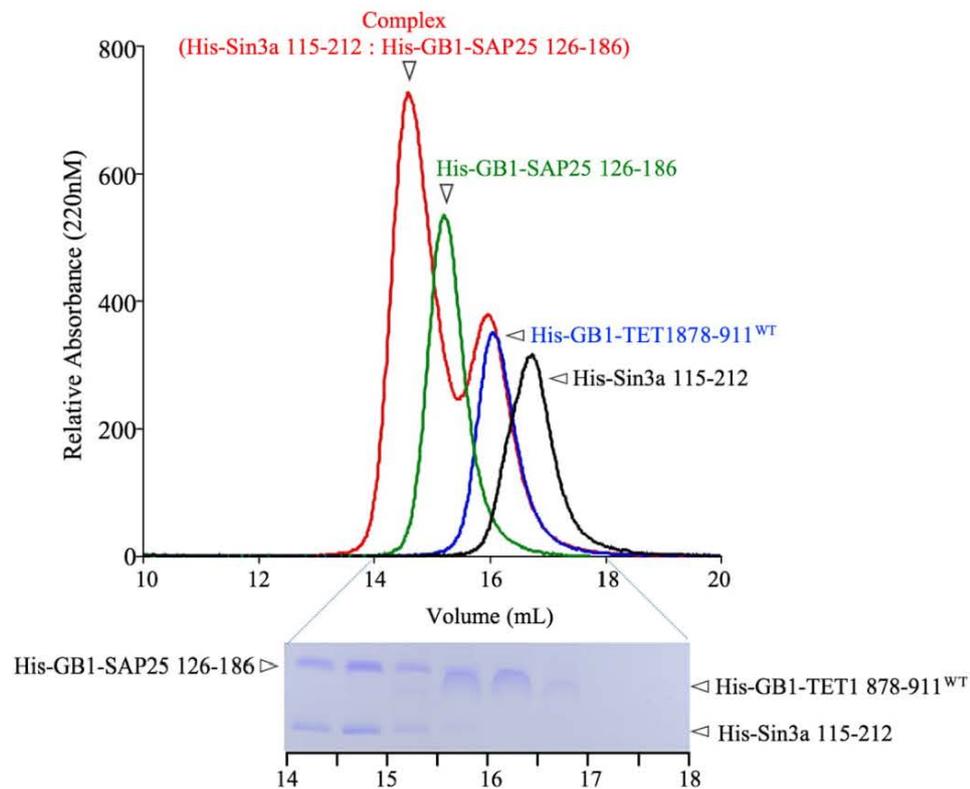


Figure 3.25: SAP25 Outcompetes TET1 for Sin3a Binding.

Gel-filtration and coupled SDS-PAGE. His-PAH1, His-GB1-SAP25-SID and His-GB1-TET1 are examined in 1:1:1 ratio. Red: His-PAH1:His-GB1-SAP25-SID complex, Green: His-GB1-SAP25-SID, Blue: TET1-SID^{WT}, Black: His-PAH1. The SDS-PAGE was conducted using a 16% gel.

3.9. Identifying the Importance of Individual Residues within the SID

The gel-filtration curve (Figure 3.26) for the complex, between His-GB1-TET1-SID^{L897A} (L897A) and His-PAH1 (red), has two maxima that coincide with each of the complex components in isolation (blue and black). There was no elution at the point wildtype His-PAH1:TET1-SID^{WT} complex came off the column in Figure 3.20. Taken together, it appears that the L897A mutation nearly abolished complex formation. This is likely due to the drastic reduction in hydrophobicity of the SID because of the loss of the leucine residue. The leucine side-chain appears to be essential in contacting residues within PAH1.

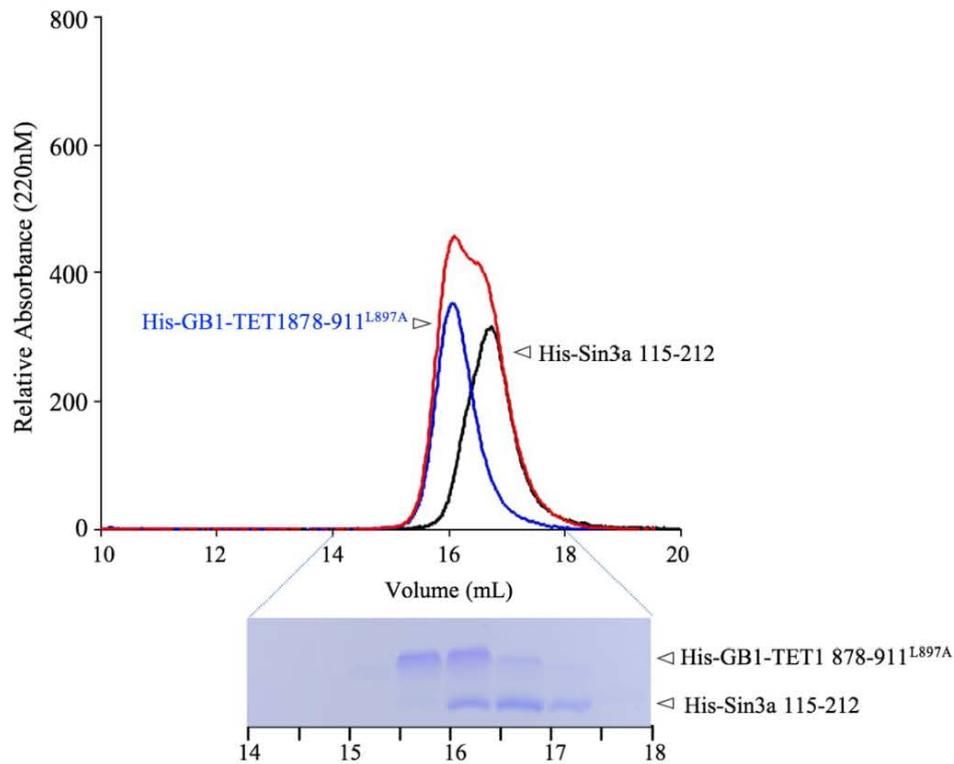


Figure 3.26: Gel-Filtration Demonstrates the L897 Residue is Critical for Sin3a Interaction.

Gel-Filtration of the His-PAH1:His-GB1-L897A complex. Red: His-PAH1:His-GB1-L897A complex, Blue: His-GB1-L897A, Black: His-PAH1. This is verified by SDS-PAGE. The SDS-PAGE was conducted using a 16% gel.

The NMR spectra (Figure 3.27) of the His-PAH1:His-GB1-L897A complex (red) looks remarkably similar to the unbound His-PAH1 (black). Gel filtration and SDS-PAGE suggest the proteins do not interact. Despite this, there exists a minimal NMR chemical shift – the overlap between the NMR spectra is imperfect. This indicates that the binding occurs, but is incredibly weak.

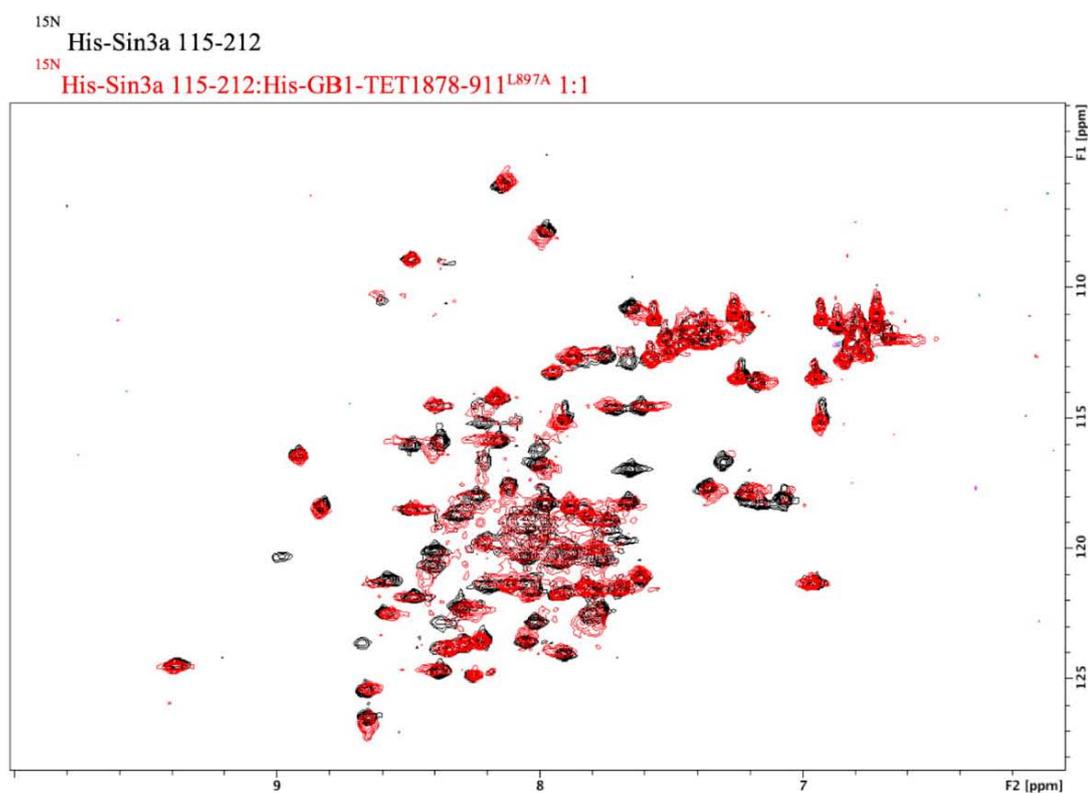


Figure 3.27: NMR Spectra Demonstrates the L897A Mutant Nearly Prevents PAH1 Binding.

The peaks are very similar to unbound His-PAH1. Unbound His-PAH1 (black). His-PAH1 binding unlabeled His-GB1-L897A (red).

When compared to the wildtype complex (black), the L897A (red) chemical shift is quite extreme (Figure 3.28). The lack of overlap between the spectra reconfirms the extremely weak binding of the L897A construct seen in Figure 3.27.

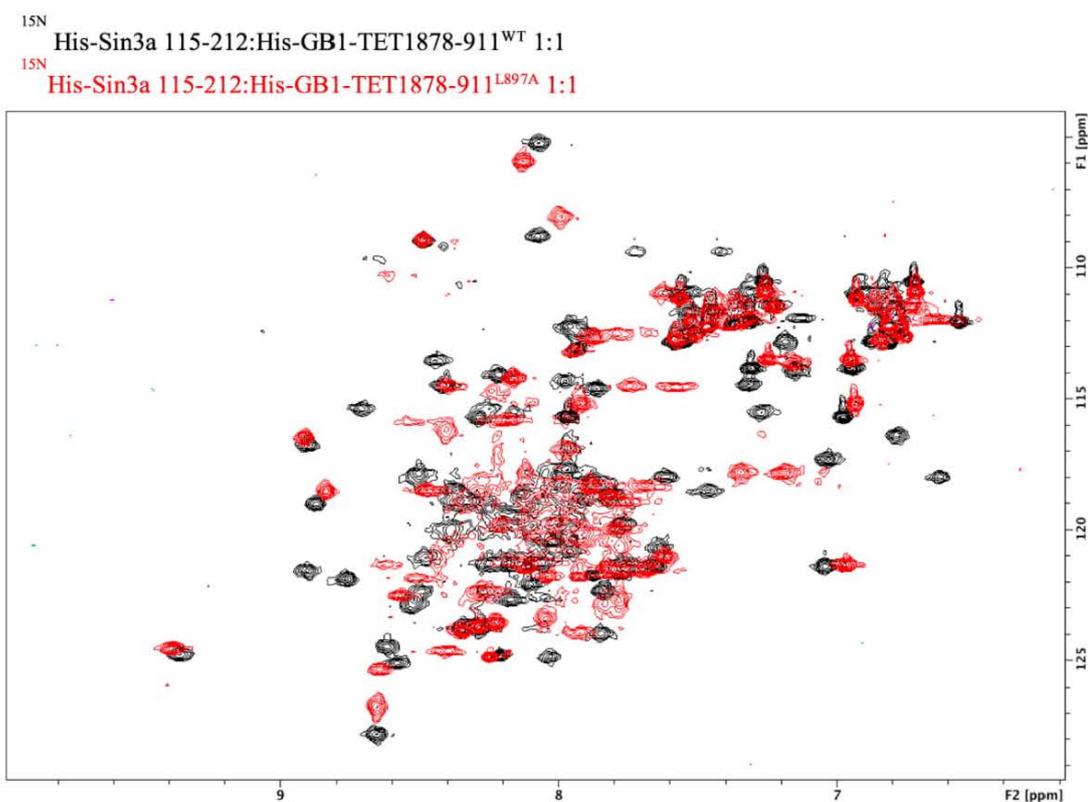


Figure 3.28: NMR Spectra Demonstrates the L897A Mutation Results in Weak His-PAH1 Binding.

The peaks do not resemble those of His-PAH1 bound to wildtype His-GB1-TET1-SID. His-PAH1: TET1-SID^{WT} (black). His-PAH1 binding His-GB1-L897A (red).

The gel-filtration curve (Figure 3.29) for the complex between His-GB1-TET1-SID^{I894A} (I894A) and His-PAH1 (red), reveals a spectra that is similar, but not identical, to the wildtype complex (Figure 3.20). The SDS-PAGE depicted in Figure 3.29, however, reveals that a portion of both (14.5mL fraction) His-PAH1 and (16mL fraction) the I894A construct are not involved in complex formation. It appears that the I894A mutation negatively impacts complex formation, though less severely than the L897A mutation. This is possibly due to a decrease in size and hydrophobicity. It is possible that I894 would be towards the edge, if L897 is at the center, of PAH1's hydrophobic pocket; explaining the difference in the residues' influence on binding.

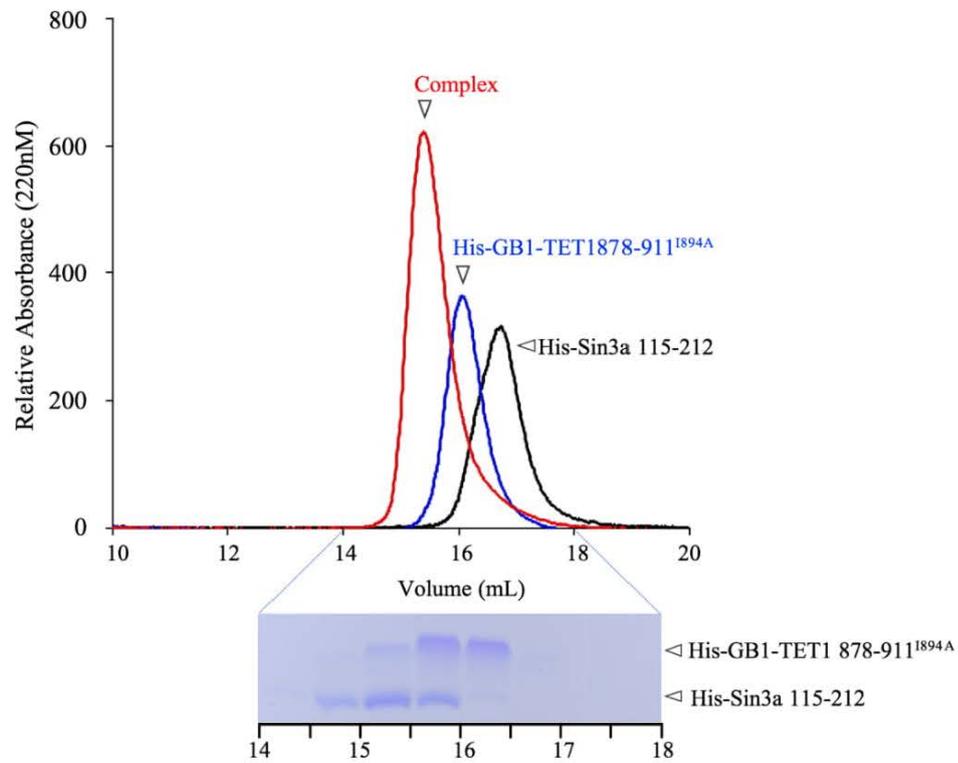


Figure 3.29: SDS-PAGE Demonstrates That the I894A Mutation Weakens His-PAH1 Binding.

Gel-Filtration of the His-PAH1:His-GB1-I894A complex. Red: His-PAH1:His-GB1-I894A complex: His-GB1-I894A, Blue: His-GB1-TET1 I894A, Black: His-PAH1. The SDS-PAGE was conducted using a 16% gel.

The NMR spectra (Figure 3.30) reveals that the binding of the I894A construct is similar to TET1-SID^{WT} indicating a subtle change in binding.

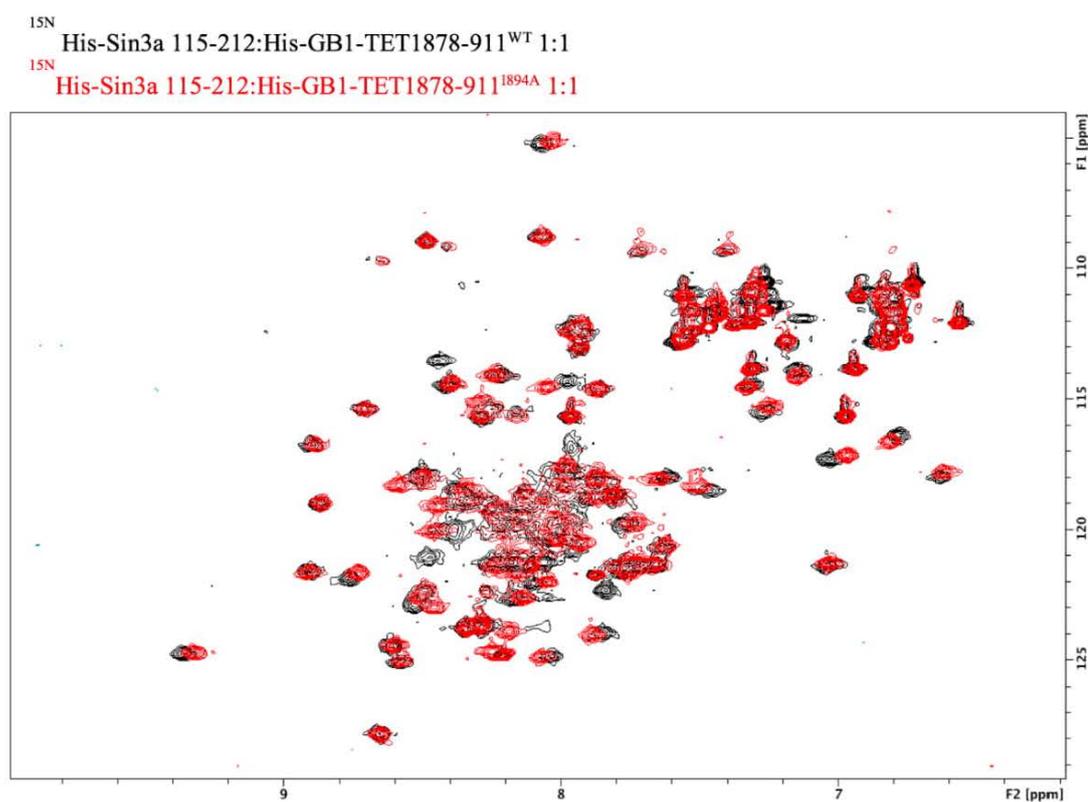


Figure 3.30: NMR Spectra Demonstrates the I894A Mutation Results in Mildly Weaker His-PAH1 Binding.

The peaks partially resemble those of His-PAH1 bound to wildtype His-GB1-TET1-SID. His-PAH1:TET1-SID^{WT} (black). His-PAH1 binding His-GB1-I894A (red).

According to the helical wheel projection (Figure 3.4), T898 would sit opposite the hydrophobic face of the amphipathic helix that forms the SID. Accordingly, both the gel-filtration and SDS-PAGE of the His-PAH1:His-GB1-TET1-SID^{T898A} (T898A) complex (Figure 3.31) reveal T898A has no effect on binding (Figure 3.20).

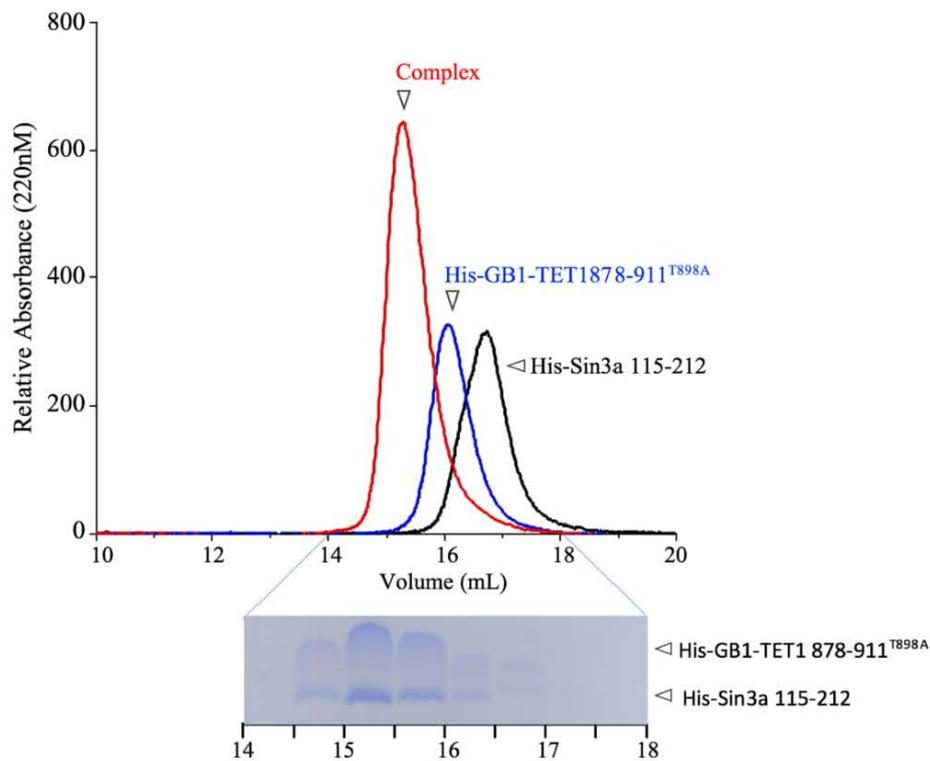


Figure 3.31: Gel-Filtration Demonstrates T898 is Unimportant in the TET1:Sin3a Interaction.

Gel-Filtration of the His-PAH1:His-GB1-T898A complex. Red: His-PAH1:His-GB1-T898A complex, Blue: His-GB1-T898A, Black: His-PAH1. This is verified by SDS-PAGE. The SDS-PAGE was conducted using a 16% gel.

The NMR spectra of T898A (Figure 3.32) is virtually identical to the wildtype complex, in agreement with the gel-filtration and SDS-PAGE data (Figure 3.31)

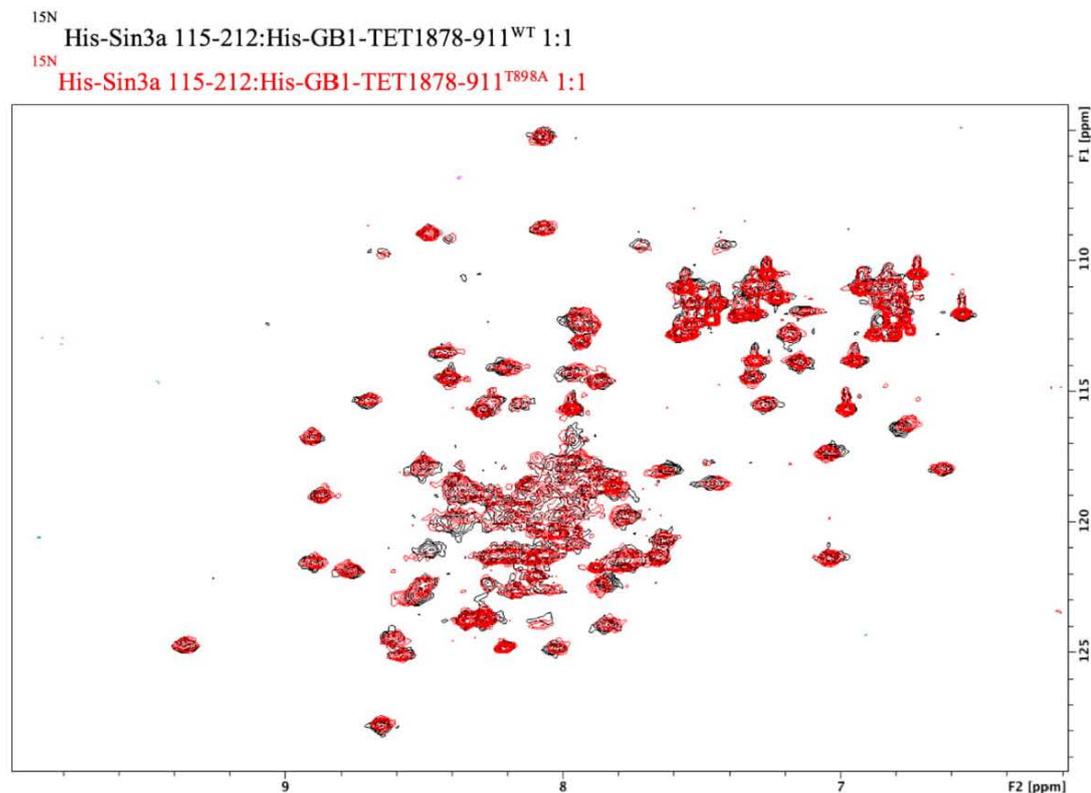


Figure 3.32: NMR Spectra Demonstrates the T898 Residue is Not Critical for His-PAH1 Binding.

The peaks strongly resemble those of His-PAH1 bound to wildtype His-GB1-TET1-SID. His-PAH1:TET1-SID^{WT} (black). His-PAH1 binding His-GB1-T898A (red).

As the threonine (T898) residue sits next to a glutamate (Q899) residue within the SID, it is a possible ‘TQ site’ for phosphorylation by the ATM, ATR and DNA-PK family of kinases (Kim et al., 1999). This led to the creation of the His-GB1-TET1 878-911^{T898E} (T898E) mutant, as glutamic acid substitution is an established method of constitutively mimicking the phosphorylation of a threonine residue (Brown et al., 1998). It should be noted that While Brown et al. (1998) indicate that the glutamic acid substitution for threonine is a suitable phosphomimetic, there are caveats to its use. Dephore and colleagues (2013) indicate that the charge of the glutamate residue is not as negative as the phosphothreonine and, relatedly, has a smaller ionic shell. These differences produce divergent chemical environments. As a result, the use of threonine substitutions to investigate phosphorylation has not always been experimentally successful (Durocher et al., 1999).

The T898E complex's (red) gel-filtration curve (Figure 3.33) does show a slightly strange right tail. This is indicative of an inadequate amount of salt in the gel-filtration buffer. Aside from this, the gel-filtration curve for the T898E complex is indistinguishable from both the wildtype complex (Figure 3.20) and the T898A complex (Figure 3.31).

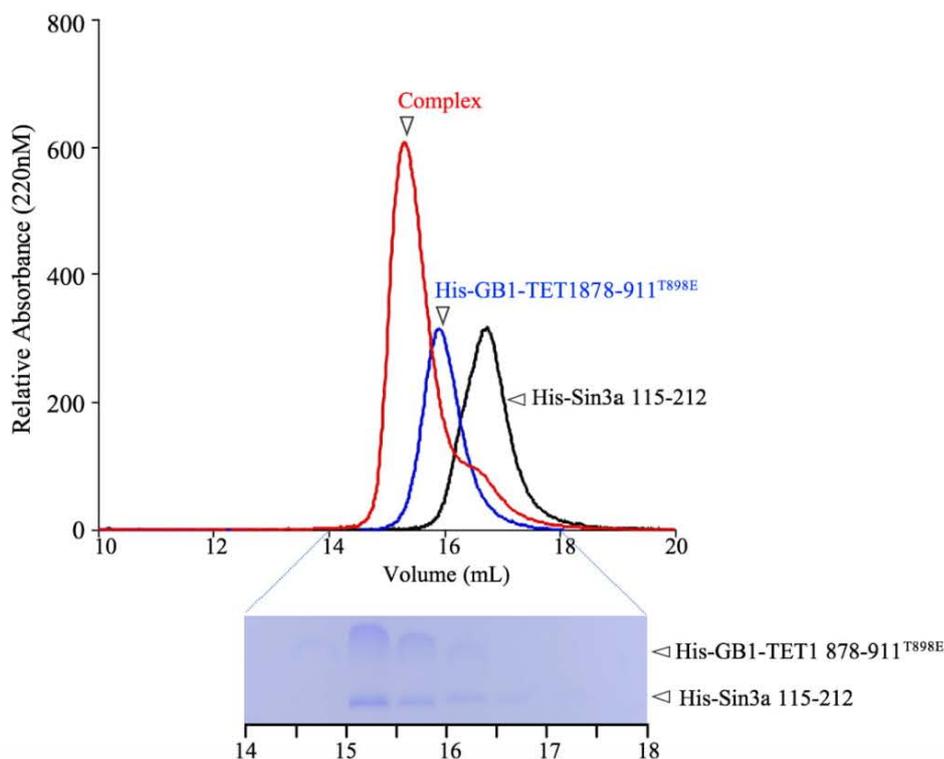


Figure 3.33: Gel-Filtration Demonstrates T898 is Unlikely to be a Phosphorylation Site.

Gel-Filtration of the His-PAH1:His-GB1-T898E complex. Red: His-PAH1:His-GB1-T898E complex, Blue: His-GB1-T898E, Black: His-PAH1. This is verified by SDS-PAGE. The SDS-PAGE was conducted using a 16% gel.

Like the T898A mutation, the NMR spectra of T898E complex (Figure 3.34) overlaps with the wildtype complex's spectra quite closely. It is not as close to the wildtype spectra as the T898A complex (Figure 3.32). Despite this, when taken together with the gel-filtration spectra and SDS-PAGE in Figure 3.33, it is unlikely that the T898 is a phosphorylation site.

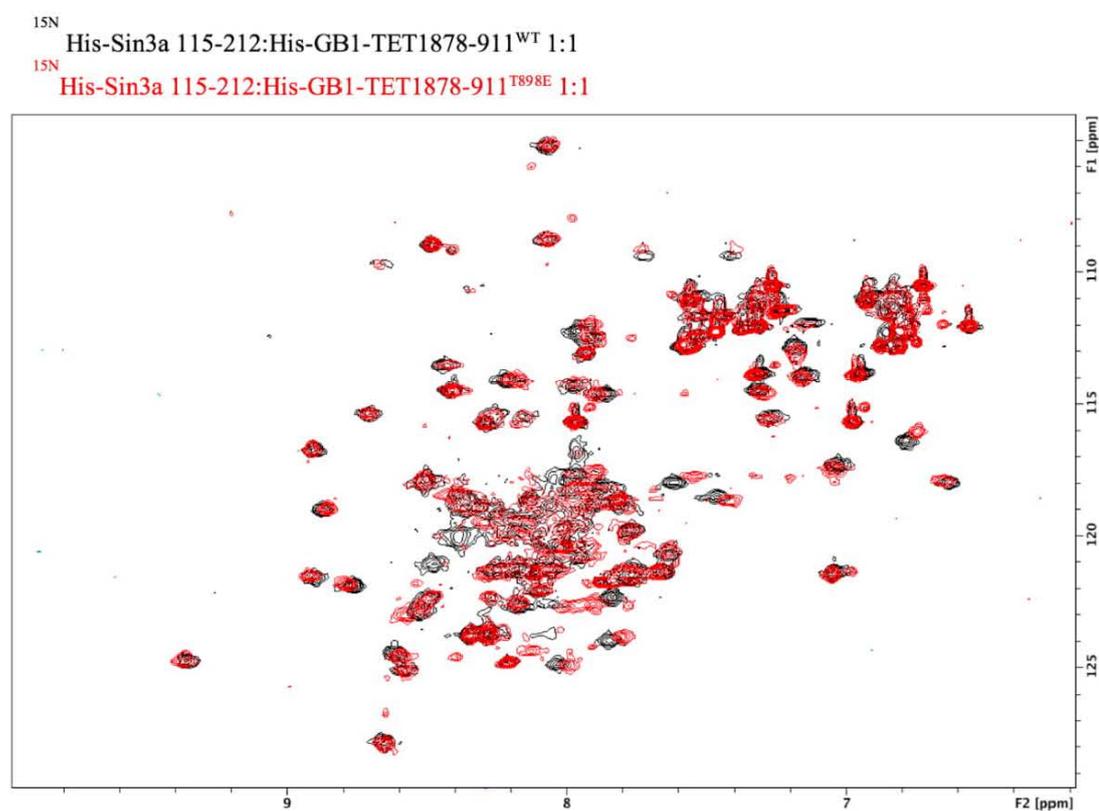


Figure 3.34: NMR Spectra Demonstrates T898E's Effects Are Nearly Identical to That of T898A and WT.

The peaks strongly resemble those of His-PAH1 bound to wildtype His-GB1-TET1-SID. His-PAH1:TET1-SID^{WT} (black). His-PAH1 binding His-GB1-T898E (red).

The gel-filtration maxima (red) for the His-PAH1:His-GB1-TET1 878-911^{L900A} (L900A) complex (Figure 3.35) shifts rightwards from the wildtype complex (Figure 3.20). Moreover, the maxima has an exaggerated right tail which appears to coincide with His-PAH1, suggesting that a portion of the protein is unbound. This is confirmed by the accompanying SDS-PAGE, where the 17mL fraction consists largely of unbound His-PAH1.

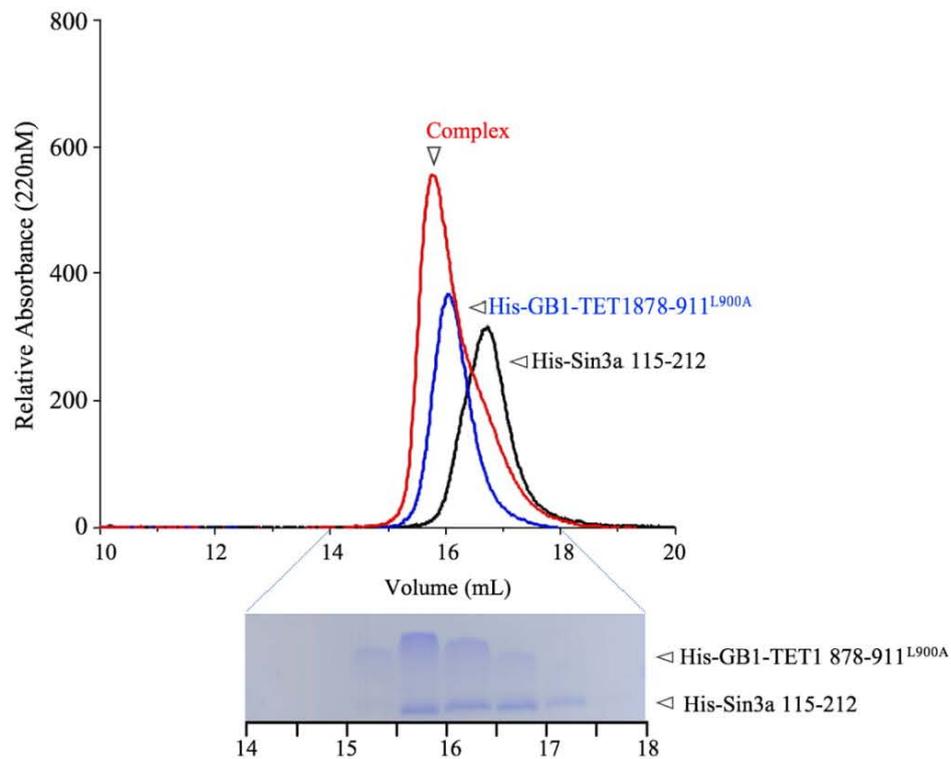


Figure 3.35: Gel-Filtration Demonstrates the L900 Residue is Important for Sin3a Interaction, Though Less Important Than L897.

Gel-Filtration of the His-PAH1:His-GB1-L900A complex. Red: His-PAH1:His-GB1-L900A complex, Blue: His-GB1-L900A, Black: His-PAH1. This is verified by SDS-PAGE. The SDS-PAGE was conducted using a 16% gel.

The clear majority of peaks in the NMR spectra for the L900A complex (Figure 3.36) have shifted in relation to those of the wildtype complex. This degree of shift is reminiscent to that seen by the L897A complex (Figure 3.28). There are some peaks that coincide with those of the wildtype complex, suggesting that, while severe, the reduction in binding is less pronounced than that caused by L897A.

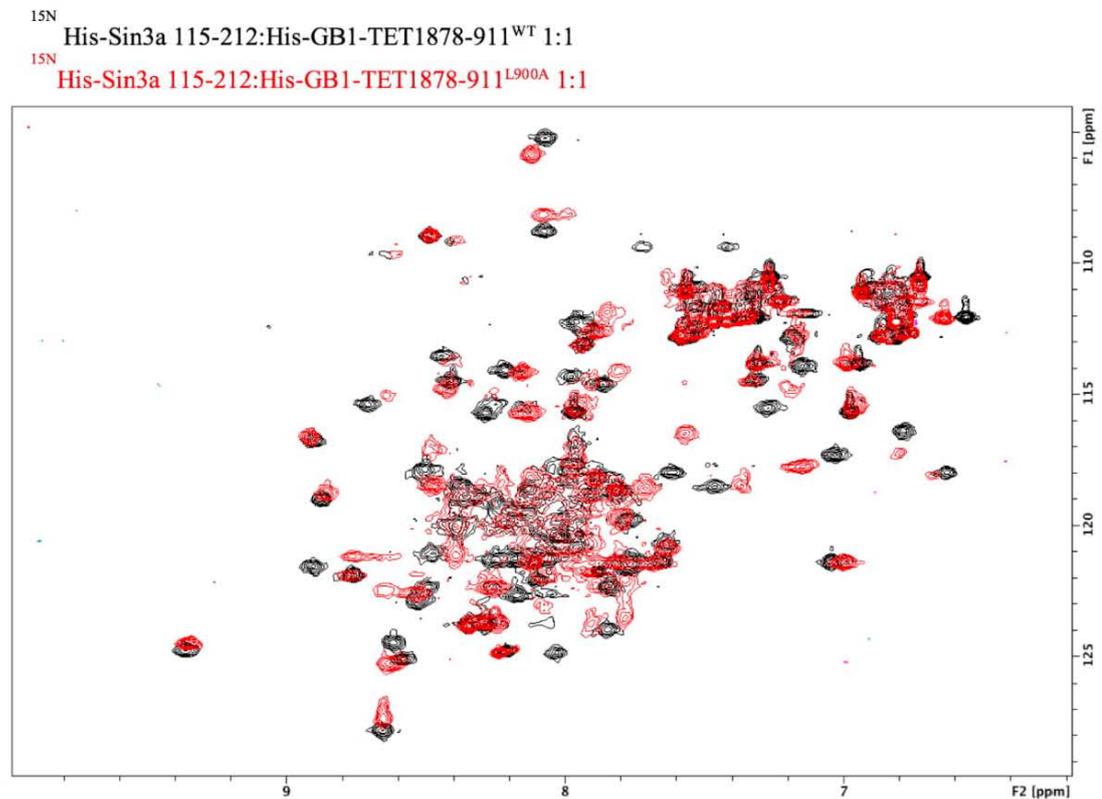


Figure 3.36: NMR Spectra Demonstrates the L900A Mutation Results in Weak His-PAH1 Binding.

The peaks do not resemble those of His-PAH1 bound to wildtype His-GB1-TET1-SID. His-PAH1:TET1-SID^{WT} (black). His-PAH1 binding His-GB1-L900A (red).

The gel filtration, SDS-PAGE and NMR (Figure 3.20-Figure 3.36) confirmed that the TET1-SID binds directly to PAH1, and in a 1:1 manner. It also answered the question as to why the GST pulldowns always presented SAP25-SID showing a much stronger preference for Sin3a than any of the TET1 constructs. Perhaps even more importantly, the triumvirate of methods revealed that the residues I894, L897 and L900 of the SID play a large role in Sin3a binding. The corresponding I894A and L900A His-GB1-TET1-SID mutants all migrate through the gel-filtration column in an identical manner (Figure 3.37). This indicates that changes in complex migration through the column are likely due to differences in binding, rather than the altered movement of one of the components through the column.

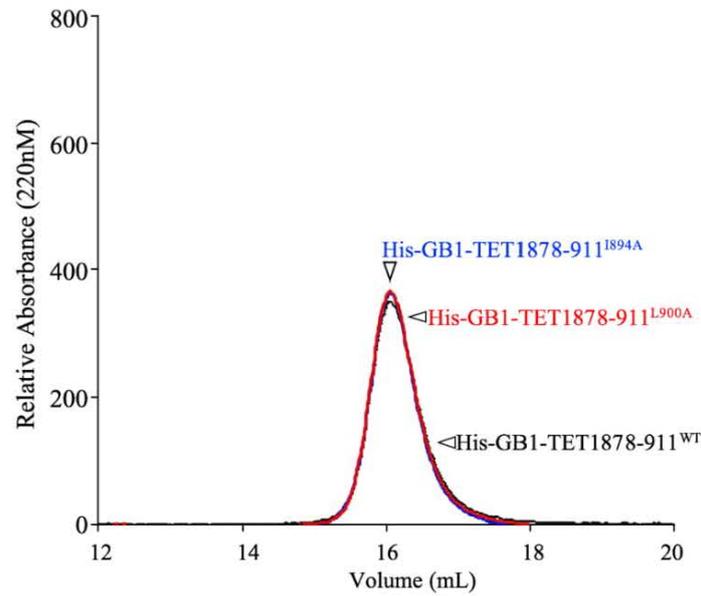


Figure 3.37: His-GB1-I894A and His-GB1-L900A Migrate Through the Column Identically to TET1-SID^{WT}.
Gel-Filtration of (Red) His-GB1-L900A, (Blue) His-GB1-L900A and (Black) TET1-SID^{WT}.

The L897 residue appears to be the critical residue for TET1 binding to PAH1. The L897A mutation demonstrated the greatest NMR chemical shift from the wildtype complex (Figure 3.28). In addition, the L897 residue is important because it is flanked by residues on either side, I894 and L900, that significantly impact PAH1 binding (Figure 3.38).

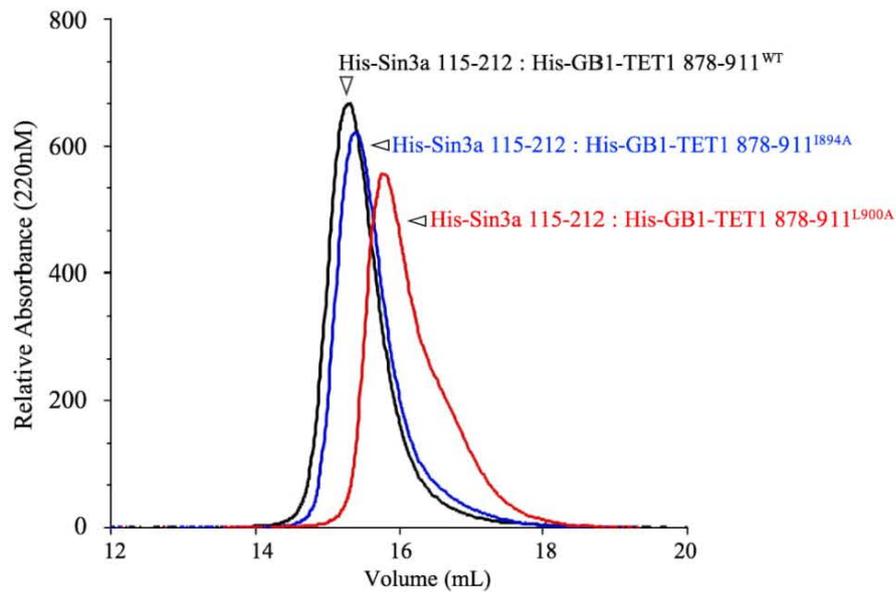


Figure 3.38: L897 Appears to be at the Center of the Interaction with Sin3a. Both the Residues Preceding (I894A) and Succeeding (L900) L897 Impact PAH1 Binding.

Gel-Filtration of (Red) His-PAH1:His-GB1-L900A complex, (Blue) His-PAH1:His-GB1-I894A complex and (Black) His-PAH1:TET1-SID^{WT} complex.

3.10. Characterizing the Full-Length Sin3a:TET1 Interaction

While GST-pulldowns, NMR and gel filtration experiments only permitted the study of solitary TET1 fragments, they incrementally revealed that L897 and L900 were central to Sin3a binding. Studying these residues within a full-length context required the stable integration of *TET1* and *TET1*^{L897A/L900A} in *TET1* KO cells, detailed in 2.11. A co-IP with these ESCs would establish the influence L897A/L900A, a < 0.1% alteration of TET1, has on binding endogenous Sin3a. The technique has the added benefit of illustrating the interaction within a cellular setting.

As this experiment depended on the absence of endogenous TET1, the B8 *TET1/2* DKO needed to first be characterized. RT-PCR established that E14 wildtype ESCs expressed both *TET1* and *TET2* on the RNA level, as seen in lanes 2 and 6 in Figure 3.39. The levels of RNA for these proteins was much lower in the B8 cells (lanes 3 and 7,

respectively). The reduced RNA level in the B8 cells was likely a result of nonsense mediated decay.

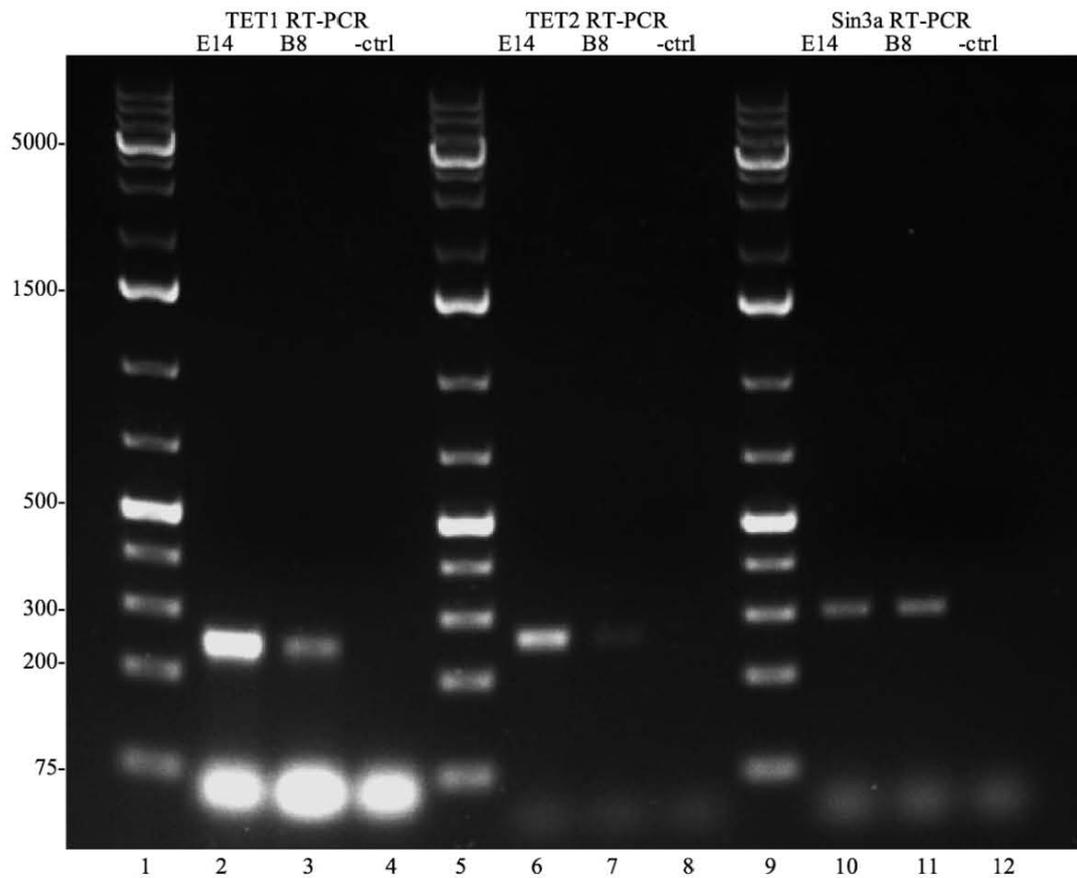


Figure 3.39: B8 Cells Are TET1/2 DKO.

RT-PCR.-ctrl: negative control. 1.5% Agarose, 3% EtBr. Ladder: GeneRuler 1kb plus.

On the protein level, it was found that only the parental E14 cells and a separate population of E14 ESCs previously described by Foster et al. (2010) expressed TET1 (> 190kDa). This is depicted in the top panel of Figure 3.40 in lanes 1 and 5. Neither the A5 *TET1* KO nor the B8 *TET1/2* DKO cells expressed TET1 (top panel, lanes 2 and 3). As differentiated cells, fibroblasts would not express significant TET1, as confirmed in lane 4 of the top panel (Gao et al., 2013).

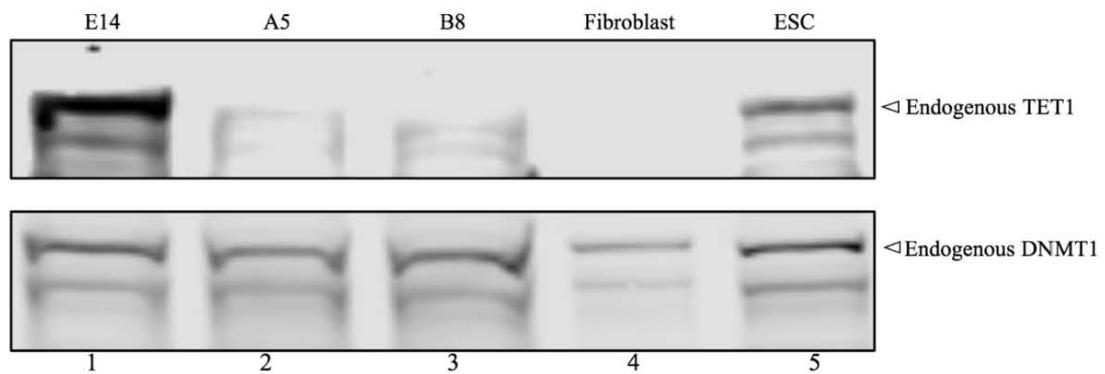


Figure 3.40: B8 Cells Do Not Express TET1 Protein.

Western Blot of endogenous TET1 and DNMT1. The cells investigated were E14 cells and the derivative A5 and B8 cells. ESC: E14 cells unrelated to the KO cells. DNMT1: ~177.5 kDa

These two groups of the E14 cells as well as the A5 cells expressed TET2 (> 190kDa); as respectively depicted in the top panel of Figure 3.41 in lanes 1, 5 and 2. Consistent with their differentiation state, fibroblasts did not express TET1 or TET2 (top panels - lane 4 in Figure 3.40 and Figure 3.41). The B8 DKO demonstrated the same lack of the TET enzymes (top panels - lane 3).

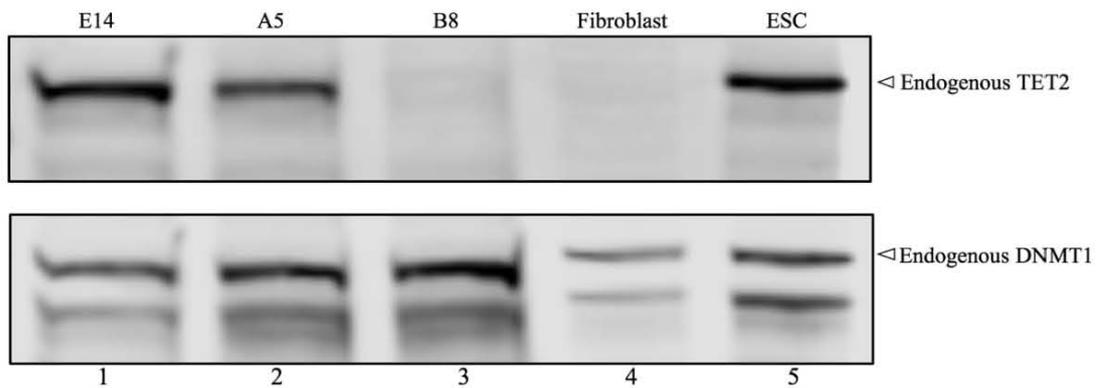


Figure 3.41: B8 Cells Do Not Express TET2 Protein.

Western Blot of endogenous TET2 and DNMT1. The cells investigated were E14 cells and the derivative A5 and B8 cells. ESC: E14 cells unrelated to the KO cells. DNMT1: ~177.5 kDa

The *TET1* reintroduced to the B8 *TET1/2* DKO cells possessed an N-terminal Flag-tag. A western blot confirmed the presence of these constructs (Figure 3.42). All cells were made in triplicate. The three wildtype *TET1* rescue clones A1, A12 and A3 correspond to lanes 2, 3 and 9. The *TET1*^{L897A/L900A} clones C2, C5 and C3 are in lanes 4, 5 and 10. The *TET1*^{ΔCXXC} clones (E6, E10 and E2), made for qRT-PCR, are shown in lanes 6, 7 and 11.

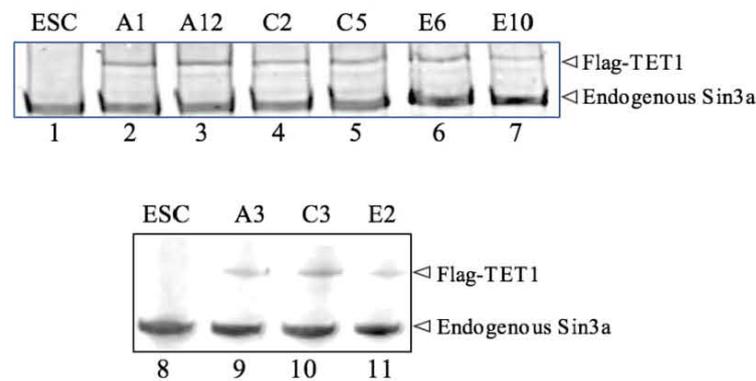


Figure 3.42: All PiggyBAC Rescue Cells Expressed Flag-TET1.

Western blot of Sin3a and flag in Flag-TET1 integrated (via PiggyBac) B8 *TET1/2* DKO cells. ESC: unaltered E14 cells.

A co-IP with the PiggyBAC cells revealed that full-length wildtype Flag-TET1 interacts with endogenous Sin3a (Figure 3.43, lane 3 bottom panel); the L897A/L900A mutations, however, prevented this interaction (lane 6, bottom panel).

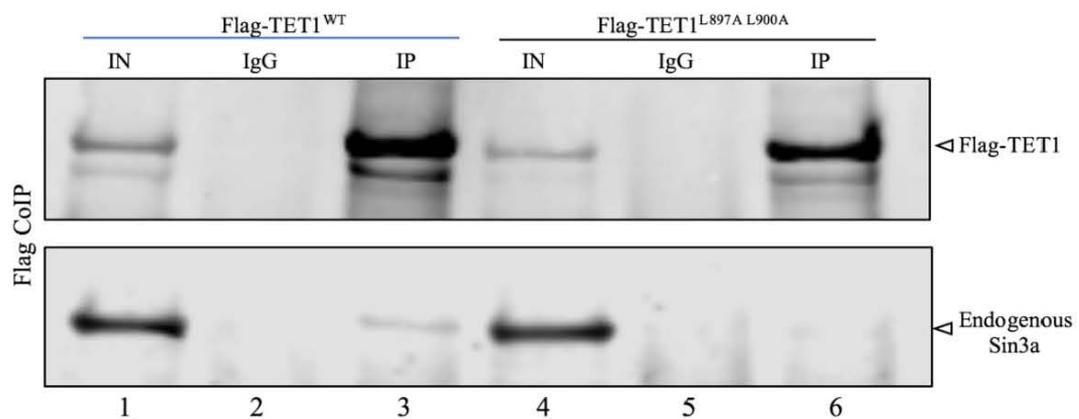


Figure 3.43: The L897A/L900A Mutations Prevent Full-Length TET1 From Interacting with Endogenous Sin3a.
 A co-IP analysis –Flag-TET1 and TET1^{L897A/L900A} proteins were used to pull down endogenous Sin3a in B8 derived ESC. In = input, IgG = negative control, IP = immunoprecipitation.

A second co-IP (Figure 3.44), using the same cells, revealed full-length endogenous Sin3a was able to interact with Flag-TET1 (top panel, lane 3); the TET1 L897A/L900A mutation abolished this interaction (top panel, lane 6). When the results presented in Figure 3.43 and Figure 3.44 are taken together, they demonstrate that the TET1 L897 and L900 residues are essential for the interaction with Sin3a. This confirms the results from the GST-pulldown, gel-filtration and NMR (Figure 3.26, Figure 3.27, Figure 3.28, Figure 3.35, Figure 3.36 and Figure 3.38) are applicable in the context of both the cell and the full-length proteins.

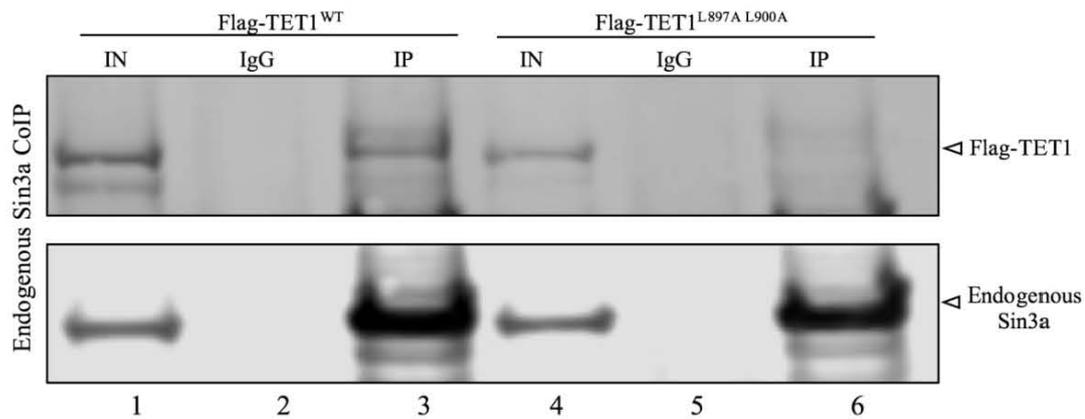


Figure 3.44: The L897A/L900A Mutations Prevent Full-Length Endogenous Sin3a from Interacting with Full-Length TET1 in ESC.

A co-IP analysis – Endogenous Sin3a was used to pull down full-length Flag-TET1 and TET1^{L897A/L900A} proteins in B8 derived ESC. In = input, IgG = negative control, IP = immunoprecipitation.

3.11. Structural Characterization of the Sin3a:TET1 Interaction

A structural model of the binding was needed to adequately illustrate the observed interaction between TET1 and Sin3a. The first attempt at generating this model entailed using X-Ray crystallography. Both the His-PAH1 and TET1-SID^{WT} peptides were used in this procedure.

3.11.1. Protein Crystallography Screens

The first crystallization trial made use of the following screens: Wizard, Hampton Research, Pact Premier, JCSG plus, MIDAS, Proplex, and Stura and Marcasol. A decision was made to forgo the gel filtration step which would have purified the complex, and in lieu of this the protein solutions were mixed in equimolar amounts and then concentrated. This was done with the knowledge that the proteins could not remain in this concentrated form for a long period of time – as they tended to precipitate out of solution. The UV280 measurement of the His-Sin3a 115-212 (PAH1) and His-GB1-TET1 878-911 mixture revealed that the protein had a concentration of 1.349mM. The corresponding Bradford measurement was curiously 591.48μM. Extrapolating from the Bradford method, as the complex was 25862.8g/mol, this could also be expressed as

15.3mg/mL. A Mosquito liquid handler was used with the 96 well crystallization plates to distribute the protein mixture between the different screens. Out of the 8 different crystallization plates, only the Molecular Dimension's ProPlex HT-96 Screen was partially successful in generating micro-crystals in two of the 96 wells. The two hits, wells F8 and F10, maintained 1M ammonium sulfate as a precipitant but contained neither an additive nor a salt. F8 contained the buffer 0.1 M sodium acetate at pH5, while F10 incorporated 0.1 M Tris at pH8.

It was possible that the micro-crystals were produced because the reaction took place too quickly due to an abundance of protein, as the concentration may have been too high. Repeating the experiment with half the protein concentration did not work, neither did reducing the temperature (4°C) to slow the reactions.

In order to better define the conditions that resulted in crystallization, the crystallization plate described in Table 3.1 was produced. This resulted in micro-crystals being formed in well B4.

		Sodium Acetate Concentration (M)					
		0.1M	0.2M	0.3M	0.4M	0.5M	0.6M
pH 3.29		A1	A2	A3	A4	A5	A6
pH 3.96		B1	B2	B3	B4	B5	B6
pH 4.89		C1	C2	C3	C4	C5	C6
pH 5.93		D1	D2	D3	D4	D5	D6
		All wells contained 1M ammonium sulfate					

Table 3.1: The His-PAH1 and TET1-SID^{WT} Solution Produced Crystals in well B4 at a pH 3.96, with 0.4M Sodium Acetate with 1M Ammonium Sulfate.

The micro-crystals produced in well B4 in the crystal screen depicted in Table 3.1, can be seen in Figure 3.45. Time constraints prevented the crystallography results from being adequately exploited. The His-PAH1 and TET1-SID^{WT} solution, when in the concentrations needed for crystallography, would precipitate over the course of hours.

This excluded the use of gel-filtration, which meant that it was possible the micro-crystals could result from either protein in isolation instead of the complex.

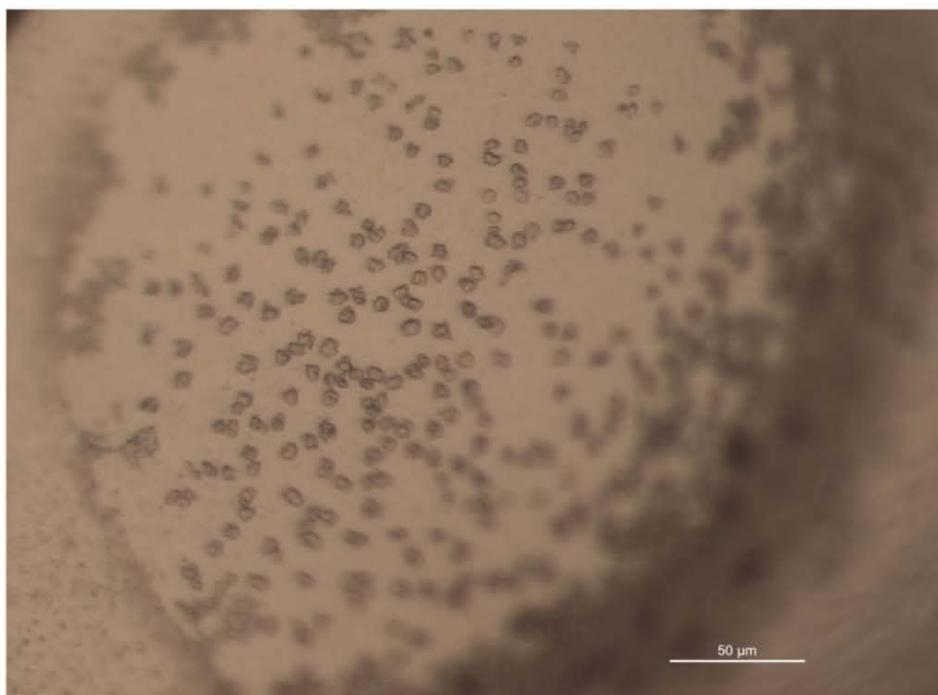


Figure 3.45: Micro-crystals Formed by the His-PAH1 and TET1-SID^{WT} Solution in Well B4. The crystals formed under the conditions described in Table 3.1.

3.11.2. An NMR Informed Model of the Sin3a:TET1 Interaction

The gel filtration and NMR mutation studies illustrated that the L900 and L897 residues have the greatest impact on Sin3a:TET1 complex formation. While significant, the I894A mutation did not impede binding to the same degree. This revealed the orientation of the TET1-SID in binding to Sin3a's PAH1. The T898E mutation had no effect on complex formation, which exposed the angle the SID took in its bound form, with respect to PAH1 (Figure 3.46). This additionally confirmed the orientation of the bound TET1-SID. If TET1-SID bound to PAH1 in the reverse configuration, T898E would protrude into the PAH1 domain and obstruct binding. The data from these mutational studies was combined with the Radakrishnan lab's NMR solution for the

PAH1:SAP25-SID complex to produce a structural model of the interaction between TET1-SID and PAH1 (Sahu et al., 2008).

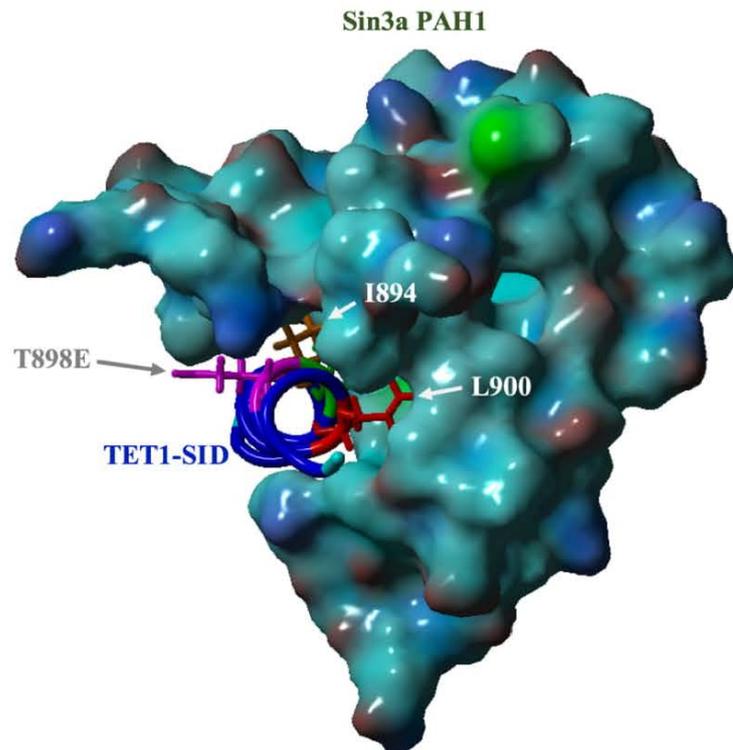


Figure 3.46: The T898E Mutation Revealed the Angle the TET1-SID Alpha Helix Binds to PAH1. The I894A, L897A and L900A Mutations Exposed the Orientation that the SID Interacts with Sin3a.

When viewing the binding within the PAH1 hydrophobic pocket, it is apparent that the residues interacting with this structure are I894, L897 and L900; with L897 occupying the central position within the pocket (Figure 3.47). A893 and A896 appear to assist in this binding as they are on the same face of the helix, but sequestered in their own hydrophobic pockets. The central location of L897 explains why a mutation of this residue has a more severe effect on Sin3a binding than the I894A or L900A mutants (Figure 3.28, Figure 3.30 and Figure 3.36).

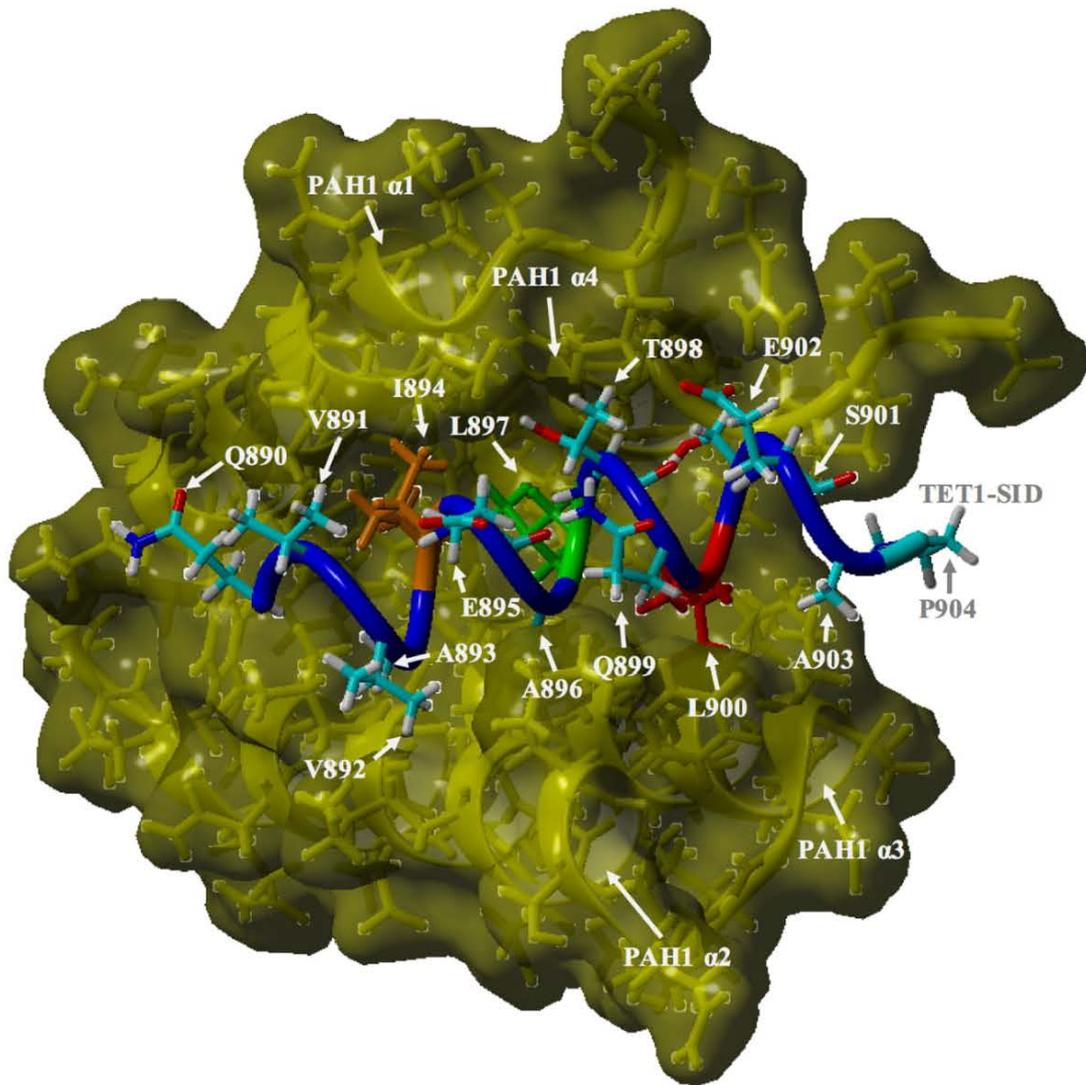


Figure 3.47: NMR Model of Sin3a's PAH1 Binding TET1 SID. TET1 L897A Occupies the Center of PAH1's Hydrophobic Pocket, it is Flanked by I894 and L900.

This model was informed by the gel filtration and NMR mutation studies, and extant PAH1 binding proteins. PAH1 is represented in yellow using a space filling model that overlays a ribbon and stick secondary structure model, whereas the TET1-SID secondary structure is represented by a tube model with stick models of the individual residues. Orange residue: I894, Green: L897, Red: L900. TET1 Sequence: 890 QVVAIEALTQLSEAP 903.

The S901 residue appears on the same hydrophobic face of the amphipathic helix, depicted in Figure 3.48 (a), as predicted by the helical wheel projection illustrated in Figure 3.4. The serine residue is polar and this part of the helix falls outside the hydrophobic PAH1 pocket (Figure 3.48 (b)), permitting the residue an aqueous environment. The Jpred prediction alluded to the SID terminating prior to this residue

(Figure 3.3). Jpred, however, also suggested that the SID terminated as abruptly as T898. The actual SID is better described as being between the Jpred and Multalin (Figure 3.1) predictions – 892-VAIEAL**TQL**-900.

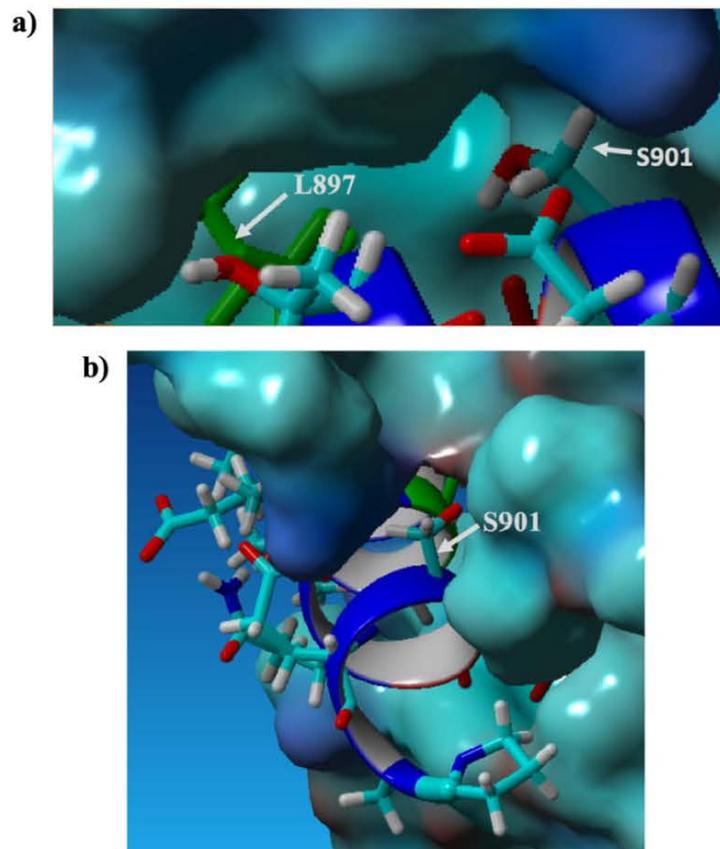
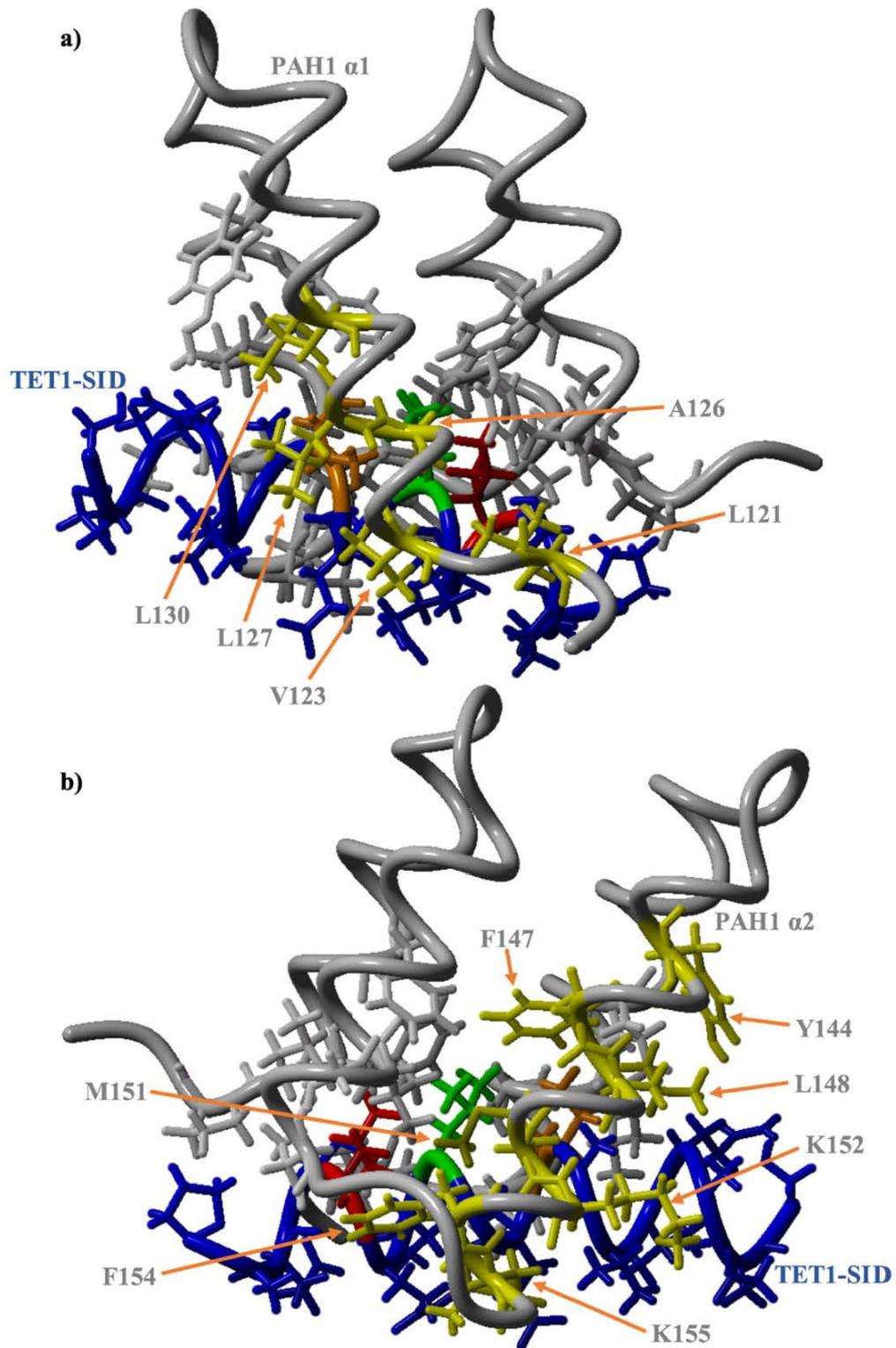


Figure 3.48: TET1 S901 is Exposed to the Aqueous Environment.

(a) S901 appears to be on the hydrophobic face of the alpha helix. The unusual presence of the polar serine 901 residue on the hydrophobic face of the SID, positioned next to L897. (b) The hydrophobic pocket ends prior to S901's position. This model was informed by the gel filtration and NMR mutation studies, and solutions of extant PAH1 binding proteins. PAH1 is represented in teal using a space filling model, whereas the TET1-SID is the central stick model containing the blue alpha helical region.

Sin3a's interaction with the TET1-SID occurs largely through the first two helices of PAH1, with the second helix maintaining the greatest number of contacts. There are 18 PAH1 residues that are responsible for binding the TET1-SID, these are indicated in Figure 3.49. The PAH1 residues form pockets into which the hydrophobic TET1 residues fit. TET1's L897, for example, is held within the hydrophobic socket formed

by the PAH1 residues L121, A126, F147, M151, F182 and F185. Moreover, some of the residues within PAH1 disproportionately contribute to TET1 binding as they are positioned between critical TET1-SID residues. One of these PAH1 residues – M151 interacts with both TET1 L897 and L900. Similarly, A126 interacts with both TET1 I894 and L897.



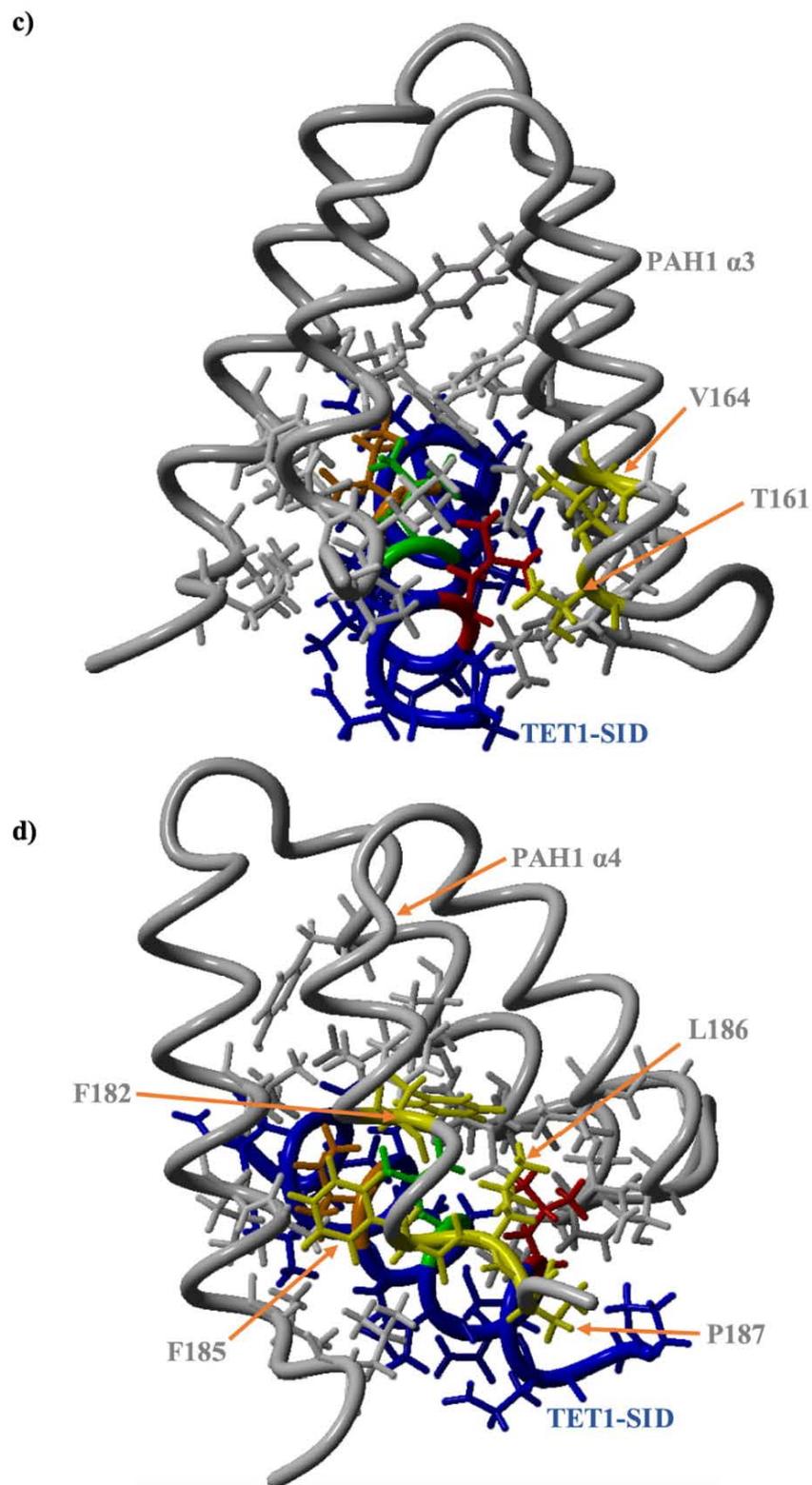


Figure 3.49: The Sin3a PAH1 (Grey) Residues That Bind to TET1's SID (Blue).

These are (a) α -Helix 1: L121, V123, A126, L127 and L130, (b) α -Helix 2: Y144, F147, L148, M151, L152, F154 and L155, (c) α -Helix 3: T161 and V164 and (d). α -Helix 4: F182, F185, L186 and P187.

3.12. Characterizing the Biological Significance of the TET1:Sin3a Interaction

In opposition to its enzymatic activity, which prevents aberrant DNA methylation and produces a transcriptionally permissive environment, TET1 is an overall negative regulator of gene expression. This activity is known to be independent of TET1's enzymatic activity. Indeed, this negative regulation occurs even in the absence of the enzyme's methylcytosine substrate (Williams et al., 2011). As the recruitment of Sin3a enables the transcriptional repression of the MAD1 and HBP1 proteins, TET1's interaction with Sin3a might similarly recruit HDAC1/2 activity (Swanson et al., 2004). Thus, the recruitment of the TET1-SID to a gene should be sufficient to repress its expression. Obstructing Sin3a binding, by introducing a L897A or L900A mutation into the TET1-SID, should abolish this regulation. This process is illustrated in Figure 3.50. A luciferase reporter assay, using Gal4 DNA binding domain (DBD) fusions of TET1-SID, tested the domain's effect on gene expression.

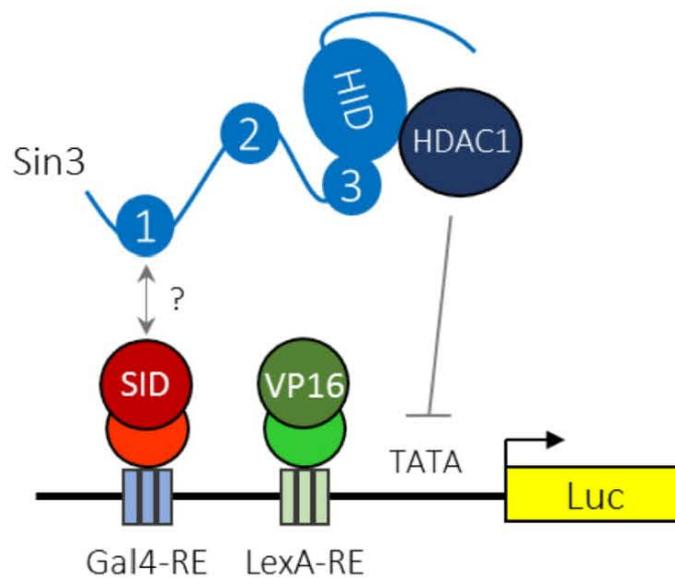


Figure 3.50: Gal4-SID Fusions Interact with Sin3a Domains and Indirectly Recruit HDACs to Repress Reporter Expression.

RE: response element, Luc: luciferase, TATA: gene promoter.

The assay revealed the interaction between the TET1-SID fusion and the Gal4 response element efficiently repressed luciferase expression (Figure 3.50). MAD1-SID

demonstrated same degree of luciferase repression. Mutation of the TET1-SID correlates closely with the NMR data (Figure 3.28, Figure 3.30 and Figure 3.36). I894A had a small effect, whereas both the L897A and L900A mutations eliminated TET1's ability to repress transcription. The reporter assay demonstrates that a large portion of TET1's negative regulation results from its interaction with Sin3a, almost certainly through the Sin3a's co-repressor components HDAC1/2 (Figure 3.51).

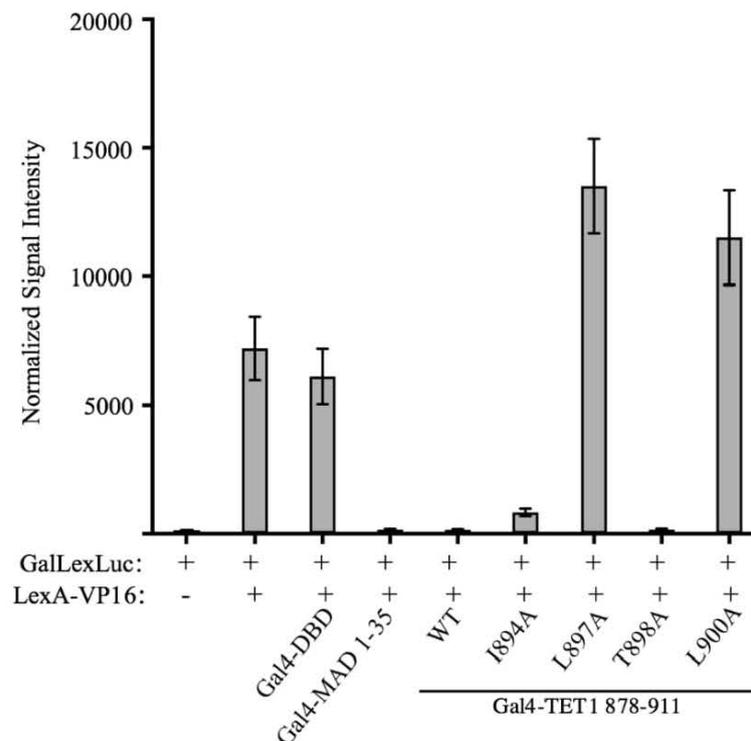


Figure 3.51: TET1's Gene Repression is Dependent on Sin3a Recruitment.

Disrupting the SID, inhibits this regulation. Gal-Lex-Luc reporter assay using Gal4 DNA binding domain protein fusions, normalized by β -Gal. Average value from three biological replicates displayed with the standard error of the mean.

3.12.1. qRT-PCR Reveals *TET1*^{WT} Rescue Increases IDAX and LIF Expression.

IDAX and Leukemia Inhibitory Factor (*LIF*) were assayed with qRT-PCR in B8 *TET1/2* DKO cells rescued with *TET1*^{WT}, *TET1*^{L897A/L900A} and *TET1*^{ΔCXXC}. The use of the mutant rescue cells in this assay interrogates, respectively, the influence of the SID and

the zinc finger domain on TET1's transcriptional regulation. Cursory data from a microarray (data not shown) suggested that *IDAX* and *LIF* were amongst the genes regulated by TET1. *IDAX* serves as an adaptor that brings TET2 to its targets. In addition, *IDAX* negatively regulates TET2 (Ko et al., 2013). Assessing TET1's influence on *IDAX* transcription would help to parse apart TET1's relationship with both *IDAX* and TET2. *TET1* expression was found to be necessary for LIF mediated STAT3 signalling and qRT-PCR would reveal if this link was more direct (Freudenberg et al., 2012). None of the qRT-PCR results proved to be of any meaningful statistical rigour.

The reintroduction of *TET1* upregulates *IDAX*. *IDAX*'s expression is increased in the *TET1*^{WT} cells as compared to the B8 *TET1/2* DKO cells (Figure 3.52). It appears that this upregulation is contingent on TET1's SID and CXXC domains, as *IDAX* expression in the *TET1*^{L897A/L900A} and *TET1*^{ΔCXXC} mutant cells was indistinguishable from the B8 DKO cells.

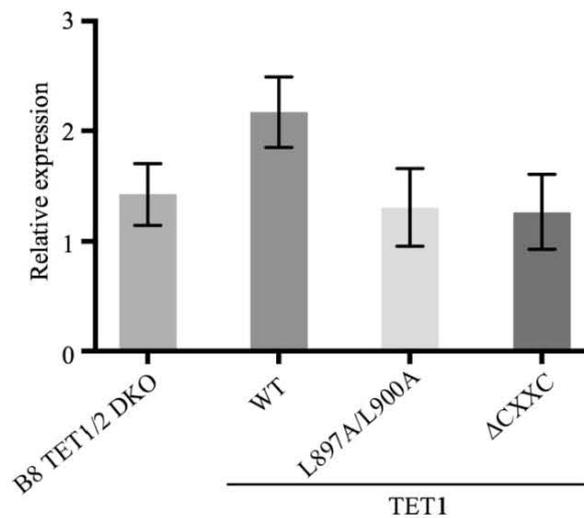


Figure 3.52: Reintroduction of *TET1*^{WT} Appears to Increase *IDAX* Expression.

qRT-PCR: *IDAX* fold change ($2^{-\Delta\Delta CT}$). Average of nine values resulting from three biological replicates, with three technical replicates for each, displayed with the standard error of the mean.

LIF transcription is higher in *TET1*^{WT} cells than in either B8 *TET1/2* DKO or *TET1*^{L897A/L900A} cells (Figure 3.53). This indicates that TET1 positively regulates LIF

expression, but implies this mechanism relies on the Sin3a:TET1 interaction. The precise difference in expression is difficult to assess because of the high variance in transcription between the three TET1^{WT} clones. Relatedly, it is not possible to determine the effect TET1's zinc finger domain has on this process. The variance in LIF transcription between the TET1^{ACXXC} and TET1^{WT} clones was comparably high.

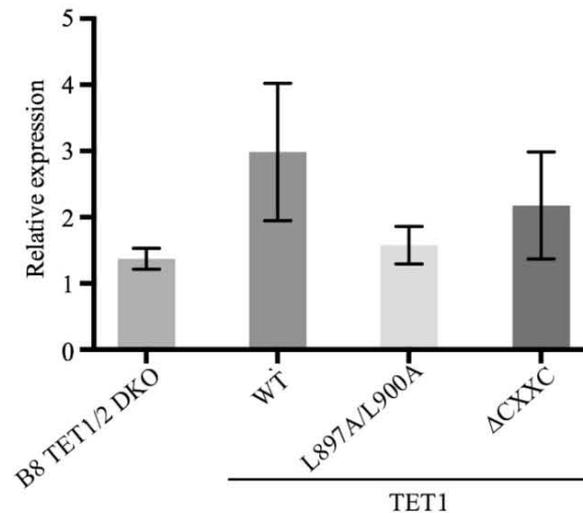


Figure 3.53: Reintroduction of TET1^{WT} Increases LIF Expression.

qRT-PCR: LIF fold change ($2^{-\Delta\Delta CT}$). Average of nine values resulting from three biological replicates, with three technical replicates for each, displayed with the standard error of the mean.

3.13. Sin3a:TET1 Complex Results

The first of the two epigenetic regulatory processes investigated in this thesis was the interaction between TET1 and Sin3a. Sin3a appeared to fundamentally alter the epigenetic function of TET1. Identifying the proteins' respective interacting domains was integral to the characterization of the Sin3a:TET1 complex. In contrast to Sin3a, TET1's major protein-protein interaction domains have yet to be exhaustively described. The existence of a Sin3a-binding TET1 paralogue enabled the identification of an evolutionarily conserved, α -helical, region present in both TET proteins (Figure 3.1-Figure 3.4 and Figure 3.16). This TET1 892-VAIEALTQLS-901 region was found to be necessary for Sin3a co-precipitation (Figure 3.5). The absence of this α -helical region from TET1 abolished the Sin3a interaction (Figure 3.6). This

evolutionarily conserved domain bound Sin3a even in the absence of the larger TET1 protein, as demonstrated by a TET1 878-911 GST-pulldown (Figure 3.7). Similar experiments identified PAH1 as the, exclusive, corresponding TET1 interaction domain in Sin3a (Figure 3.13 and Figure 3.14). NMR, supplemented by gel filtration and SDS-PAGE, illustrated the interaction between the two domains (Figure 3.20 and Figure 3.21). This not only confirmed the pulldown data, but affirmed that the interaction was equimolar (Figure 3.22). These techniques, also established the influence individual TET1-SID residues (T898, I894, L897 and L900) had on the interaction with Sin3a (Figure 3.26-Figure 3.38). As this data permitted the structural modelling of the TET1-SID:PAH1 complex (Figure 3.47), the mutant constructs also revealed the corresponding residues within PAH1 that directly engaged the SID (Figure 3.49). The proline mutations of Sin3a's L127 and L130 residues, found to interact with TET1 I894, disabled PAH1 and prevented full-length Sin3a from binding the TET1-SID. While less severe, the L897A/L900A mutations in full-length TET1 similarly abolished its interaction with full-length Sin3a. TET1-SID Gal4 DBD fusions were used to address the function of the Sin3a:TET1 interaction. The ability of the TET1-SID fusions to repress the transcription of a luciferase reporter gene correlated with their ability to bind to Sin3a (Figure 3.51). This closely followed the NMR and gel filtration data. Taken together, the data suggests that TET1's non-enzymatic, transcriptionally repressive, function largely derives from the hydrophobic interaction between TET1 892-901 and 18 residues within Sin3a's PAH1. This binding likely recruits the histone deacetylase activity of the Sin3a co-repressor complex.

Chapter 4. Results: HDAC1/2's Role in KDCr

The initial detection of histone crotonylation was not accompanied by a similar discovery of novel, crotonylation-specific, epigenetic proteins. Instead, the writers, readers and erasers of crotonylation were found to be previously described histone acetylation-related proteins (Sabari et al., 2015, Li et al., 2016, Bao et al., 2014). EP300/CBP, lysine crotonyltransferases, were previously thought to only function as HATs (Marmorstein, 2001). Similarly, the YEATS domain proteins, several of which preferentially bind the crotonylation mark, were initially thought to be readers of acetylation (Schulze et al., 2009). Unlike the groups of proteins discovered to be involved in other aspects of crotonylation regulation, only a single decrotonylase has been found. SIRT3, a histone deacetylase, is the only enzyme revealed to have *in vivo* decrotonylation activity (Bao et al., 2014). The nuclear distribution of class I HDACs (Figure 1.13) suggest that they may play a role in the removal of alternative histone acylations. This is supported by the finding that HDAC8 removes fatty-acylations (Aramsangtienchai et al., 2016). This activity was previously thought to be the exclusive function of, the SIRT3 paralogue, SIRT6. The ubiquity of the HDAC1/2 enzymes (Figure 1.14) indicate that they may play an important role in the lysine decrotonylation (KDCr) process.

4.1. Metabolic Effects on Crotonylation

EP300, an acetyltransferase known to confer the H3K18Ac mark, was the first crotonyltransferase to be discovered. In this new role, the crotonyltransferase also targeted H3K18 (Sabari et al., 2015). Antibodies against H3K18Ac and H3K18Cr permitted the investigation into, the crotonyl moiety precursor, crotonate's influence on acetylation and crotonylation in E14 ESCs. The addition of crotonate increased the crotonylation of H3K18 in a concentration correlated manner (Figure 4.1 - right). A similar level of acetylation could be observed on the H3K18 residue (Figure 4.1 - left).

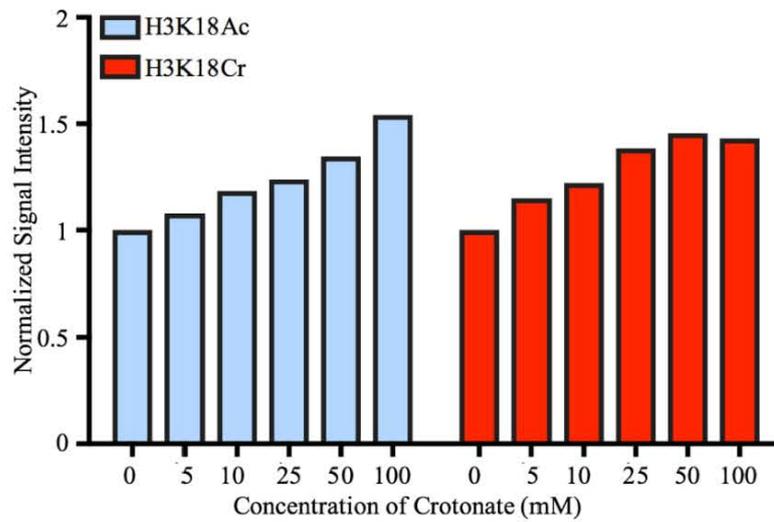


Figure 4.1: The Addition of Crotonate Increases Histone Crotonylation, But Also Histone Acetylation.

E14 ESC crotonate concentration course. Quantitative western blot for H3K18Ac (blue), H3K18Cr (red) both normalized to H3 (All ~17.5kDa).

4.2. Measuring Decrotonylase Activity

The deacylase assay worked on the premise that if the BOC-Lys(Ac/Cr)-AMC molecule could be deacylated, trypsin would be able to cleave the amide bond that attaches the AMC group to the substrate. This resulting fluorescence allowed the quantification of the deacylation (Figure 4.2).

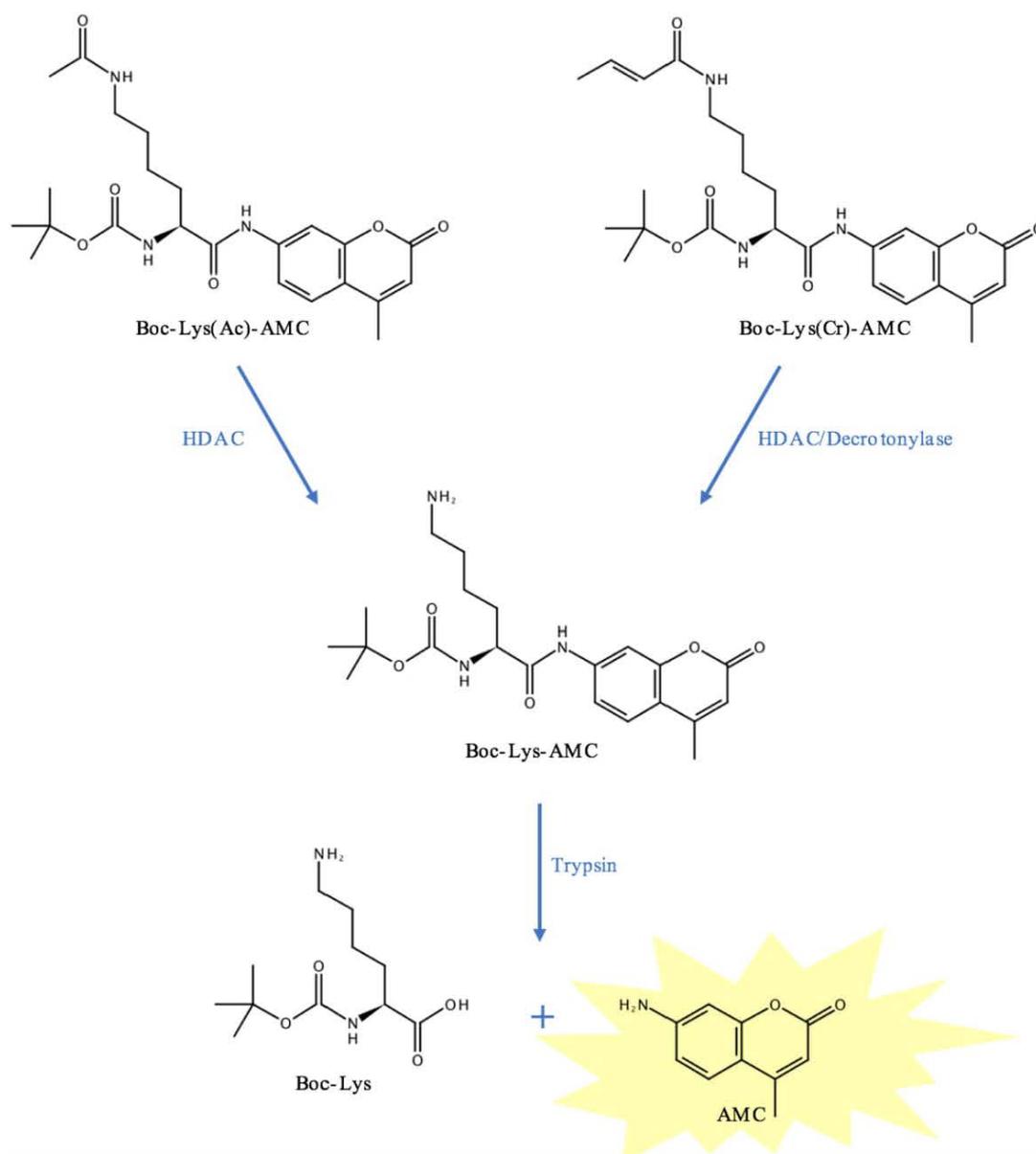


Figure 4.2: The BOC-Lys(Ac/Cr)-AMC Molecule Permits the Measurement of Deacetylation.

The acylation protects the fluorogenic AMC group from being cleaved by trypsin.

4.3. Measuring Cellular Decrotonylation Activity

The deacetylase/decrotonylase enzymes contained in HEK293 cells would naturally be present in HEK293 lysate. A portion of this lysate was added to the BOC-Lys(Ac)-AMC substrate in a concentration dependent manner and over a range that demonstrates significant deacetylase activity. This is illustrated in Figure 4.3 (a). A separate portion

of lysate was added to the Boc-Lys(Cr)-AMC substrate. While the Boc-Lys(Ac)-AMC experiment only required a two-hour incubation, this was insufficient to observe decrotonylation. Longer incubation demonstrated that Boc-Lys(Cr)-AMC substrate decrotonylation was also dependent on the concentration of lysate added. This is depicted in in Figure 4.3 (a) and (b). When compared to deacetylation, decrotonylase activity was less efficient in terms of both time and total substrate processed.

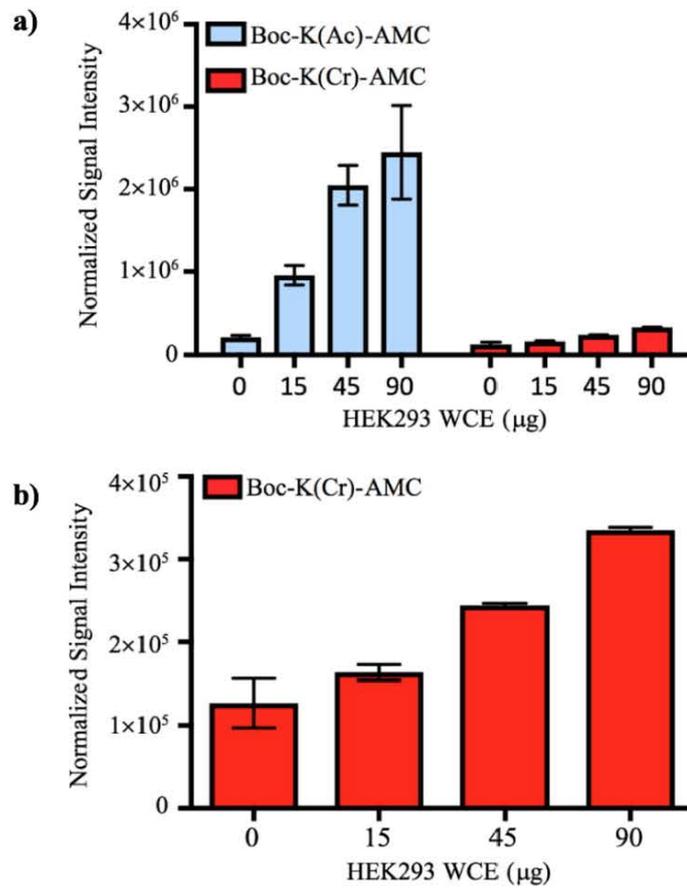


Figure 4.3: Mammalian Cell Lysate Contains a Lysine Decrotonylase.

(a) Boc-Lys(Ac)-AMC and Boc-Lys(Cr)-AMC incubated with different concentrations of HEK293 cell lysate for 2h and 18h, respectively. (b) Boc-Lys(Cr)-AMC results in isolation.

4.4. Are Classical HDACs Involved in Decrotonylation?

Panobinostat (LBH589) is a broad spectrum inhibitor of Zn^{2+} HDAC enzymes (HDACs 1-11) and facilitates the hyperacetylation of histones (Anne et al., 2013). The application of LBH, to E14 ESCs, resulted in the increase of H3K18Ac as well as H3K18Cr in a dose dependent manner (Figure 4.4). This suggested Zn^{2+} dependent HDACS, in addition to sirtuins, were responsible for histone decrotonylation

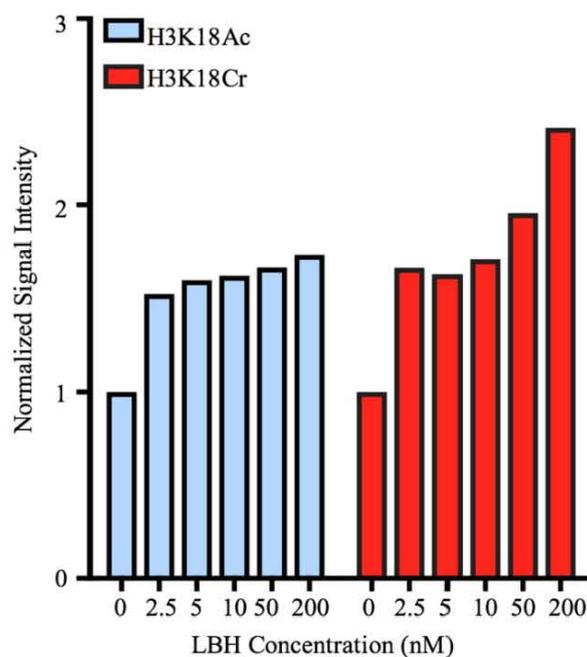


Figure 4.4: Classical HDACs Are Histone Decrotonylases.

E14 ESC LBH concentration course. Quantitative western blot for H3K18Ac, H3K18Cr both normalized to H3 (All ~17.5kDa).

In the absence of the HDAC activity, acetylation doubles within the first half hour (Figure 4.5, a and b). This takes longer than two hours for crotonylation (Figure 4.5, c and d). The absence of decrotonylase activity permits the rate of crotonylation to be observed. H3K18 crotonylation occurs at a much slower rate than acetylation.

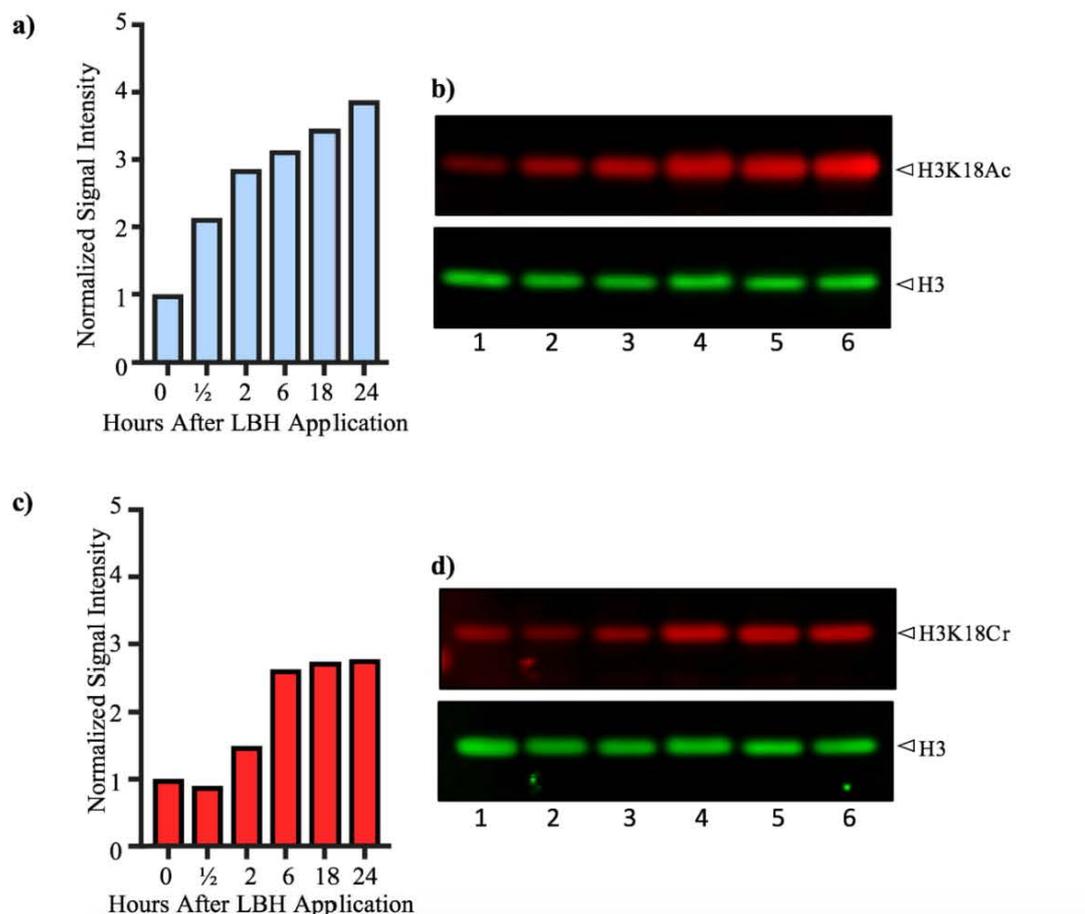


Figure 4.5: Histone Crotonylation is Slower than Histone Acetylation.

50nM LBH time course, E14 ESC. (a) Quantitative western blot for H3K18Ac, normalized by H3. (b) western blot, top panel: H3K18Ac, bottom panel: H3. (c) Quantitative western blot for H3K18Cr, normalized by H3. (d) western blot, top panel: H3K18Cr, bottom panel: H3.

4.5. Are HDAC1/2 Involved in Decrotonylation?

E14 *HDAC1/2* inducible DKO cells do not contain observable amounts of HDAC1/2 (~60kDa) protein three days after tamoxifen induction (Figure 4.6). These cells would enable three separate experiments. The influence of the two proteins could be observed, through their absence, in a decrotonylation assay. The H3K18Ac and H3K18Cr antibodies would reveal the specific effects HDAC1/2 have on H3K18, while a pan-crotonyl-lysine (pan-Kcr) antibody would illustrate the proteins' broader effects on histone crotonylation.

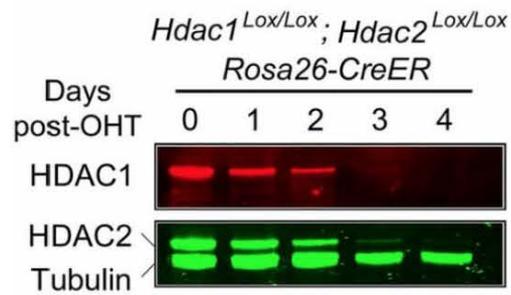


Figure 4.6: HDAC DKO Cells Contain Negligible HDAC1/2 3 Days After Induction (Jamaladdin et al., 2014). Western blot for HDAC1 (top panel), and HDAC2 and Tubulin (bottom panel) before tamoxifen induction (Day 0) and up to 4 days after tamoxifen induction.

The absence of HDAC1/2 results in decreased deacetylase and decrotonylase activity (Figure 4.7). In agreement with the result in Figure 4.3, despite a constant amount of substrate, deacetylation far exceeded decrotonylase activity. The data revealed that HDAC1/2 are responsible for the majority of the deacetylase and decrotonylase activity in ESCs.

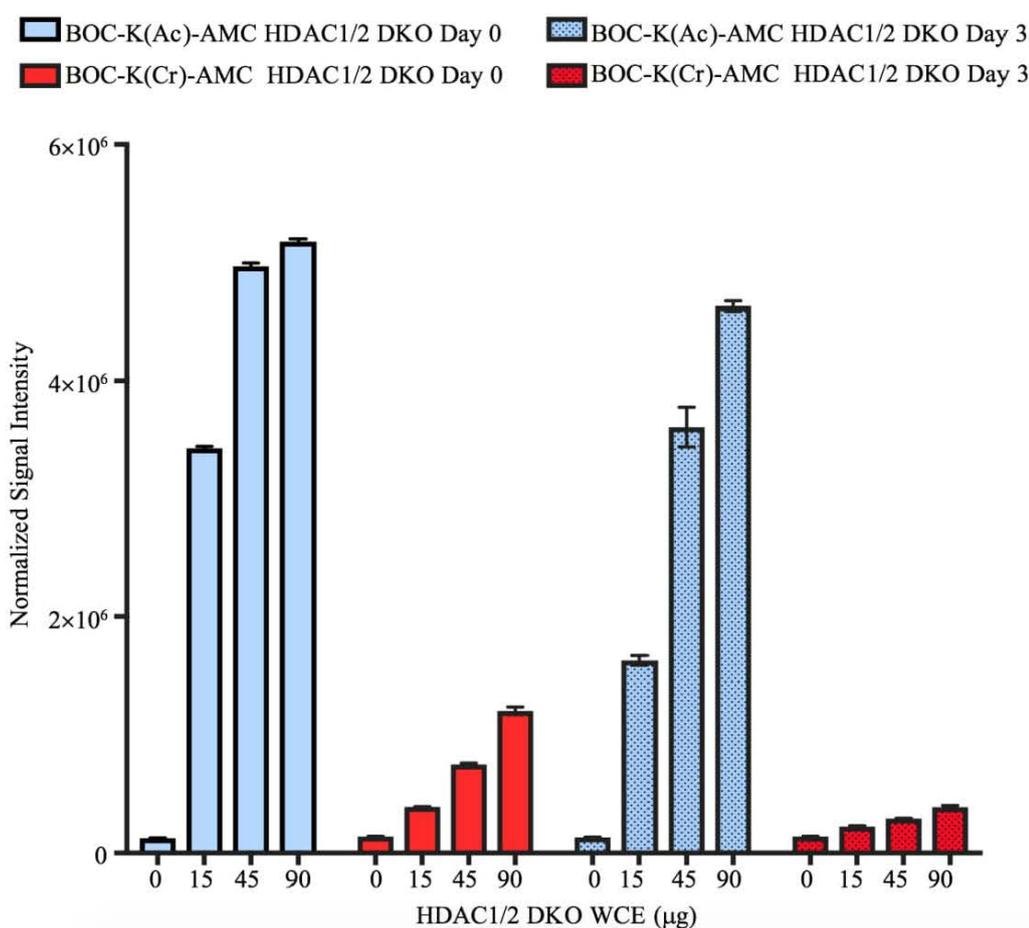


Figure 4.7: The Absence of HDAC1/2 Reduces Deacetylation and Decrotonylation Activity.

Deacylation assay using HDAC1/2 DKO cells. (Blue) BOC-Lys(Ac)-AMC with un-induced DKO lysate. (Blue patterned) BOC-Lys(Ac)-AMC with induced DKO lysate. (Red) BOC-Lys(Cr)-AMC with un-induced DKO lysate. (Red patterned) BOC-Lys(Cr)-AMC with induced DKO lysate.

When the HDAC1/2 DKO deacetylation assay is looked at in isolation, it is apparent that the increasing amounts of lysate begin to negate the DKO's effect (Figure 4.8). For 15mg/mL of lysate, the absence of HDAC1/2 results in a reduction of > 52% lysine deacetylation activity. For 90mg/mL the decrease is < 11%, suggesting that the assay was being saturated by lysine deacetylase activity. The other HDACs present in the lysate compensated for the loss of HDAC1/2 lysine deacetylation, albeit with reduced efficiency, as indicated by the 15mg/mL result.

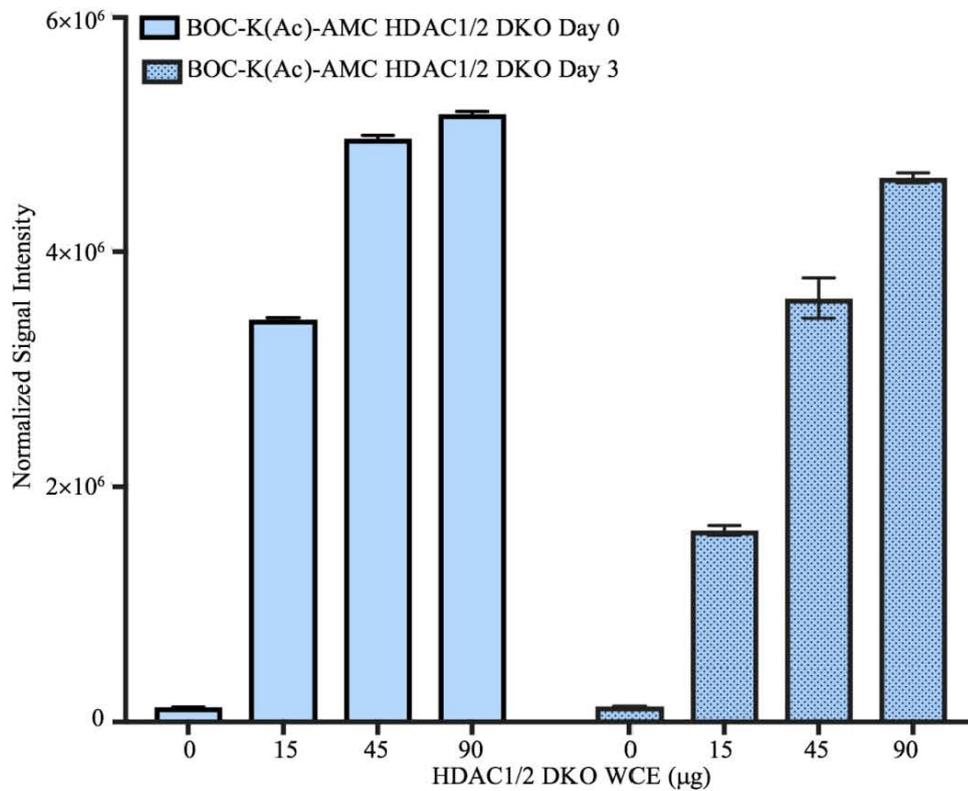


Figure 4.8: Alternative HDACs Can Deacetylate Lysine Residues Upon HDAC1/2 Loss, Albeit Less Efficiently.
 (Blue) BOC-Lys(Ac)-AMC with un-induced DKO lysate. (Blue patterned) BOC-Lys(Ac)-AMC with induced DKO lysate.

The naturally low decrotonylase activity present within the lysate ensured that saturation was not achieved even when investigating the wildtype cells (Figure 4.7). While HDAC1/2 are not the only decrotonylases present in the ESCs, their loss results in the forfeiture of the bulk of decrotonylation activity (Figure 4.9). This is true even at 90mg/mL, where the absence of HDAC1/2 results in a reduction of ~68% of lysine decrotonylation activity.

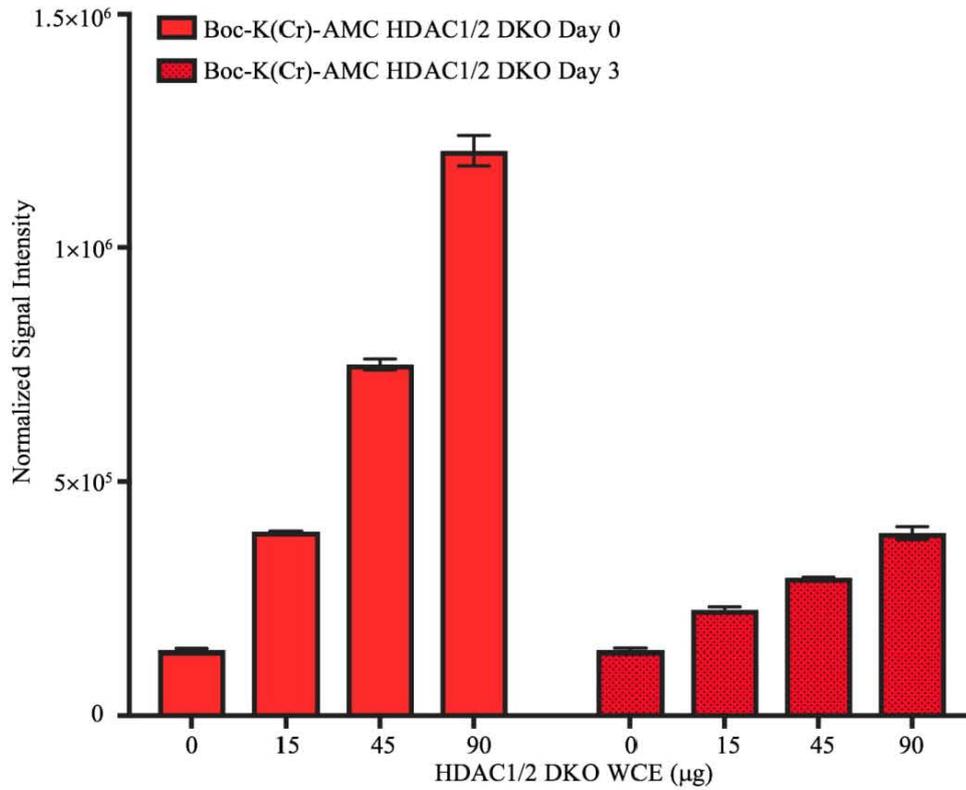


Figure 4.9: HDAC1/2 Are Responsible for the Bulk of Decrotonylation Activity in ESCs.

(Red) BOC-Lys(Cr)-AMC with un-induced DKO lysate. (Red patterned) BOC-Lys(Cr)-AMC with induced DKO lysate.

The use of H3K18Ac and H3K18Cr antibodies with the *HDAC1/2* DKO cells revealed that the loss of the two HDACs resulted in an increase in acetylation of just under 43% (Figure 4.10, a and b), whereas crotonylation increased by 80% (Figure 4.10, a and c). HDAC1/2 appear to regulate crotonylation, approximately, twice as tightly as they control H3K18 acetylation levels.

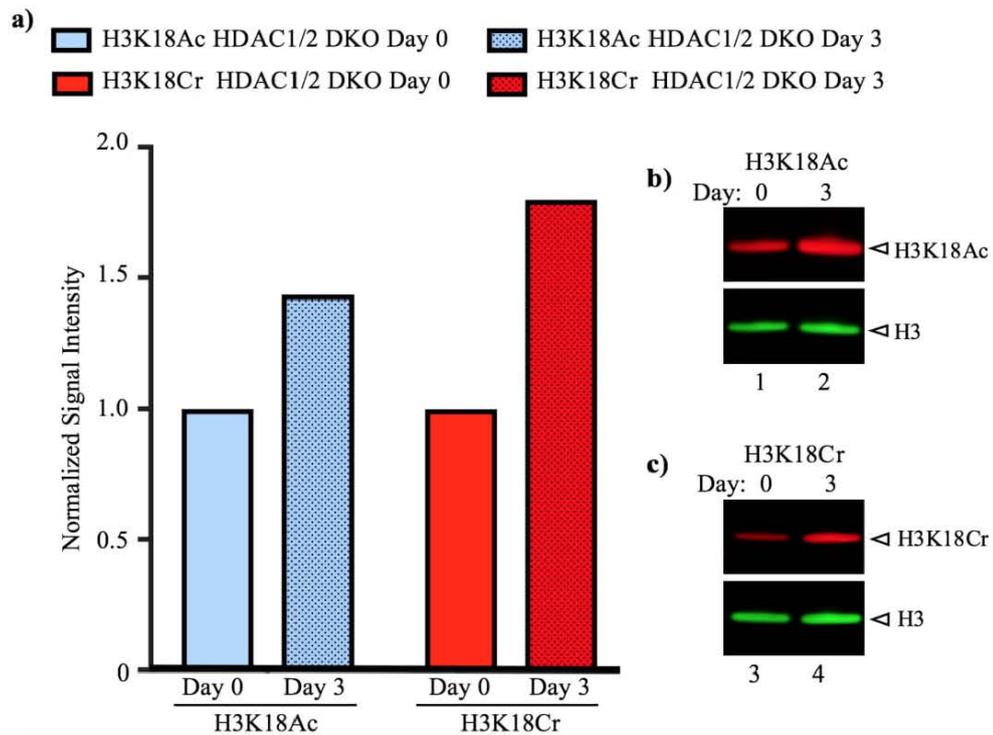


Figure 4.10: HDAC1/2 Control H3K18 Crotonylation Twice as Tightly as Acetylation.

HDAC1/2 DKO ESC. (a) Quantitative western blot for (blue) H3K18Ac and (red) H3K18Cr, normalized by H3. (b) western blot, top panel: H3K18Ac, bottom panel: H3. (c) western blot, top panel: H3K18Cr, bottom panel: H3.

Application of the pan-KCr antibody demonstrated that in the absence of HDAC1/2, H4's crotonylation increased more than 2.5-fold, while the cumulative crotonylation of H2a, H2b and H3 only increased 1.56-fold (Figure 4.11). The three H2/3 histones demonstrate 3-fold higher crotonylation than H4 in the presence of HDAC1/2. After HDAC1/2 loss, the disparity reduces to 1.85-fold. This means that the three crotonylation sites on H4 are closer to being at full occupancy, after HDAC1/2 loss, as compared to the cumulative 16 sites on H2a, H2b and H3 in ESCs (Tan et al., 2011). This is consistent with their being additional layers regulating histone crotonylation – there is a low level of decrotonylation activity even in the absence of HDAC1/2 as illustrated in Figure 4.9.

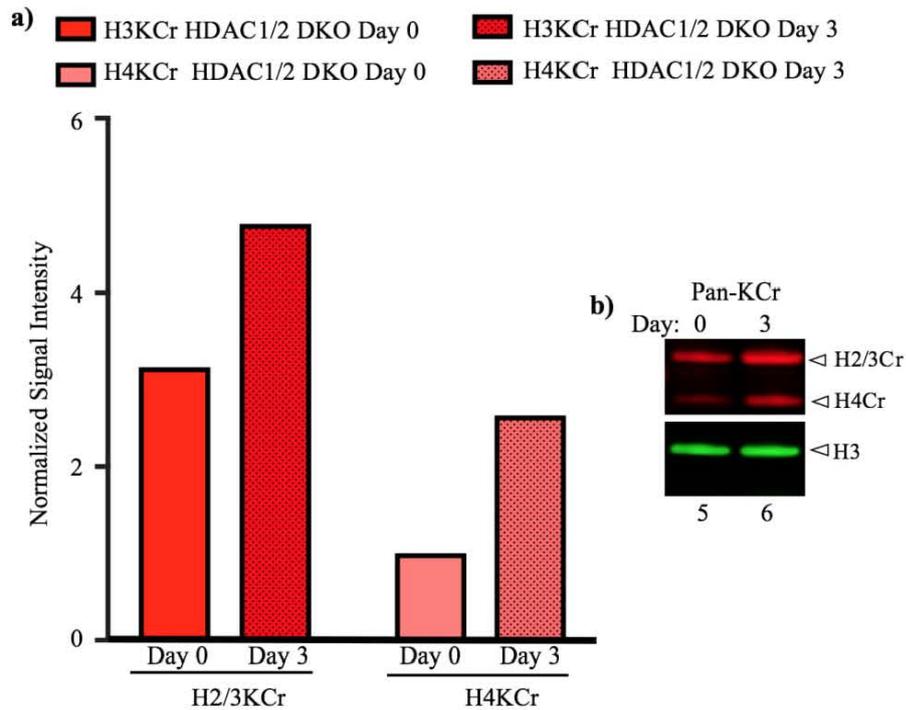


Figure 4.11: HDAC1/2 DKO Has the Greatest Effect on H4 Crotonylation.

(a) Quantitative western blot for H2/3KCr and H4KCr, normalized by H3. (b) western blot, top panel: pan-crotonyllysine, bottom panel: H3 (H2a, H2b and H3: ~17kDa. H4: ~13kDa).

An immunoprecipitation extracted the histone modifying enzymes LSD1 and HDAC2 from E14 ESC lysate. These proteins were subsequently used in a BOC-K(Cr)-AMC decrotonylase assay. Both HDAC2 and LSD1 were found to have decrotonylase activity, but HDAC2 had 51% more activity than LSD1 (Figure 4.12). It should be noted that LSD1, a histone demethylase, does not possess any innate decrotonylase activity. As a portion of the cell's LSD1 occurs within the CoREST complex, this results in the co-immunoprecipitation of CoREST constituents with LSD1. HDAC1/2 present within this complex mediate the observed decrotonylation of the BOC-K(Cr)-AMC substrate.

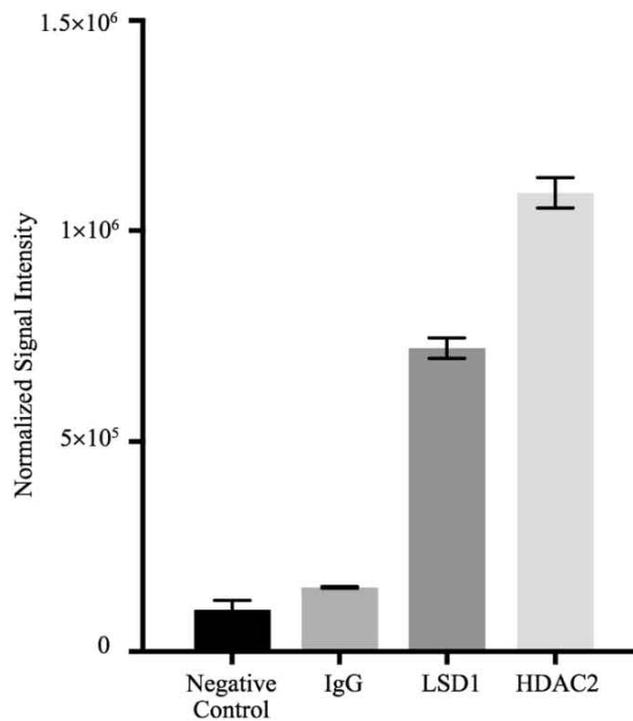


Figure 4.12: LSD1 Appears to Recruit Deacetylase Activity. HDAC2 Demonstrates Decrotonylase Activity. LSD1 and HDAC2 were immunoprecipitated, using 5 μ g antibody/replicate, from cell lysate and used in a BOC-K(Cr)-AMC decrotonylase assay.

4.6. HDAC1/2 Decrotonylation Results

The second epigenetic regulatory process investigated was the role that the classical histone deacetylases, HDAC1/2, had on lysine decrotonylation. The discovery of SIRT3 demonstrated the involvement of histone deacetylases in this process. HDAC1/2 were implicated in lysine decrotonylation by their distribution, activity and ubiquity. The panobinostat inhibition of HDAC1-11 led to H3K18 hypercrotonylation (Figure 4.4). This implicated the classical HDACs in histone decrotonylation. It also revealed that sirtuin activity, at cellular concentrations, was similarly insufficient to prevent either hypercrotonylation or hyperacetylation of H3K18 (Figure 4.5). E14 HDAC1/2 inducible DKO ESCs permitted a more direct observation of HDAC1/2's epigenetic influence. A deacetylase assay, revealed HDAC1/2 to be responsible for approximately two thirds of the ESC lysine decrotonylation activity as compared to half for deacetylation (Figure 4.9). HDAC1/2 were similarly found to regulate crotonylation more tightly in ESCs,

where their absence resulted in 43% more H3K18Ac, but 80% more H3K18Cr (Figure 4.10). The influence of HDAC1/2 is not even across histones in ESCs. Crotonylation of H2a, H2b and H3 was increased 56% on average, in the absence of HDAC1/2, whereas H4 crotonylation increased 85% (Figure 4.11). Direct evidence of HDAC2's lysine decrotonylation activity was displayed in a deacylase assay using immunoprecipitated HDAC2 (Figure 4.12). An immunoprecipitation of LSD1 resulted in two thirds the amount of lysine decrotonylation as the HDAC2 experiment. This indicated that HDAC1/2 decrotonylation activity could be recruited, by LSD1, through the CoREST complex.

Chapter 5. Discussion: Sin3a-TET1

5.1. TET1: The Great Epigenetic Loom

The true hero of Homer's *Odyssey* is not the eponymous character, but rather his wife – Penelope (Homer, 1999). In the wake of her husband's presumed death following the Trojan War, Penelope avoids ceding the crown of Ithaca to any of her 108 suitors, and in doing so escapes the social inequity of the age, through guile and little else. Penelope does this by having the suitors wait for her to finish weaving a burial shroud for her father in law – the hero Laertes, of Argonauts fame. She uses her great loom to diligently weave this woolen shroud during the day; but unbeknownst to her suitors, she unweaves the cloth every night under the light of a torch. Her great loom weaves and unweaves this cloth for three years, permitting her husband to return to Ithaca and likely saving her son from the suitors' machinations. TET1 appears to be part of the central nexus, or great epigenetic loom, through which epigenetic marks are woven and unwoven.

5.2. A Thousand Strings

TET1 stands at the point of confluence of more than six different epigenetic pathways, including: DNA demethylation, histone deacetylation, lysine acetylation, histone methylation, protein O-GlcNAcylation, protein phosphorylation, and covalent and non-covalent PARylation (Streubel et al., 2017, Zhong et al., 2017, Neri et al., 2013, Bauer et al., 2015, Ciccarone et al., 2015). TET1 appears to be tangentially related to several additional modifications, like SUMOylation through Sin3a's association with ING2; or histone acylation via the 5fC mark which appears to be a major determinant of EP300's distribution (Ythier et al., 2010, Song et al., 2013). The DNA demethylase is also directly tied several epigenetic mechanisms by metabolism. TET1's reliance on α -ketoglutarate, and therefore the citric acid cycle, means it can be inhibited by fumarate and succinate (Laukka et al., 2016). These are both precursors for many of the acylations imparted by EP300 onto histones. The consumption of α -ketoglutarate also metabolically links TET1 to many of the lysine demethylases (Culhane and Cole, 2007). TET1 similarly influences, and is influenced by, the availability of NAD⁺ through the PARylation pathway, metabolically connecting the DNA demethylase to class III

HDACs (Barkauskaite et al., 2013). The pattern that is most readily distilled from the cacophony of transcriptional regulation is the involvement of Sin3a, either directly or indirectly, in the extended TET1 epigenetic network. The mapping of the interaction between TET1 and Sin3a will, hopefully, contribute to the understanding of the mechanisms behind some of these PTM pathways.

5.3. Identification and Characterizing the TET1/Sin3a Interaction

Finding the Sin3a interaction domain (SID) of TET1 was facilitated by the discovery, by Deplus and colleagues (2013), that Sin3a interacted with TET1 and TET3, but not TET2. The search for a non-enzymatic region that was conserved in TET1 and TET3, but missing in their paralogue, revealed TET1 892-VAIEALTQLS-901 and TET3 259-IAIEALTQLS-268 (Figure 3.1). Bioinformatics tools suggested this was the region responsible for binding Sin3a (Figure 3.2-3.4 and Figure 3.16). Only sequences containing the putative SID could bind Sin3a (Figure 3.5), its deletion abrogated this ability (Figure 3.6). The SID largely in isolation proved to be capable of binding Sin3a (Figure 3.7), proving that the region was both necessary and sufficient for the interaction. While Sin3a interacted with the TET1-SID, the inactivation of PAH1 (L127P/L130P) prevented this interaction (Figure 3.13). Curiously, when both of the binding partners were present in truncated form, the signal from the interaction was atypically low (Figure 3.11) this was only partially remedied by using the slightly longer TET1 peptide (868-927) (Figure 3.12). A GST-pulldown using only PAH1, in isolation, demonstrated that this domain was sufficient for the interaction with the TET1 SID (Figure 3.14).

Gel filtration (Figure 3.20) confirmed TET1-SID and PAH1 formed a complex. The accompanying SDS-PAGE revealed that the two complex components simultaneously eluted from the gel filtration column, with the same relative concentration – indicating complex formation. Indeed, when PAH1 was labelled with N¹⁵ and its (TET1-SID) bound and unbound NMR spectra (Figure 3.21) were compared, it was apparent that all the PAH1 peaks shifted. This illustrated the PAH1 residues moving to accommodate the binding of TET1-SID. Unlike the histone deacetylase interaction domain (Figure 1.15),

Sin3a's canonical PAH domains only interact with a single partner at any given time. Examples of this include the SAP25:PAH1, MAD1:PAH2 and SAP30:PAH3 complexes (Swanson et al., 2004, Sahu et al., 2008, Xie et al., 2011). The use of an excess of PAH1, and subsequently TET1-SID, in gel filtration experiments, determined that they were equimolar binding partners (Figure 3.22). This was corroborated by NMR. A stoichiometric mixture of wildtype TET1-SID and PAH1 produced a distinctive NMR spectra that diverged from unbound PAH1 (Figure 3.21).

A 1:1 interaction did not preclude the presence of multiple interaction sites on either protein. Opi1, for example, is a protein that interacts with PAH1, but has two disparate sites that are individually capable of binding to this SIN3 domain (Wagner et al., 2001, Sahu et al., 2008). A GST-pulldown, bolstered by s2D (Figure 3.16), dismissed the possibility of there being multiple PAH1 binding sites in TET1. The demethylase was cut along domain boundaries and the only TET1 fragment (746-951) capable of interacting with PAH1 contained the, Multalin identified (Figure 3.1), SID (Figure 3.15). While it is generally accepted that SIDs only correspond to a single PAH domain, this was experimentally verified (Sahu et al., 2008). This was done using full-length Sin3a containing a, PAH1 inactivating, double L127P/L130P mutation. The lack of an operative PAH1 domain prevented, the otherwise functional, Sin3a from binding to TET1-SID (Figure 3.13). More conclusively, the disruption of the PAH1:TET1-SID interaction prevented the full-length proteins from interacting with each other. A, relatively innocuous, double alanine mutation within the SID of full-length TET1 blocked it from binding endogenous full-length Sin3a (Figure 3.43 and Figure 3.44).

5.4. Viewing Sin3a:TET1 from a Structural Perspective

A structural model of the Sin3a:TET1 complex was made using data derived from previous NMR solutions of PAH1 complexes (Nomura et al., 2005, Sahu et al., 2008). This model, illustrated in Figure 3.47, was informed by GST-pulldowns, gel filtration and NMR experiments. The TET1-SID L897A, L900A and I894A mutants suggested that the L897 residue was especially important for complex formation (Figure 3.26- Figure 3.38). This indicated that it occupied a central position within the PAH1

hydrophobic cleft. The orientation of the bound SID was informed by the T898E and I894A mutations. The perfect binding of the T898E mutation to PAH1 (Figure 3.33 and Figure 3.34) would not be possible if the SID bound Sin3a in the reverse orientation. In this configuration, the glutamic acid residue would occupy the same space as PAH1's second α -helix (Figure 3.49). The intermediate impact of I894A on binding (Figure 3.17, Figure 3.29 and Figure 3.30) suggested TET1's I894 occupies the opening, between the first and second α -helix, of PAH1's hydrophobic cleft (Figure 3.49). T898E mutation additionally confirmed the angle the TET1-SID alpha helix took within PAH1 (Figure 3.46). The model seamlessly accommodates the presence of the hydrophilic S901 residue on TET1-SID's hydrophobic face. The serine residue extends just beyond PAH1's hydrophobic pocket (Figure 3.48).

The Radhakrishnan group identified a motif (Figure 1.17) present in the SIDs of PAH1 binding proteins (Sahu et al., 2008). The motif was defined around the positioning of the bulky hydrophobic residues (Φ), within the SID, between short side chain (s) and non-proline (x) residues. With this information, it was proposed that the model could predict the orientation in which the SID bound PAH1. TET1-SID contains an eight-residue section, 894-IEALTQLS-903, that loosely conforms to the type I (Φ x Φ Φ sx Φ s) motif shared by NRSF and Pf1. TET1's SID only contains three bulky hydrophobic residues, and this results in TET1 A896 being designated as bulky by the Radhakrishnan motif. As the I894A mutation suggested, and the T898E TET1-SID mutation established, a type I orientation is not possible. The type I configuration would additionally result in S901 occupying a hydrophobic space. Instead, TET1's SID adopts the same orientation, in binding PAH1, as SAP25's SID. Unlike SAP25, TET1's SID does not conform to the proposed type II (s Φ x Φ s Φ x Φ) motif (Figure 5.1).

Type I Motif	Φ	x	Φ	Φ	s	x	Φ	s
TET1 894-901	I	E	A	L	T	Q	L	S
Type II Motif	s	Φ	x	s	Φ	Φ	x	Φ
SAP25 137-144	A	L	A	G	L	L	Q	M
TET1 896-903	A	L	T	Q	L	S	E	A
TET1 893-900	A	I	E	A	L	T	Q	L
TET1 890-897	Q	V	V	A	I	E	A	L

Figure 5.1: While TET1 Conforms to the Type I Motif, It Binds PAH1 in SAP25's Type II Orientation.

TET1 does not appear to conform to the Sahu et al. (2008) proposed model of SIN3 PAH1 binding. Hydrophobic amino acids involved in Sin3a binding are represented in bold. Amino acids that do not fit the model are indicated in red, the intensity of which corresponds to the degree of model mismatch. Residues that match the model are indicated in green.

The TET1-SID:PAH1 structural model identified 18 key residues within PAH1 that bind TET1's SID (Figure 3.49). These residues appear to place a significance on some of the shorter amino acids, like A893 and A896, within the TET1-SID. These alanine residues occur on the hydrophobic face of the SID and occupy a PAH1 hydrophobic pocket that would likely not tolerate a significantly longer residue. Moreover, K155 forms a lip at the bottom of the PAH1's second α -helix which holds the A896 residue, and the larger TET1-SID, in position. Each bulky hydrophobic residue within the TET1-SID (I894, L897 and L900) is held in a hydrophobic cage formed by 6-7 PAH1 residues. TET1 I894, for example, is wedged in a pocket formed by Sin3a's L121, V123, A126, L127, L130 F147 and L148. Several PAH1 amino acids contribute disproportionately to the Sin3a:TET1 interaction. This is due to their proximity to multiple bulky hydrophobic TET1-SID residues. While a T161W substitution would only expel TET1 L900 from the hydrophobic pocket, a M151A mutation in PAH1 would loosen the pocket holding both the TET1 L897 and L900 residues, weakening the hydrophobic interaction. Conversely, an A126W mutation would diminish the capacity of the hydrophobic pocket, displacing both TET1 I894 and L897 (Figure 3.49). The importance of the 18 PAH1 residues, that directly bind TET1, is conserved across Sin3a complexes. SAP25's SID is bound by the very same 18 PAH1 residues (Sahu et al., 2008). Indeed, the significance of these residues extends even beyond PAH1

complexes. PAH1 bears ~54% sequence similarity to PAH2. There are 17 PAH2 amino acids that interact directly with MAD1's SID and 13 of these residues have cognates in PAH1 that bind to both TET1 and SAP5 (Sahu et al., 2008).

5.5. Relative Strength of the Sin3a:TET1 Interaction

MAD1 and SAP25 appeared to have greater affinity for Sin3a than TET1. This was, obliquely, suggested during pulldowns that used a short Sin3a and TET1 sequences; the combination resulted in problematic signal intensity (Figure 3.11 and Figure 3.12). This issue was, however, not observed when using the SAP25-SID and MAD1-SID, despite their brevity. A GST-pulldown comparing SAP25-SID and TET1-SID's ability to bind full-length Sin3a, corroborated these findings (Figure 3.23). A competitive binding gel filtration experiment, using an equimolar mixture of SAP25-SID, TET1-SID and PAH1, permitted a more direct observation of this phenomenon. SAP25-SID demonstrated such a comparatively high affinity for PAH1, that it competitively inhibited the formation of the TET1-SID:PAH1 complex (Figure 3.25). The findings of the GST-pulldowns and the gel filtration experiment were supported in the literature. Both SAP25 (134 ± 13 nM) and MAD1 (400 ± 100 nM) are defined as tight Sin3a binders (van Ingen et al., 2004, Sahu et al., 2008). An additional, fourth, bulky hydrophobic group on the interacting face of the MAD1 and SAP25 SIDs permits a tighter interaction with the PAH domains' hydrophobic cleft (Figure 5.2).

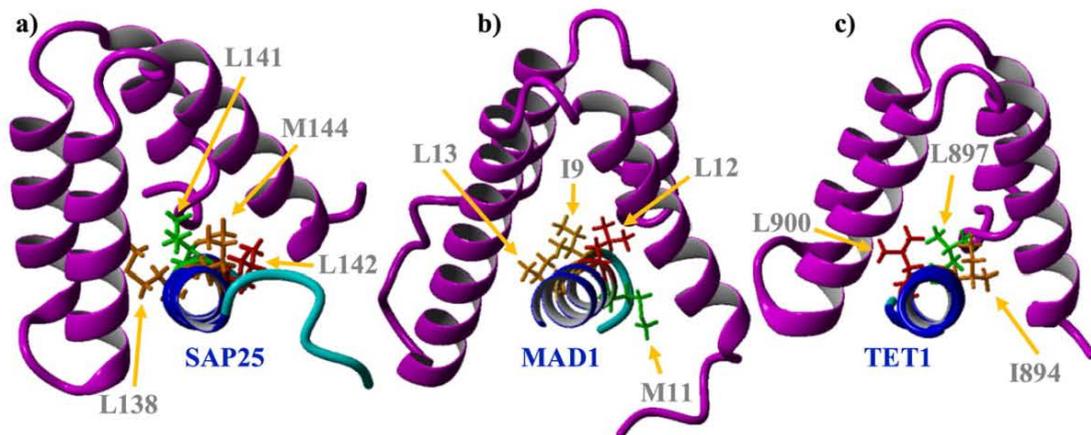


Figure 5.2: SAP25 and MAD1's SID Contains an Extra Hydrophobic Residue That Likely Permits Stronger PAH1 Binding.

(a) SAP25 (blue) binds Sin3a's PAH1 (magenta) with four hydrophobic residues: L138 (orange), L141 (green), L142 (red) and M144 (orange). (b) MAD1 (blue) binds Sin3a's PAH1 (magenta) with four hydrophobic residues: L13 (orange), I9 (orange), L12 (red) and M11 (green). (c) TET1 (blue) binds Sin3a's PAH1 (magenta) with only three hydrophobic residues: I894 (orange), L897 (green) and L900 (red).

MAD1's binding affinity to PAH2 is enhanced by a long-distance electrostatic attraction between MAD1's E23 and K165 of Sin3b's PAH2 (Figure 5.3, a). The presence of the MAD1 21-RRER-24 charge motif (++-+) increases MAD1's affinity for PAH2 by 3.5-fold (van Ingen et al., 2004). Sin3a has an analogous residue (K134) to Sin3b's K165 (Figure 1.18). Relative to MAD1, TET1 binds the PAH domain in the opposite orientation (Sahu et al., 2008). The electrostatic motif is present, but inverted in TET1: 882-KDRR-885 (+-++). It is also largely conserved in TET3 (213-KEKN-216). Jpred (Figure 3.3) and s2D (Figure 3.16) predict the region between TET1's D833 and the SID is unstructured; which might bring D833 closer to PAH1's K134 than Figure 5.3 (b) suggests. The relatively unstructured unbound PAH2 uses the electrostatic interaction to attract MAD1 (5-20) close enough to PAH2 for hydrophobic interactions to take over (Brubaker et al., 2000). While PAH1 maintains its structure in unbound form, it is plausible that a similar interaction might occur. The use of a TET1-SID D883A mutant might provide insight on this topic.

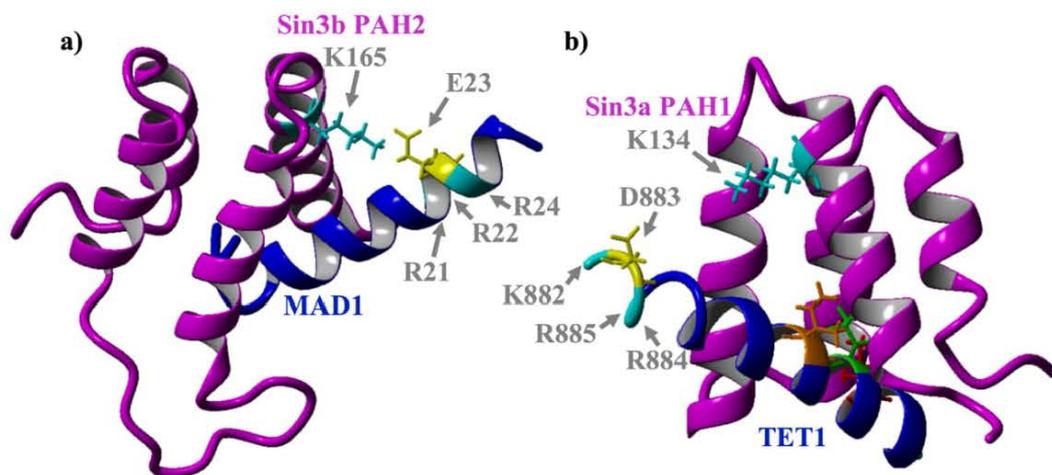


Figure 5.3: *TET1* May Have a Long Distance Electrostatic Interaction with *Sin3a*'s PAH1, like *MAD1* with *Sin3b*'s PAH2.

(a) *Sin3b*'s PAH2 (magenta) has a positively charged K165 residue (cyan) that interacts with *MAD1*'s (blue) negatively charged E23 (yellow) (van Ingen et al., 2004). (b) *TET1*'s (blue) D883 (yellow) residue likely interacts with K134 (cyan) in *Sin3a*'s PAH1 (magenta).

5.6. A Dissenting View

During the course of this project, Zhong and colleagues (2017) determined, using a co-IP; that the terminal, DSBH containing region of *TET1*, was responsible for the interaction with *Sin3a* (Figure 5.4).

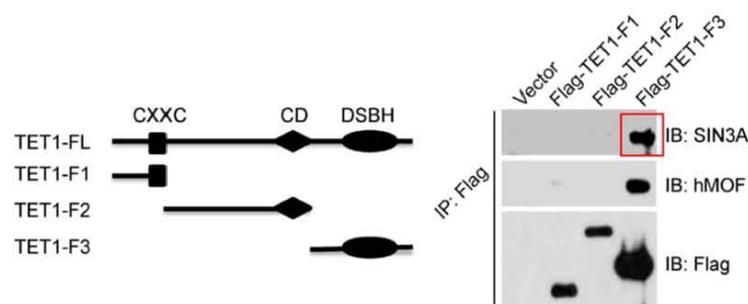


Figure 5.4: *A Co-IP Revealing the C-terminal End of TET1 is Responsible for Sin3a Binding.*

The band in the red box reveals that the terminal fragment of *TET1* interacts with *Sin3a*. *TET1*-F1: 1-654, *TET1*-F2: 620-1590, *TET1*-F3: 1418-2138 (Zhong et al., 2017).

They demonstrated the interaction a second time (Figure 5.5), in a GST-pulldown, using purified GST-TET1 DSBH (TET1-F3: 1413-2138) as well as purified His-Sin3a proteins.

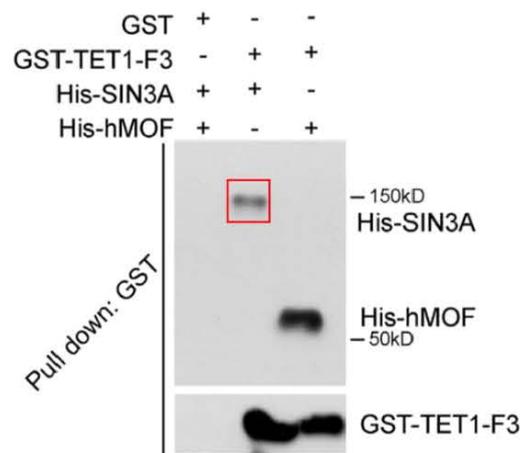


Figure 5.5: GST-Pulldown Shows TET1-F3:1418-2138 Interacts with His-Sin3a.

GST-Pulldown using purified proteins produced in *Bl21* cells. The band in the red box indicates an interaction between the C-terminal fragment of TET1 and Sin3a (Zhong et al., 2017).

Zhong and colleagues' (2017) work make two separate refutations of the data presented in this thesis. The first was that a region within TET1 1418-2138 is responsible for Sin3a binding (Figure 5.4 and Figure 5.5). The second was that the SID described in this thesis, TET1 892-903, cannot interact with Sin3a. This latter point was made by the co-IP that depicts TET1-F2 (TET1 620-1590) as being unable to bind Sin3a (Figure 5.4).

In addressing the first issue, it should be noted that the Sun lab's proposed SID, TET1-F3 (TET1 1418-2138), shares 65% sequence identity with TET2. TET2 does not interact with Sin3a (Deplus et al., 2013). The possibility of there being multiple TET1 SIDs was ruled out by several experiments. TET1 exclusively interacts with Sin3a's PAH1, and there is only a single corresponding SID throughout the entirety of TET1, as detailed in chapter 3.6. One GST-pulldown conducted to establish this, covers the Sun lab's proposed SID. Figure 3.15, demonstrates that neither TET1 1256-1799 nor 1800-2136 interact with PAH1. More pertinently, a double alanine mutation to the, Multalin identified, SID prevented the full-length TET1 from interacting with full-length Sin3a (Figure 3.43 and Figure 3.44). This would not be the case if TET1 contained an

additional SID that fell outside the four-amino acid region disrupted by the double mutation. The finding in this thesis that the TET1 892-903 region was responsible for the interaction was first suggested by bioinformatics tools and later demonstrated by co-IPs and GST-pulldowns. This was corroborated by gel-filtration, NMR and even functionally via a reporter assay. It is therefore difficult to reconcile this with the Sun lab's second assertion, without further information; though as the comprehensive theory of active DNA methylation will eventually and undoubtedly reveal – there are elements of truth in many of the numerous, divergent, currently proposed mechanisms. Each of the theorized processes likely illustrates a portion of a much larger, intricately regulated, epigenetic picture. As with the TET1-Sin3a relationship, it is always possible for multiple apparently opposing theories to eventually harmonize.

5.7. Biological Implications of the Sin3a:TET1 Complex

Despite being a DNA demethylase and its involvement in multiple epigenetic processes, many of which occur through the Sin3a/Fam60 complex, TET1's net epigenetic effect is transcriptionally repressive (Streubel et al., 2017, Williams et al., 2011). The transcriptional significance of other repressor proteins, like MAD1 and HBP1, has been linked to their ability to recruit HDAC activity through binding Sin3a (Swanson et al., 2004). As TET1 is also known to recruit Sin3a, it seems likely that TET1's net transcriptional repression derives from the same mechanism (Williams et al., 2011). A luciferase reporter assay was used to examine the effect the TET1-SID had on transcription. The mutations used in the NMR and gel filtration experiments were incorporated into the reporter assay to determine their, individual, functional impact. While TET1-SID did not appear to bind to Sin3a as tightly as MAD1-SID (discussed in 5.4), this difference did not translate in terms of biological function. The presence of either protein resulted in the near complete inhibition of reporter expression (Figure 3.51). The T898A mutation, in agreement with NMR and gel filtration data, demonstrated identical transcriptional repression. NMR and gel filtration revealed that the individual L897A mutation nearly abolished TET1-SID's interaction with PAH1. This same effect was displayed in the reporter assay, where the mutant TET1-SID was unable to repress luciferase expression. While NMR demonstrated that L900A had a

slightly less severe impact on PAH1 binding, as compared to L897A, the reporter assay did not make this nominal distinction. The NMR spectra of the I894A complex only moderately deviated from the wildtype complex, the mutation similarly had a mild but observable effect in the reporter assay. The reporter assay confirms that the TET1-SID negatively regulates gene expression. Additionally, the degree of the transcriptional repression is directly proportional to the TET1-SID's ability to bind PAH1. This suggests that TET1 negatively regulates gene expression through HDAC1/2 activity via Sin3a, as illustrated in Figure 3.50. The similar levels of transcriptional repression displayed by TET1-SID and MAD1-SID is partially explained by their shared reliance on Sin3a for this activity.

The qRT-PCR experiment indicated that *TET1*^{WT} cells expressed higher levels of both *IDAX* and *LIF* than their parental B8 *TET1/2* DKO cells (Figure 3.52 and Figure 3.53). TET1's demethylation of the *IDAX* and *LIF* loci may contribute to their increased expression. This is consistent with the observation that *IDAX* was not upregulated in *TET1*^{ΔCXXC} cells, as TET1's enzymatic activity is recruited to DNA via the protein's zinc finger domain. The expression of *IDAX* and *LIF* was similarly not increased in *TET1*^{L897A/L900A} cells, which instead maintained a similar level of transcription to the B8 *TET1/2* DKO cells. This signified that the, TET1 mediated, upregulation was dependent on the DNA demethylase's ability to interact with Sin3a. This is curious in the case of *LIF*, as the cytokine is known to promote pluripotency via STAT3 signaling (Raz et al., 1999). Sin3a antagonizes this pathway through the deacetylation of STAT3, inhibiting the transcription factor's nuclear recruitment (Icardi et al., 2012). The result for *IDAX* is equally difficult to interpret, as there is a paucity of information on *IDAX*'s relationship with Sin3a.

The PiggyBAC integration of the *TET1* genes, into the B8 *TET1/2* DKO cells, successfully characterized the effect of the L897A/L900A TET1 mutations on the binding of full-length TET1 and Sin3a. The technique was less suitable for qRT-PCR analysis. PiggyBac's method of random chromosomal integration likely contributed to the excessive variance in gene expression observed between clones (Figure 2.3). This

was especially the case regarding the TET1^{WT} and TET1^{ΔCXXC} rescue cells, as seen in the qRT-PCR analysis of *LIF*.

5.7.1. TET1^{ΔCXXC}'s Biological Consequences

There was excessive variation between clones, and not enough time, to adequately describe the TET1^{ΔCXXC} rescues. Good et al. (2017) recently discovered and characterized a naturally occurring isoform of TET1 that lacks the CXXC domain. The protein results from an alternative transcription start site in the gene's second exon. This variant (TET1-Alt) is not expressed in ESCs, just as the full-length protein is largely absent in adult tissues. TET1-Alt has a different binding profile to the full-length protein. The TET1 isoform is unable to prevent the aberrant methylation of CpG islands, a phenotype associated with cancer formation. TET1 was originally discovered as an MLL fusion lacking a CXXC domain (Ono et al., 2002). This likely explains the development of the AML phenotype despite a global increase in 5hmC (Huang et al., 2013). TET2 similarly lacks a CXXC domain, though this is partially counterbalanced by TET2's interactions with zinc finger adaptor proteins like IDAX or even the DNMTs (Ko et al., 2013, Zhang et al., 2017). Multiple mutations within *TET2* are known to contribute to AML (Gaidzik et al., 2012). Future experimentation might establish if a mutation induced disruption of the TET2:CXXC-protein interaction drives the disease phenotype.

5.8. Sin3a:TET1 Complex Conclusion

Sin3 evolutionarily predates the TET proteins, whose ancestry dates back to the early jawed vertebrates (Iyer et al., 2009, Vidal et al., 1990). The evolutionary conservation of the TET1-SID extends similarly far into the past. The domain can be found in reptiles, birds and even fish with perfect sequence similarity. This ancient protein-protein interaction occurs through the hydrophobic face of the TET protein's SID, an amphipathic α -helix, which in the case of TET1 occurs at 892-VAIEALTQLS-901. The bulky hydrophobic residues within this motif, I894, L897 and L900, bind the Sin3a PAH1 domain's hydrophobic cleft. The smaller residues within the SID, especially

A896, also contribute to the binding; though, their contribution ensures that the interaction between the domains is close-fitting. Comparable to SAP25's interaction with Sin3a, there are 18 PAH1 residues (Figure 3.49) that interact directly with the TET1-SID (Sahu et al., 2008). The TET1-SID also adopts the same (C- to N-terminus) orientation as SAP25 in binding PAH1; this is reversed in MAD1's interaction with PAH2. Similar to SAP25:PAH1 and MAD1:PAH2 complexes, Sin3a:TET1 maintains 1:1 stoichiometry. As a DNA demethylase, TET1 prevents aberrant DNA methylation, however, it is through the interaction with Sin3a that TET1's epigenetic influence transcends its enzymatic activity. Like MAD1-MAX, TET1's transcriptionally repressive effects derive from its association with Sin3a, which permits the recruitment of the Sin3a co-repressor complex's histone deacetylase activity (Laherty et al., 1997, Ayer et al., 1995). TET3's SID, 259-IAIEALTQLS-268, likely permits it to bind to Sin3a in the same manner as TET1. The combination of TET3's SID and zinc finger domain would similarly permit the recruitment of HDAC1/2 activity to a subset of the enzyme's gene targets, enabling their transcriptional repression. TET3, like the other TET proteins, is known to interact with OGT (Ito et al., 2014). It would be tempting to speculate on the existence of a TET3:Sin3a:OGT megadalton complex; though, in keeping with TET3's expression pattern, such a complex would only be possible relatively early in development (Gu et al., 2011).

Chapter 6. Discussion: HDAC1/2's Role in KDCr

6.1. Crotonylation

The second focus of this thesis was to establish the role of, the Sin3a co-repressor complex components, HDAC1/2 in lysine decrotonylation. SIRT3 was found to have *in vivo* decrotonylase activity. The sirtuins' reliance on NAD⁺ ties their activity closely to metabolism, allowing epigenetic processes to respond to the environment. Low nutrient conditions, that lead to a higher NAD⁺:NADH ratio, promote sirtuin activity (Yu and Auwerx, 2009). There are circumstances where crotonylation regulation is unrelated to nutritional environment, as in the example of acute kidney injury (Sabari et al., 2017). Classical HDACs would provide a more appropriate regulatory response in these situations. While the HDAC1/2 proteins have intrinsic lysine deacetylase activity, their epigenetic influence is augmented by their involvement in key epigenetic complexes. This is especially true of the CoREST complex which links histone deacetylation to LSD1's histone demethylation (Shi et al., 2005). HDAC1/2's ubiquity, nuclear localization and influence on the related acetylation pathway made them ideal lysine decrotonylase candidates. A novel lysine decrotonylation assay supplemented by inducible *HDAC1/2* DKO ESCs permit lysine crotonylation to be observed in the absence of HDAC1/2 activity, enabling the specific influence of these two proteins to be identified.

6.2. Crotonylation and Metabolism

EP300 could use either Crotonyl-CoA or Acetyl-CoA to acylate H3K18 (Sabari et al., 2015, Mizzen et al., 1999). Sabari et al. (2015) demonstrated that the addition of modest concentrations of (10mM) sodium crotonate could significantly increase crotonylation levels. The experiment detailed in Figure 4.1 reveals that the addition of crotonate does elevate H3K18 crotonylation levels. The rise in crotonylation, however, was very slight and was accompanied by an equally significant increase in acetylation. This was true even of concentrations that were an order of magnitude higher than the one used by Sabri and colleagues (2015). The marginal increase in crotonylation and accompanying elevation in acetylation can be explained through metabolism. Acetyl-CoA can be converted into and from crotonyl-CoA (Eggers and Steinbuechel, 2013, Sabari et al.,

2015). This explanation is countered by Wei et al. (2017), who corroborated Sabri and colleagues' (2015) data; disputing the results displayed in Figure 4.1. It is possible that H3K18 crotonylation in ESCs is nearly at full occupancy to begin with, which means that any added crotonate would not yield a dramatic change. A similar phenomenon was observed when using the pan-crotonylation antibody (Figure 4.11), where H4 appeared to be preferentially crotonylated.

6.3. Corroboration of HDAC1/2's Role in Decrotonylation

It was found that mammalian cell lysate could decrotonylate lysine residues in a manner that was directly proportional to the concentration of the lysate (Figure 4.3). This verified the findings of Bao and colleagues (2014) who discovered *in vivo* SIRT3 decrotonylase activity. The experiment in Figure 4.3 simultaneously validated the decrotonylase assay. Application of the pan-HDAC inhibitor (LBH) to cells increased H3K18 acetylation, but also increased lysine crotonylation in a dose dependent manner (Figure 4.4). As SIRT3 is not zinc dependent, and therefore immune to LBH activity, this result suggests that classical HDACs play a more important role in this process (Bao et al., 2014, Marks and Xu, 2009). Xu and colleagues (2017), confirmed these findings. They observed that SAHA and LBH, both inhibitors of classical HDACs, increase crotonylation levels. Xu et al., however, theorize that HDAC1 and HDAC3 are responsible for this activity. HDAC3 is a nuclear HDAC and is the second most highly expressed deacetylase in ESCs, which makes its decrotonylase activity plausible. HDAC1 shares ~85% sequence identity with HDAC2, but only ~60% with HDAC3; it would therefore be more credible if HDAC2 maintained any conserved HDAC1 decrotonylase activity. The comparative ubiquity of HDAC2 to HDAC3 in ESCs (Figure 1.14) makes an equally compelling argument, as does the demonstrable lysine decrotonylase activity of immunoprecipitated HDAC2 (Figure 4.12). The use of *HDAC1/2* inducible knockout cells (Figure 4.6), revealed the two enzymes to be responsible for the greater part of the cells' total decrotonylase activity (Figure 4.7 and Figure 4.9). This explains the results of Figure 4.5, where the absence of the classical HDACs resulted in the accumulation of H3K18 crotonylation. Extrapolating from the H3K18 and the BOC-K(Cr)-AMC assays (Figure 4.5 and Figure 4.7), it appears that

crotonylation and decrotonylation take longer than the equivalent acetylation processes. This supports the observation, by Montellier et al. (2012), that crotonylation may provide a means of acylating histones in an environment hostile to lysine acetylation. HDAC1/2 appear to control the process of decrotonylation far more tightly than they do deacetylation (Figure 4.10) and the enzymes appear to preferentially target lysine residues on H4 in ESCs (Figure 4.11).

The immunoprecipitated LSD1 displayed 66.23% of HDAC2's decrotonylation activity (Figure 4.12). LSD1 does not have any native deacetylase/deacylase activity. This function likely involves CoREST, as LSD1 and HDAC1/2 co-occur in this ubiquitous complex, permitting LSD1 to recruit HDAC1/2 deacylase activity (Shi et al., 2005). The reduced LSD1 associated activity, as compared to HDAC2 in the decrotonylation assay, speaks to this indirect effect. LSD1's ability to mediate decrotonylation through the recruitment of the CoREST complex likely signifies the involvement of the other HDAC1/2 complexes in histone decrotonylation. The existence of the SIN3 co-repressor complex constituents, and the histone crotonylation mark, in yeast suggests a sort of regulatory co-evolution (Kelly and Cowley, 2013, Tan et al., 2011). Taken together, it seems reasonable that the Sin3a co-repressor complex would play an important role in lysine decrotonylation.

6.4. Crotonylation on Non-Histone Proteins

Just as the histones were discovered not to be unique in bearing the acetylation modification, non-histone protein crotonylation appears to be reasonably widespread. Xu and colleagues (2017) found 1,024 proteins crotonylated at nearly 2,700 sites. Another group determined the number of crotonylated proteins to be 2,021, in an adenocarcinoma cell line (Wu et al., 2017). This makes the PTM similarly ubiquitous to acetylation (Choudhary et al., 2009). A simple motif began to emerge from the crotonylated non-histone proteins, where crotonylated lysine residues were disproportionately found to be either followed or preceded by an aspartic acid or glutamate residue (Xu et al., 2017). This very basic motif is evolutionarily conserved quite tightly and is even maintained in plants (Sun et al., 2017).

6.5. The Regulatory Importance of Cr in Non-Histone Proteins

Adding to the regulatory complexity of the non-histone crotonylation pathway, there appears to be a positive transcriptional feedback loop. Crotonylation of histones increases transcription. Similarly, the crotonylation of several non-histone proteins, involved in chromatin condensation, results in the inhibition of the protein's native transcriptional repression. HP1 α binds to methylated histones and its localization results in the recruitment of HP1 β and histone methyltransferase SETDB1, and the enhanced H3K9 trimethylation independent of the polycomb group of proteins (Verschure et al., 2005). This confluence of events, resulting from HP1 α activity, leads to large-scale chromatin condensation. The crotonylation of HP1 α doesn't permit the binding of the protein to H3K9me3, which likely has expansive implications on chromatin remodeling (Wei et al., 2017). Wei and colleagues (2017) supplementary data reveal that the lysine methyltransferase KMT2E is endogenously crotonylated, though the effects of this specific PTM is less clear.

HDAC1 and HDAC2 are both acetylated (Qiu et al., 2006, Eom et al., 2014). Indeed, the two HDACs share a large degree of sequence similarity, included in this are five of the six lysine residues that are acetylated in HDAC1 (Luo et al., 2009). HDAC1 acetylation impedes its lysine deacetylase activity (Qiu et al., 2006). The crotonylation of the HDAC1 appears to similarly reduce the enzyme's activity on histones (Wei et al., 2017). While it remains to be seen if HDAC2 is crotonylated, it is known that the acetylation of HDAC1 inhibits HDAC2 function (Luo et al., 2009). The implication of this finding is that HDAC1 crotonylation might also constrain HDAC2 activity.

6.6. The Role of HDAC1/2 in Non-Histone Protein Decrotonylation

HDAC1/2 likely maintain the same decrotonylase activity on histone and non-histone substrates. Indeed, the BOC-K(Cr)-AMC assays only measured the general lysine decrotonylation activity of the HDACs (Figure 4.3, Figure 4.7 and Figure 4.12). While this assay's substrate was not complex enough to simulate a non-histone protein, there is evidence in the literature that supports the idea that classical HDACs are responsible

for the decrotonylation of non-histone proteins. Echoing the experiment illustrated in Figure 4.4, Xu et al. (2017) discovered that the application of TSA, but not the sirtuin inhibitor nicotinamide, increases the level of NPM1 crotonylation. This specific example of non-histone protein crotonylation, likely speaks to a larger trend where classical HDACs, especially HDAC1/2, function as the principal mammalian protein decrotonylases. Wu and colleagues' (2017) findings support this, the application of SAHA, a class I and II HDAC inhibitor, revealed an increase of crotonylation on 2,021 proteins, at over ten thousand individual sites. Mirroring their histone decrotonylation activity in ESCs, HDAC1/2 would expectedly play a central regulatory role in this expanded crotonylation network.

6.7. TET1's Role in Decrotonylation

There appears to be a trend in the TET proteins, where the ability to directly interact with DNA, through a zinc finger domain, appears to be coupled with the capacity to bind Sin3a. As evidenced in this thesis, TET1 is able to recruit Sin3a repressor complex activity in a reporter assay. As HDAC1/2 form the enzymatic core of the Sin3a repressor complex and are, in addition, amongst the most potent regulators of histone crotonylation, it seems logical that TET1/3 might target this activity to chromosomes via both their zinc finger and Sin3a interaction domains. This is in keeping with the broader pattern seen in alternative acylations; they are not governed by novel proteins, but are instead managed by known proteins functioning in novel regulatory roles. The difference in binding strength between TET1 and MAD1, for Sin3a, did not translate into any deviation in biological effect. It therefore seems likely that this may be the case in the likely role that TET1 has in regulating crotonylation through Sin3a.

Chapter 7. Appendix

Evolutionary Conservation of the TET1-SID

DauidsMyotisBat-TET1	--IEALaQLS
BrandtsBat-TET1	--IEALaQLS
RockDove-TET1	VAIEALTrLS
AdeliePenguin-TET1	VAIEALTrLS
WesternPaintedTurtle-TET1	VAIEALTkLS
NorthernGreaterGalago-TET1	VAIEALTkLS
GreenSeaTurtle-TET1	VAIEALTkLS
Human-TET1	VAIEALTQLS
Chimpanzee-TET1	VAIEALTQLS
Mouse-TET1	VAIEALTQLS
BlueTit-TET1	VAIEALTQLS
Finch-TET1	VAIEALTQLS
AmericanFlamingo-TET1	VAIEALTQLS
DomesticCat-TET1	VAIEALTQLS
FlyingFoxBat-TET1	VAIEALTQLS
WhiteTailedDeer-TET1	VAIEALTQLS
Goat-methylcytosineDioxygenase	VAIEALTQLS
AfricanElephant-TET1	VAIEALTQLS
BelugaWhale-TET1	VAIEALTQLS
NineBandedArmadillo-TET1	VAIEALTQLS
FloridaManatee-TET1	VAIEALTQLS
SpermWhale-TET1	VAIEALTQLS
WildBoar-TET1	VAIEALTQLS
DomesticDog-TET1	VAIEALTQLS
SeaOtter-TET1	VAIEALTQLS
NorthAmericanBeaver-TET1	VAIEALTQLS
ThirteenLinedGroundSquirrel-TET1	VAIEALTQLS
NakedMoleRat-TET1	VAIEALTQLS
GuineaPig-TET1	VAIEALTQLS
BactrianCamel-TET1	VAIEALTQLS
TasmanianDevil-TET1	VAIEALTQLS
PhillipineTarsier-TET1	VAIEALTQLS
RiceFish-TET1Like	VAIEALTQL-
SouthernPlatyfish-TET3	-AIdALTQLS
ZebraMbunaCichlid-TET3	iAIdALTQLS
Zebrafish-TET3	iAIEALTQLS
WhaleShark-TET3	iAIEALTQLS
RainbowTrout-TET3Like	iAIEALTQLS
DomesticCat-TET3	iAIEALTQLS
Human-TET3	iAIEALTQLS
Mouse-TET3	iAIEALTQLS
Finch-TET3	iAIEALTQLS
SmallYellowCroaker-TET2	iAIEALTQL-
GreaterAmberjack-TET1	iAIEALTQL-
ZebraMbunaCichlid-TET1	iAIEALTQL-
SouthernPlatyfish-TET1	iAIEALTQL-
Killifish-TET3	iAIEALTQL-
MexicanTetra-TET1	iAIEALTQL-
Mudskipper-TET2Like	iAIEALTQL-

Table 7.1: Table of TET1-SID Evolutionary Conservation

Buffers and Solutions	
Cryostorage Media:	60% M10/M15+LIF, 40% Fetal Bovine Serum and 10% Dimethyl sulfoxide
Gal4 VP16 Reporter Assay β-gal Stock Solution:	60mM Na_2HPO_4 , 40mM NaH_2PO_4 and 10mM KCl and 1mM MgCl_2
Gal4 VP16 Reporter Assay Cell Lysis Buffer:	6.25mM Tris-HCl at pH7.8, 10mM DTT, 10mM EDTA, 50% Glycerol and 5% Triton X-100
Gel Filtration Buffer:	50mM Tris-HCl at pH 8.0, 150mM NaCl and 1mM DTT
HDAC Buffer:	50mM NaCl and 50mM Tris-HCl at pH 7.5
HDAC Lysis Buffer:	50mM NaCl and 50mM Tris-HCl at pH 7.5, 5% Glycerol and 0.3% Triton
His Column Buffer A:	20mM Tris-HCl at pH8, 0.5M NaCl and 20mM Imidazole Adjusted to pH8 (with ~0.2mL Concentrated HCl)
His Column Buffer B:	20mM Tris-HCl at pH8, 0.5M NaCl and 0.5M Imidazole Adjusted to pH8 (with ~6mL Concentrated HCl)
Histone Extraction Buffer:	0.2M H_2SO_4
IP Buffer:	250mM NaCl, 10% Glycerol, 0.5% IEGPAL and 50mM Tris-HCl at pH7.5
M10 Media:	89% Dulbecco's Modified Eagle's Medium, 10% Fetal Bovine Serum and 1% Penicillin/Streptomycin/L-glutamine
M15+LIF Media:	83% Knock-Out Dulbecco's Modified Eagle's Medium, 15% Fetal Bovine Serum, 1% Penicillin/Streptomycin/L-glutamine, 5mM β -mercaptoethanol and 1000U/ml Leukemia Inhibitory Factor
NMR Buffer:	20mM Tris-HCl at pH 7.4, 50mM NaCl

	and 1mM DTT
PCR Reaction:	10% 10×Standard Taq Reaction Buffer (contains MgCl ₂), 200μM of all four dNTPs, 0.2μM forward primer, 0.2μM reverse primer, 0.5μg DNA template, 1U/40μL Taq DNA Polymerase, sterile nuclease-free water
Q Column Buffer A:	20mM Tris-HCl at pH8 and 60mM NaCl
Q Column Buffer B:	20mM Tris-HCl at pH8 and 1M NaCl
qRT-PCR Sensifast Reaction:	62.5% Sensifast, 2mM Forward and Reverse Primer, 31.25% Nuclease-Free Autoclaved Water
Western Blot Loading Buffer:	4×LDS Sample Buffer and 0.375M DTT
Western Blot Running Buffer:	MES SDS Running Buffer and 0.5% NuPAGE Antioxidant
Western Blot Transfer Buffer:	TG Buffer with 10% Methanol and 0.125% SDS Solution
Western Blot Wash Buffer:	PBS with 0.1% Tween 20
Whole Cell Extract Buffer:	Western Blot Loading Buffer and 0.1% TSA

Table 7.2: Table of Buffers and Solutions.

Apart from the constructs that used in prior publications, the HALO-TET1 plasmid [Promega, Cat: FHC23878] and the HALO-hSin3a plasmid [Promega, Cat: FHC11647] were used to generate the protein fragments used in this thesis.

Primer/Gene Block	Sequence
T1 746-72-F:	ATCATTTTGGCAAAGAATTCaccATGgattacaaggatcatgacattgac tacaagacgatgacgacaagACCAAATCTCCAAAATTGTTT
T1 951-72-R:	GGAGGGAGAGGGGCGGAATTCtcaTTTTGGTGGCTCTCCC TGTA
T1 1255-72-R:	GGAGGGAGAGGGGCGGAATTCtcaCACCTTTTCCTCTGCT GATTC
T1 1800-72-F:	CATCATTTTGGCAAAGAATTCaccATGgattacaaggatcatgacattgac actacaagacgatgacgacaagACCTTAGGGAGTAACACTGAG
T1 2136-72-R:	GGAGGGAGAGGGGCGGAATTCtcaTCAGACCCAATGGTT ATAGGG
PL72TET1-887R:	GGAGGGAGAGGGGCGGAATTCtcaTGTTAATCTCCTATCT TTCAT
PL72TET1-909R:	GGAGGGAGAGGGGCGGAATTCtcaGGAGGAATTCTCTGA TGGGGC
TET1-Δ891-902-5':	AGGAGATTAACATTGGAGCAAGCCCCATCAGAGAATTC CTCC
TET1-Δ891-902-3':	GGAGGAATTCTCTGATGGGGCTTGCTCCAATGTTAATCT CCT
PL02 GST-878F:	TACTTCCAATCCatgCTCTTATCGTTAATGAAAGAT
PL02 GST-910R:	TATCCACCTTTACTGtcaTGGGGAGGAATTCTCTGATGG
PL02 GST-868F:	TACTTCCAATCCatgCCAAAAGATGGATCTCCCGTT
PL02 GST-927R:	TATCCACCTTTACTGtcaCAAAGTGGCTGTTCTCTGCTC
PL02 GST-848F:	TACTTCCAATCCatgGCTAGTATACACAATGAAGGT
PL02 GST-952R:	TATCCACCTTTACTGtcaTGGTGGCTCTCCCTGTAAGTT
PL42TET1-746F:	AGGAGATATACATatgACCAAATCTCCAAAATTGTTT
PL42TET1-951R:	GAAGTACAGGTTCTCtcaTTTTGGTGGCTCTCCCTGTAAG

PL02Sin3a115F:	TACTTCCAATCCatg CAGCAGCAATTCAGAGGCTG
PL02Sin3a212R:	TATCCACCTTTACTGtca GGAATCTGATGAACCTGGCC
PL02Sin3a292F:	TACTTCCAATCCatg ACGGCCCCATCCTTGCAGAAC
PL02Sin3a388R:	TATCCACCTTTACTGtca AAGCACGGAGCTGTTGGCATC
PL02Sin3a454F:	TACTTCCAATCCatg GATGCCAGCAAACATGGTGGT
PL02Sin3a530R:	TATCCACCTTTACTGtca CAGATGTACAGACTCCTTATA
PL02Sin3a115F:	TACTTCCAATCCatg CAGCAGCAATTCAGAGGCTG
PL02Sin3a212R:	TATCCACCTTTACTGtca GGAATCTGATGAACCTGGCC
PL42TET1-2F:	AGGAGATATACATatg TCTCGATCCCGCCATGCAAGG
PL42TET1-383R:	GAAGTACAGGTTCTCtca TGGGTCAGCACCAGGAAGTTC
PL42TET1-384F:	AGGAGATATACATatg GTTTCATGGTGAGGCCCTGGGT
PL42TET1 745R:	GAAGTACAGGTTCTCtca TCGTTTCTTGTCAACTTCAGA
PL42TET1-746F:	AGGAGATATACATatg ACCAAATCTCCAAAATTGTTT
PL42TET1-951R:	GAAGTACAGGTTCTCtca TTTTGGTGGCTCTCCCTGTAAG
PL42TET1-952F:	AGGAGATATACATatg CTTAATCACTGTCCATCTTTG
PL42TET1-1255R:	GAAGTACAGGTTCTCtca CACCTTTTCCTCTGCTGATTC
PL42TET1-1256F:	AGGAGATATACATatg AAGGTTGAACCATTGGATTCA
PL42TET1-1799R:	GAAGTACAGGTTCTCtca TGGCAGTGACGAAGGCTTACT
PL91TET1_878-911:	5'- GTA TTT TCA GGG CGC CCT CTT ATC GTT AAT GAA AGA TAG GAG ATT AAC ATT GGA GCA AGT GGT AGC CAT AGA GGC CCT GAC TCA ACT CTC AGA AGC CCC ATC AGA GAA TTC CTC CCC ATG AAA TTC GAG CTC CGT C -3'
PL91TET1_878-911A893DA896D:	5'- TAC TTC CAA TCC ATG CTC TTA TCG TTA ATG AAA GAT AGG AGA TTA ACA TTG GAG CAA GTG GTA GAC ATA GAG GAC CTG ACT CAA CTC TCA GAA GCC CCA TCA GAG AAT TCC TCC CCA TGA CAG TAA AGG TGG ATA -3'
PL91TET1_878-911I894A:	5'- TAC TTC CAA TCC ATG CTC TTA TCG TTA ATG AAA GAT AGG AGA TTA ACA TTG GAG CAA GTG GTA GCC GCC GAG GCC CTG ACT CAA CTC TCA GAA GCC CCA TCA GAG AAT TCC TCC CCA TGA CAG TAA AGG TGG ATA -3'
PL91TET1_878-911I894E:	5'- TAC TTC CAA TCC ATG CTC TTA TCG TTA ATG AAA GAT AGG AGA TTA ACA TTG GAG CAA GTG GTA GCC GAG GAG GCC CTG ACT CAA CTC TCA GAA GCC CCA TCA GAG AAT TCC TCC CCA TGA CAG TAA AGG TGG ATA -3'

PL91TET1_878-911I894Q	5'- TAC TTC CAA TCC ATG CTC TTA TCG TTA ATG AAA GAT AGG AGA TTA ACA TTG GAG CAA GTG GTA GCC CAG GAG GCC CTG ACT CAA CTC TCA GAA GCC CCA TCA GAG AAT TCC TCC CCA TGA CAG TAA AGG TGG ATA -3'
PL91TET1_878-911L897AL900A	5'- TAC TTC CAA TCC ATG CTC TTA TCG TTA ATG AAA GAT AGG AGA TTA ACA TTG GAG CAA GTG GTA GCC ATA GAG GCC GCC ACT CAA GCC TCA GAA GCC CCA TCA GAG AAT TCC TCC CCA TGA CAG TAA AGG TGG ATA -3'
PL91TET1_878-911L897EL900E:	5'- TAC TTC CAA TCC ATG CTC TTA TCG TTA ATG AAA GAT AGG AGA TTA ACA TTG GAG CAA GTG GTA GCC ATA GAG GCC GAG ACT CAA GAG TCA GAA GCC CCA TCA GAG AAT TCC TCC CCA TGA CAG TAA AGG TGG ATA -3'
PL42Sin3a115-212:	5'- AGG AGA TAT ACA TAT GCA GCA GCA ATT TCA GAG GCT GAA GGT GGA GGA TGC GCT ATC TTA TCT TGA CCA GGT GAA GCT GCA GTT TGG TAG TCA GCC TCA GGT CTA CAA TGA TTT CCT TGA CAT CAT GAA GGA ATT TAA ATC TCA GAG CAT CGA CAC CCC AGG AGT GAT TAG TCG TGT GTC CCA GCT ATT CAA AGG CCA CCC CGA TCT GAT AAT GGG ATT CAA CAC CTT CTT GCC CCC TGG CTA CAA AAT TGA GGT GCA AAC CAA TGA CAT GGT GAA TGT GAC AAC TCC TGG CCA GGT TCA TCA GAT TCC CTG AGA GAA CCT GTA CTT C -3'
PL01Sin3a115F:	TACTTCCAATCCatg CAGCAGCAATTT CAGAGGCTG
PL01Sin3a212R:	TATCCACCTTTACTGtca GGAATCTGATGAACCTGGCC
PL91SAP25_126-186:	5'- GTA TTT TCA GGG CGC CTC CTC CAC CTG GCT CAG CGA GGC AGA GAT GAT CGC CCT TGC TGG TCT GCT GCA GAT GAG CCA GGG AGA GCA GAC ACC CAA CTG CGT GGC AAG CTC TCT GCC TTC CAC CAG CTG CCC AGA CCC TGT CTC TGT CTC TGA AGA CCC AGG TCC CAG TGG TGA CCA GAG CTG TTC TGG GAC TGA CAC ATG AAA TTC GAG CTC CGT C -3'

Table 7.3: Table of Primers and GBlocks.

Red denotes the N and C terminal arms of homology needed for insertion into the required vectors. Green denotes a flag tag (DYKDHDIDYKDDDDK).

Vector Name	Organism	Antibiotic Resistance	Tag	Features	Vector Family
pLEICS-01	<i>E. Coli</i>	Amp	N-HIS6	TEV Extra SM	C
pLEICS-02	<i>E. Coli</i>	Amp	N-GST	TEV Extra SM	C
pLEICS-42	Cell-free Wheat Germ Lysate	Amp	C-HIS ₆	TEV Extra ENLYFQ	B
pLEICS-72	Mouse Embryonic Stem Cell	AMP	EGFP		72
pLEICS-84	Mammalian cells	Amp	GAL4		A
PLEICS-91	<i>E. Coli</i>	Amp	N-GB1/HIS6	TEV Extra GA	A

Table 7.5: Table of Cloning Vectors.

Families refer to plasmids that contain the same arms of homology, An example of this are that primers or a gBlock designed for pLEICS-01 would similarly fit inside pLEICS-02 as they both belong to Family C. The arms of homology are listed below (Yang, 2017).

Family A vectors homology region:

N-ter: **GTATTTTCAGGGCGCC...**

C-ter: **GACGGAGCTCGAATTTC...**

Family B vectors homology region:

N-ter: **AGGAGATATACATATG...**

C-ter-no tag: **GAAGTACAGGTTCTCTCA...**

Family C vectors homology region:

N-ter: **TACTTCCAATCCATG...**

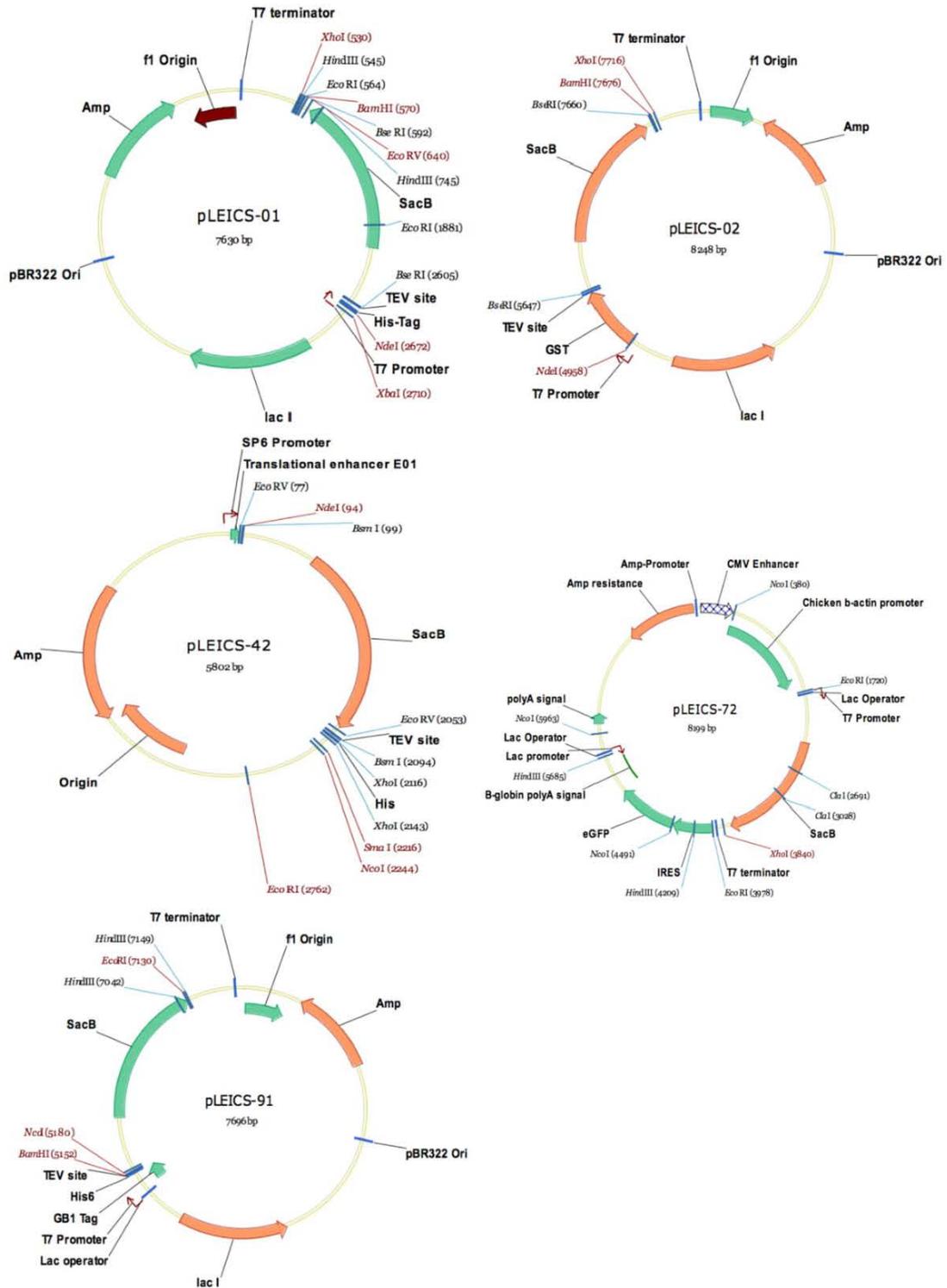
C-ter: **TATCCACCTTTACTGTCA...**

Family 72 vector homology region:

N-ter: **CATCATTTTGGCAAAGAATTCatg ...**

C-ter: **GGAGGGAGAGGGGCGGAATTCtea ...**

Vector Maps from the University of Leicester Protex Service



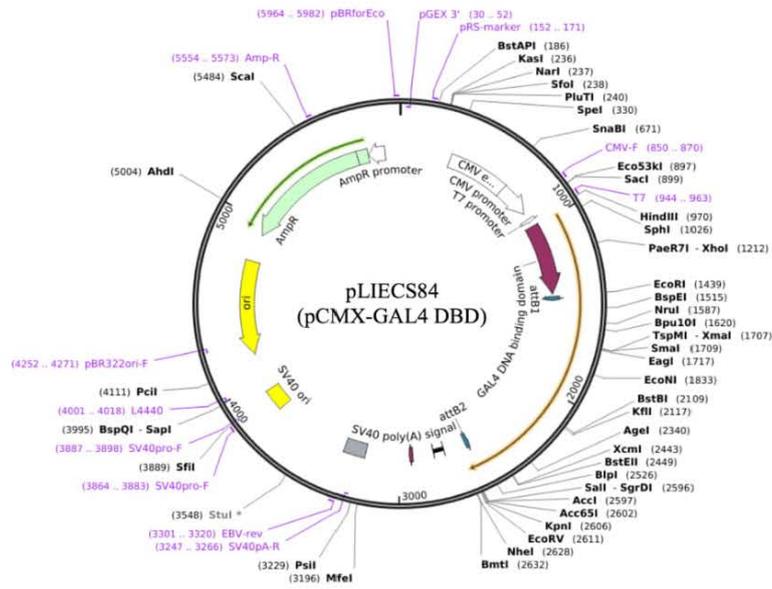


Figure 7.1: Vectors Provided by the University of Leicester Protex Service (Yang, 2017, Zhao et al., 2008)

qRT-PCR Primers	
GAPDH-L1:	GGGAAGCCCATCACCATCTTC
GAPDH-R1:	AGAGGGGCCATCCACAGTCT
qPCR-TET1DNA2668F:	CAAGTGGTAGCCATAGA
qPCR-TET1DNA2831R:	GGGTCTTGGAGGTCTTT
TET2a-F:	AGAGAAGACAATCGAGAAGTCGG
TET2a-R:	CCTTCCGTACTCCCAAACATCAT
L1_AF:	TCTGGTGAGTGGAACACAGC
L1_AR:	AGTCTCGAGTGGAGCGGAAG
L1_TF:	GCCTAAGCCACAGCAGCA
L1_TR:	GCTGTCAGGTTCTCTGGCG
LIF-F:	CCCAGCATCCCAGAACCATT
LIF-R:	TTTGTCAACCAAGGCCAAGT
IDAXaF:	CTGCCCGCAGAATCATTCT
IDAXaR:	CAGACGCCACAGTTGATGAG
EP300microF:	GACCAGTTCCCCACATCCTG
EP300microR:	GAACCACACACACAAGGGGA

Table 7.6: Primers Used for qRT-PCR.

RT-PCR Primers	
TET1-RT-L1:	CTGTTCGACGGCTGTGATAGT
TET1-RT-R1:	TCCCAAACCTTACAGCCGTTGA
TET2-RT-L2:	CCACAGAGACCAGCAGAACAT
TET2-RT-R2:	TCTTCTTGCAACTCTTGGGCA
Sin3a-RT-L1:	AACACCTGGCCAAGTTCATCA
Sin3a-RT-R1:	CCACAGGCTGATTGTTTTGCA

Table 7.7: Primers Used for RT-PCR.

Sequencing Primers for PiggyBAC TET1	
PB SP1:	GGAGCCTACCTAGACTCAGC
TET1 SP2:	CTAGGACAGGCCTTTGGTGC
TET1 SP3:	TATACCACTTTGCTACCGAC
TET1 SP4:	AGAATTCGGCAAGACATTGG
TET1 SP5:	AGCATCCACAAAGTCACATG
TET1 SP6:	CTGGCTCAAACGAGGTCCAT
TET1 SP7:	CATGGAGAATAGGTATGGTC
TET1 SP8:	CATGAATAATGGAAGCACTG
TET1 SP9:	GCATTCACAGCTTGGCGAA
Sequencing primers for PiggyBAC TET1's L897A/L900A region:	
TET1-871LLAAF:	TCCCGTTCAACCAAGTCTCT
TET1-953LLAAR:	AAGTTT TGGTGGCTCTCCCT
Sequencing primers for PiggyBAC TET1's ΔCXXC (891-902)	
TET1RTPCR-CXXC-F:	TCTTCCATGCTTCACTGGGT
TET1RTPCR-CXXC-R:	T GGGGCCTCTTGTTTTTCCTT

Table 7.8: Primers Used to Sequence PiggyBAC TET1 Constructs.

Antibody	Manufacturer and Catalog/Product Number	Dilutions Used
Anti-acetyl-Histone H3 (Lys18) Rabbit pAb	Merck Millipore Catalogue number: 07-354	WB 1:10,000
Anti-HaloTag Monoclonal Antibody. Mouse	Promega Part No. G921A	WB 1:1000
Anti-HDAC1 antibody [EPR460(2)] Rabbit mAb	Abcam Cat: Ab 109411	WB 1:1000
Anti-HDAC2, clone 3F3	Millipore Catalog number: 05-814	WB 1:1000 IP: 1:20
Anti-KDM1 / LSD1 antibody [EPR6825]	Abcam Cat: Ab129195	WB: 1:10,000 IP: 1:4
Anti-Methylecytosine dioxygenase TET1 Rabbit pAb	Merck Millipore Cat: 09-872	WB 1:200
Anti-Sin3a antibody [EPR6780] Rabbit mAb	Abcam Cat: Ab129087	WB 1:1000 IP:1:20
Anti-TET2 antibody Rabbit pAb	Abcam Cat: Ab124297	WB 1:250
Crotonyl-Histone H3 Lys18 Rabbit pAb	PTM Biolabs LLC Item code: PTM-517	WB 1:2000
GAL4 (DBD) (RK5C1) Mouse mAb	Santa Cruz Biotechnology Inc. Cat: Sc-510	WB 1:500
HDAC2 clone 3F3 Mouse mAb	Merck Millipore Cat: 05-814	WB 1:1000
Monoclonal ANTI-FLAG M2 Mouse mAb	Sigma-Aldrich Cat: F1804	WB 1:1000 IP: 1:20
Myc-Tag Mouse mAb	Cell Signaling Technology	WB 1:1000

	Cat: 9B11	IP: 1:20
Pan anti-crotonyllysine Rabbit pAb	PTM Biolabs LLC Item code: PTM-501	WB 1:1000
VP16 (1-21) Mouse mAb	Santa Cruz Biotechnology Inc. Cat: Sc-7545	WB 1:500
Sin3a (AK-11) Rabbit pAb	Santa Cruz Biotechnology Inc. Cat: Sc-767	WB 1:1000

Table 7.9: Table of Antibodies.

WB dilution is the ratio of antibody to blocking buffer. IP dilution is the ratio of antibody to volume of un-compacted sepharose beads (Yang, 2017).

Chapter 8. Bibliography

- ABRAJANO, J. J., QURESHI, I. A., GOKHAN, S., ZHENG, D., BERGMAN, A. & MEHLER, M. F. 2009. REST and CoREST Modulate Neuronal Subtype Specification, Maturation and Maintenance. *PLOS ONE*, 4, e7936.
- AKAHORI, H., GUINDON, S., YOSHIZAKI, S. & MUTO, Y. 2015. Molecular Evolution of the TET Gene Family in Mammals. *Int J Mol Sci*, 16, 28472-85.
- ALLFREY, V. G., FAULKNER, R. & MIRSKY, A. E. 1964. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proceedings of the National Academy of Sciences of the United States of America*, 51, 786-794.
- ALLIS, C. D., CHICOINE, L. G., RICHMAN, R. & SCHULMAN, I. G. 1985. Deposition-related histone acetylation in micronuclei of conjugating Tetrahymena. *Proc Natl Acad Sci U S A*, 82, 8048-52.
- ANDRES, M. E., BURGER, C., PERAL-RUBIO, M. J., BATTAGLIOLI, E., ANDERSON, M. E., GRIMES, J., DALLMAN, J., BALLAS, N. & MANDEL, G. 1999. CoREST: a functional corepressor required for regulation of neural-specific gene expression. *Proc Natl Acad Sci U S A*, 96, 9873-8.
- ANNE, M., SAMMARTINO, D., BARGINEAR, M. F. & BUDMAN, D. 2013. Profile of panobinostat and its potential for treatment in solid tumors: an update. *OncoTargets and therapy*, 6, 1613-1624.
- ARAMSANGTIENCHAI, P., SPIEGELMAN, N. A., HE, B., MILLER, S. P., DAI, L., ZHAO, Y. & LIN, H. 2016. HDAC8 Catalyzes the Hydrolysis of Long Chain Fatty Acyl Lysine. *ACS Chem Biol*, 11, 2685-2692.
- ARANY, Z., HUANG, L. E., ECKNER, R., BHATTACHARYA, S., JIANG, C., GOLDBERG, M. A., BUNN, H. F. & LIVINGSTON, D. M. 1996. An essential role for p300/CBP in the cellular response to hypoxia. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 12969-12973.
- ARNAUDO, A. M. & GARCIA, B. A. 2013. Proteomic characterization of novel histone post-translational modifications. *Epigenetics Chromatin*, 6, 24.
- ATCHLEY, D. W. & NICHOLS, E. G. 1925. The influence of protein concentration on the conductivity of human serum. *Journal of Biological Chemistry*, 65, 729-734.
-

- AYER, D. E., LAWRENCE, Q. A. & EISENMAN, R. N. 1995. Mad-max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. *Cell*, 80, 767-776.
- BANNING, C. 1946. Food Shortage and Public Health, First Half of 1945. *The Annals of the American Academy of Political and Social Science*, 245, 93-110.
- BANNISTER, A. J., SCHNEIDER, R. & KOUZARIDES, T. 2002. Histone Methylation. *Cell*, 109, 801-806.
- BANTSCHIEFF, M., HOPF, C., SAVITSKI, M. M., DITTMANN, A., GRANDI, P., MICHON, A. M., SCHLEGL, J., ABRAHAM, Y., BECHER, I., BERGAMINI, G., BOESCHE, M., DELLING, M., DUMPELFELD, B., EBERHARD, D., HUTHMACHER, C., MATHIESON, T., POECKEL, D., READER, V., STRUNK, K., SWEETMAN, G., KRUSE, U., NEUBAUER, G., RAMSDEN, N. G. & DREWES, G. 2011. Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat Biotechnol*, 29, 255-65.
- BAO, X., WANG, Y., LI, X., LI, X.-M., LIU, Z., YANG, T., WONG, C. F., ZHANG, J., HAO, Q. & LI, X. D. 2014. Identification of 'erasers' for lysine crotonylated histone marks using a chemical proteomics approach. *eLife*, 3, e02999.
- BARKAUSKAITE, E., JANKEVICIUS, G., LADURNER, A. G., AHEL, I. & TIMINSZKY, G. 2013. The recognition and removal of cellular poly(ADP-ribose) signals. *FEBS J*, 280, 3491-507.
- BARKER, D. J. 1995. Fetal origins of coronary heart disease. *BMJ : British Medical Journal*, 311, 171-174.
- BARLOW, D. P. & BARTOLOMEI, M. S. 2014. Genomic Imprinting in Mammals. *Cold Spring Harbor Perspectives in Biology*, 6.
- BARTOLOMEI, G., LEUTERT, M., MANZO, M., BAUBEC, T. AND HOTTIGER, M.O. 2016. Analysis of Chromatin ADP-Ribosylation at the Genome-wide Level and at Specific Loci by ADPr-ChAP. *Molecular Cell*, 61, 474-85.
- BAUBEC, T., IVANEK, R., LIENERT, F. & SCHUBELER, D. 2013. Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell*, 153, 480-92.
- BAUER, C., GÖBEL, K., NAGARAJ, N., COLANTUONI, C., WANG, M., MÜLLER, U., KREMMER, E., ROTTACH, A. & LEONHARDT, H. 2015.

- Phosphorylation of TET Proteins Is Regulated via O-GlcNAcylation by the O-Linked N-Acetylglucosamine Transferase (OGT). *Journal of Biological Chemistry*, 290, 4801-4812.
- BENYSHEK, D. C., JOHNSTON, C. S. & MARTIN, J. F. 2006. Glucose metabolism is altered in the adequately-nourished grand-offspring (F3 generation) of rats malnourished during gestation and perinatal life. *Diabetologia*, 49, 1117-1119.
- BERGER, S. L. 2007. The complex language of chromatin regulation during transcription. *Nature*, 447, 407-12.
- BJERLING, P., SILVERSTEIN, R. A., THON, G., CAUDY, A., GREWAL, S. & EKWALL, K. 2002. Functional Divergence between Histone Deacetylases in Fission Yeast by Distinct Cellular Localization and In Vivo Specificity. *Molecular and Cellular Biology*, 22, 2170-2181.
- BORMAN, S. 1987. Eluent, Effluent, Eluate, and Eluite. *Analytical Chemistry*, 59, 99A-99A.
- BROWN, M. C., PERROTTA, J. A. & TURNER, C. E. 1998. Serine and threonine phosphorylation of the paxillin LIM domains regulates paxillin focal adhesion localization and cell adhesion to fibronectin. *Mol Biol Cell*, 9, 1803-16.
- BROWNELL, J. E., ZHOU, J., RANALLI, T., KOBAYASHI, R., EDMONDSON, D. G., ROTH, S. Y. & ALLIS, C. D. 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, 84, 843-51.
- BRUBAKER, K., COWLEY, S. M., HUANG, K., LOO, L., YOCHUM, G. S., AYER, D. E., EISENMAN, R. N. & RADHAKRISHNAN, I. 2000. Solution structure of the interacting domains of the Mad-Sin3 complex: implications for recruitment of a chromatin-modifying complex. *Cell*, 103, 655-65.
- CAMPOREALE, G., SHUBERT, E. E., SARATH, G., CERNY, R. & ZEMPLINI, J. 2004. K8 and K12 are biotinylated in human histone H4. *Eur J Biochem*, 271, 2257-63.
- CHAMPAGNE, K. S. & KUTATELADZE, T. G. 2009. Structural Insight Into Histone Recognition by the ING PHD Fingers. *Current drug targets*, 10, 432-441.

- CHEN, T., UEDA, Y., DODGE, J.E., WANG, Z. AND LI, E. 2003. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Molecular and Cellular Biology*, 23, 5594–605.
- CHEN, Y., SPRUNG, R., TANG, Y., BALL, H., SANGRAS, B., KIM, S. C., FALCK, J. R., PENG, J., GU, W. & ZHAO, Y. 2007. Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics*, 6, 812-9.
- CHEN, Z.-X. & RIGGS, A. D. 2011. DNA Methylation and Demethylation in Mammals. *Journal of Biological Chemistry*, 286, 18347-18353.
- CHOUDHARY, C., KUMAR, C., GNAD, F., NIELSEN, M. L., REHMAN, M., WALTHER, T. C., OLSEN, J. V. & MANN, M. 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*, 325, 834-40.
- CHOUDHARY, C., WEINERT, B. T., NISHIDA, Y., VERDIN, E. & MANN, M. 2014. The growing landscape of lysine acetylation links metabolism and cell signalling. *Nat Rev Mol Cell Biol*, 15, 536-50.
- CHUN, A. C. & JIN, D. Y. 2003. Transcriptional regulation of mitotic checkpoint gene MAD1 by p53. *J Biol Chem*, 278, 37439-50.
- CICCARONE, F., VALENTINI, E., ZAMPIERI, M. & CAIAFA, P. 2015. 5mC-hydroxylase activity is influenced by the PARylation of TET1 enzyme. *Oncotarget*, 6, 24333-24347.
- CLARK, M.D., MARCUM, R., GRAVELINE, R., CHAN, C.W., XIE, T., CHEN, Z., DING, Y., et al. 2015. Structural insights into the assembly of the histone deacetylase-associated Sin3L/Rpd3L corepressor complex. *Proceedings of the National Academy of Sciences*, 112, E3669-78.
- COMPTON, S. J. & JONES, C. G. 1985. Mechanism of dye response and interference in the Bradford protein assay. *Anal Biochem*, 151, 369-74.
- CORPET, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res*, 16, 10881-90.
- COWLEY, S. M., IRITANI, B. M., MENDRYSA, S. M., XU, T., CHENG, P. F., YADA, J., LIGGITT, H. D. & EISENMAN, R. N. 2005. The mSin3A

- Chromatin-Modifying Complex Is Essential for Embryogenesis and T-Cell Development. *Molecular and Cellular Biology*, 25, 6990-7004.
- COWLEY, S. M., KANG, R. S., FRANGIONI, J. V., YADA, J. J., DEGRAND, A. M., RADHAKRISHNAN, I. & EISENMAN, R. N. 2004. Functional analysis of the Mad1-mSin3A repressor-corepressor interaction reveals determinants of specificity, affinity, and transcriptional response. *Mol Cell Biol*, 24, 2698-709.
- CULHANE, J. C. & COLE, P. A. 2007. LSD1 and the chemistry of histone demethylation. *Curr Opin Chem Biol*, 11, 561-8.
- DAI, L., PENG, C., MONTELLIER, E., LU, Z., CHEN, Y., ISHII, H., DEBERNARDI, A., BUCHOU, T., ROUSSEAU, S., JIN, F., SABARI, B. R., DENG, Z., ALLIS, C. D., REN, B., KHOCHBIN, S. & ZHAO, Y. 2014. Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat Chem Biol*, 10, 365-370.
- DANIEL, J. M., SPRING, C. M., CRAWFORD, H. C., REYNOLDS, A. B. & BAIG, A. 2002. The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. *Nucleic Acids Research*, 30, 2911-2919.
- DAWLATY, M. M., BREILING, A., LE, T., RADDATZ, G., BARRASA, M. I., CHENG, A. W., GAO, Q., POWELL, B. E., LI, Z., XU, M., FAULL, K. F., LYKO, F. & JAENISCH, R. 2013. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell*, 24, 310-23.
- DAWLATY, M. M., GANZ, K., POWELL, B. E., HU, Y.-C., MARKOULAKI, S., CHENG, A. W., GAO, Q., KIM, J., CHOI, S.-W., PAGE, D. C. & JAENISCH, R. 2011. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell stem cell*, 9, 166-175.
- DEATON, A. M. & BIRD, A. 2011. CpG islands and the regulation of transcription. *Genes Dev*, 25, 1010-22.
- DEBELOUCHINA, G.T., GERECHT, K. AND MUIR, T.W. 2017. Ubiquitin utilizes an acidic surface patch to alter chromatin structure. *Nature Chemical Biology*, 13, 105–10.

- DEPHOURE, N., GOULD, K. L., GYGI, S. P. & KELLOGG, D. R. 2013. Mapping and analysis of phosphorylation sites: a quick guide for cell biologists. *Mol Biol Cell*, 24, 535-42.
- DELVECCHIO, M., GAUCHER, J., AGUILAR-GURRIERI, C., ORTEGA, E. & PANNE, D. 2013. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. *Nat Struct Mol Biol*, 20, 1040-6.
- DEPLUS, R., DELATTE, B., SCHWINN, M. K., DEFRANCE, M., MENDEZ, J., MURPHY, N., DAWSON, M. A., VOLKMAR, M., PUTMANS, P., CALONNE, E., SHIH, A. H., LEVINE, R. L., BERNARD, O., MERCHER, T., SOLARY, E., URH, M., DANIELS, D. L. & FUKS, F. 2013. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J*, 32, 645-55.
- DEREEPER, A., GUIGNON, V., BLANC, G., AUDIC, S., BUFFET, S., CHEVENET, F., DUFAYARD, J. F., et al. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, 36, W465–W469.
- DHALLUIN, C., CARLSON, J.E., ZENG, L., HE, C., AGGARWAL, A.K. AND ZHOU, M.M. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, 39, 491–6.
- DHAWAN, S., GEORGIA, S., TSCHEN, S.-I., FAN, G. & BHUSHAN, A. 2011. Pancreatic beta cell identity is maintained by DNA methylation-mediated repression of Arx. *Developmental cell*, 20, 419-429.
- DHORDAIN, P., LIN, R. J., QUIEF, S., LANTOINE, D., KERCKAERT, J. P., EVANS, R. M. & ALBAGLI, O. 1998. The LAZ3(BCL-6) oncoprotein recruits a SMRT/mSIN3A/histone deacetylase containing complex to mediate transcriptional repression. *Nucleic Acids Research*, 26, 4645-4651.
- DONOHUE, D. R., COLLINS, L. B., WALI, A., BIGLER, R., SUN, W. & BULTMAN, S. J. 2012. The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Mol Cell*, 48, 612-26.
- DOS SANTOS, R. L., TOSTI, L., RADZISHEUSKAYA, A., CABALLERO, I. M., KAJI, K., HENDRICH, B. & SILVA, J. C. 2014. MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. *Cell Stem Cell*, 15, 102-10.

- DOVEY, O. M., FOSTER, C. T. & COWLEY, S. M. 2010. Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 8242-8247.
- DROZDETSKIY, A., COLE, C., PROCTER, J. & BARTON, G. J. 2015. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res*, 43, W389-94.
- DUROCHER, D., HENCKEL, J., FERSHT, A. R. & JACKSON, S. P. 1999. The FHA Domain Is a Modular Phosphopeptide Recognition Motif. *Molecular Cell*, 4, 387-394.
- EGGERS, J. & STEINBUHEL, A. 2013. Poly(3-hydroxybutyrate) degradation in *Ralstonia eutropha* H16 is mediated stereoselectively to (S)-3-hydroxybutyryl coenzyme A (CoA) via crotonyl-CoA. *J Bacteriol*, 195, 3213-23.
- EHRlich, M. 2002. DNA methylation in cancer- too much, but also too little. *Oncogene*, 21, 5400 – 5413.
- EOM, G. H., NAM, Y. S., OH, J. G., CHOE, N., MIN, H.-K., YOO, E.-K., KANG, G., NGUYEN, V. H., MIN, J.-J., KIM, J.-K., LEE, I.-K., BASSEL-DUBY, R., OLSON, E. N., PARK, W. J. & KOOK, H. 2014. Regulation of Acetylation of Histone Deacetylase 2 by p300/CBP-Associated Factor/Histone Deacetylase 5 in the Development of Cardiac Hypertrophy. *Circulation Research*, 114, 1133-1143.
- FANG, C. Y., SHEN, C. H., WANG, M., CHEN, P. L., CHAN, M. W., HSU, P. H. & CHANG, D. 2015. Global profiling of histone modifications in the polyomavirus BK virion minichromosome. *Virology*, 483, 1-12.
- FATEMI, M., HERMANN, A., GOWHER, H. AND JELTSCH, A. 2002. Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *European Journal of Biochemistry*, 269, 4981–4.
- FEIGERLE, J. T. & WEIL, P. A. 2016. The C Terminus of the RNA Polymerase II Transcription Factor IID (TFIID) Subunit Taf2 Mediates Stable Association of Subunit Taf14 into the Yeast TFIID Complex. *Journal of Biological Chemistry*, 291, 22721-22740.
- FIDALGO, M., HUANG, X., GUALLAR, D., SANCHEZ-PRIEGO, C., VALDES, V. J., SAUNDERS, A., DING, J., WU, W. S., CLAVEL, C. & WANG, J. 2016.

- Zfp281 Coordinates Opposing Functions of Tet1 and Tet2 in Pluripotent States. *Cell Stem Cell*, 19, 355-69.
- FILIPPAKOPOULOS, P. & KNAPP, S. 2012. The bromodomain interaction module. *FEBS Letters*, 586, 2692-2704.
- FLEISCHER, T. C., YUN, U. J. & AYER, D. E. 2003. Identification and characterization of three new components of the mSin3A corepressor complex. *Mol Cell Biol*, 23, 3456-67.
- FLYNN, E. M., HUANG, O. W., POY, F., OPPIKOFER, M., BELLON, S. F., TANG, Y. & COCHRAN, A. G. 2015. A Subset of Human Bromodomains Recognizes Butyryllysine and Crotonyllysine Histone Peptide Modifications. *Structure*, 23, 1801-14.
- FOSTER, C. T., DOVEY, O. M., LEZINA, L., LUO, J. L., GANT, T. W., BARLEV, N., BRADLEY, A. & COWLEY, S. M. 2010. Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol Cell Biol*, 30, 4851-63.
- FREUDENBERG, J. M., GHOSH, S., LACKFORD, B. L., YELLABOINA, S., ZHENG, X., LI, R., CUDDAPAH, S., WADE, P. A., HU, G. & JOTHI, R. 2012. Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. *Nucleic Acids Res*, 40, 3364-77.
- FUKS, F., BURGERS, W.A., GODIN, N., KASAI, M. AND KOUZARIDES, T. 2001. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *The EMBO Journal*, 20, 2536-44.
- GAIDZIK, V. I., PASCHKA, P., SPATH, D., HABDANK, M., KOHNE, C. H., GERMING, U., VON LILIENFELD-TOAL, M., HELD, G., HORST, H. A., HAASE, D., BENTZ, M., GOTZE, K., DOHNER, H., SCHLENK, R. F., BULLINGER, L. & DOHNER, K. 2012. TET2 mutations in acute myeloid leukemia (AML): results from a comprehensive genetic and clinical analysis of the AML study group. *J Clin Oncol*, 30, 1350-7.
- GAO, L., CUETO, M. A., ASSELBERGS, F. & ATADJA, P. 2002. Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. *J Biol Chem*, 277, 25748-55.

- GAO, Y., CHEN, J., LI, K., WU, T., HUANG, B., LIU, W., KOU, X., ZHANG, Y., HUANG, H., JIANG, Y., YAO, C., LIU, X., LU, Z., XU, Z., KANG, L., CHEN, J., WANG, H., CAI, T. & GAO, S. 2013. Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell*, 12, 453-69.
- GELMAN, L., ZHOU, G., FAJAS, L., RASPE, E., FRUCHART, J. C. & AUWERX, J. 1999. p300 interacts with the N- and C-terminal part of PPARgamma2 in a ligand-independent and -dependent manner, respectively. *J Biol Chem*, 274, 7681-8.
- GLOZAK, M. A., SENGUPTA, N., ZHANG, X. & SETO, E. 2005. Acetylation and deacetylation of non-histone proteins. *Gene*, 363, 15-23.
- GOLL, M. G., KIRPEKAR, F., MAGGERT, K. A., YODER, J. A., HSIEH, C.-L., ZHANG, X., GOLIC, K. G., JACOBSEN, S. E. & BESTOR, T. H. 2006. Methylation of tRNA^{Asp} by the DNA Methyltransferase Homolog Dnmt2. *Science*, 311, 395.
- GOOD, C. R., MADZO, J., PATEL, B., MAEGAWA, S., ENGEL, N., JELINEK, J. & ISSA, J.-P. J. 2017. A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer. *Nucleic Acids Research*, 45, 8269-8281.
- GOUDARZI, A., ZHANG, D., HUANG, H., BARRAL, S., KWON, O. K., QI, S., TANG, Z., BUCHOU, T., VITTE, A. L., HE, T., CHENG, Z., MONTELLIER, E., GAUCHER, J., CURTET, S., DEBERNARDI, A., CHARBONNIER, G., PUTHIER, D., PETOSA, C., PANNE, D., ROUSSEAU, S., ROEDER, R. G., ZHAO, Y. & KHOCHBIN, S. 2016. Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation Are Hallmarks of Highly Active Gene Promoters. *Mol Cell*, 62, 169-80.
- GOWHER, H., LIEBERT, K., HERMANN, A., XU, G. & JELTSCH, A. 2005. Mechanism of Stimulation of Catalytic Activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *Journal of Biological Chemistry*, 280, 13341-13348.
- GRANDJEAN, V., O'NEILL, L., SADO, T., TURNER, B. & FERGUSON-SMITH, A. 2001. Relationship between DNA methylation, histone H4 acetylation and gene
-

- expression in the mouse imprinted Igf2-H19 domain. *FEBS Letters*, 488, 165-169.
- GRZENDA, A., LOMBERK, G., ZHANG, J. S. & URRUTIA, R. 2009. Sin3: master scaffold and transcriptional corepressor. *Biochim Biophys Acta*, 1789, 443-50.
- GU, T. P., GUO, F., YANG, H., WU, H. P., XU, G. F., LIU, W., XIE, Z. G., SHI, L., HE, X., JIN, S. G., IQBAL, K., SHI, Y. G., DENG, Z., SZABO, P. E., PFEIFER, G. P., LI, J. & XU, G. L. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*, 477, 606-10.
- GUCCIONE, E., BASSI, C., CASADIO, F., MARTINATO, F., CESARONI, M., SCHUCHLAUTZ, H., LÜSCHER, B., et al. 2007. Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature*, 449, 933-37.
- HANSEN, J. C., GHOSH, R. P. & WOODCOCK, C. L. 2010. Binding of the Rett syndrome protein, MeCP2, to methylated and unmethylated DNA and chromatin. *IUBMB Life*, 62, 732-8.
- HASHIMOTO, H., LIU, Y., UPADHYAY, A. K., CHANG, Y., HOWERTON, S. B., VERTINO, P. M., ZHANG, X. & CHENG, X. 2012. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res*, 40, 4841-9.
- HAYES, J. J. & HANSEN, J. C. 2001. Nucleosomes and the chromatin fiber. *Curr Opin Genet Dev*, 11, 124-9.
- HEIDEMANN, K. A., CROWE, S., JACOB, T., CARL ASCOLI, A. 2015. Localization of HDAC1 using high resolution microscopy. In: MICROSYSTEMS, R. I. I. A. L. (ed.) https://www.rockland-inc.com/uploadedFiles/About_Rockland/Collaborations/Rockland_Leica_Localization_of_HDAC1_poster.pdf. Gilbertsville, PA 19525: Rockland Immunochemicals Inc.
- HEIJMANS, B. T., TOBI, E. W., STEIN, A. D., PUTTER, H., BLAUW, G. J., SUSSER, E. S., SLAGBOOM, P. E. & LUMEY, L. H. 2008. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci U S A*, 105, 17046-9.

- HESS, J. L. 2004. MLL: a histone methyltransferase disrupted in leukemia. *Trends Mol Med*, 10, 500-7.
- HIRSCHEY, M. D. & ZHAO, Y. 2015. Metabolic Regulation by Lysine Malonylation, Succinylation, and Glutarylation. *Mol Cell Proteomics*, 14, 2308-15.
- HOMER. 1999. The Odyssey. Edited by Butler, S., Tinsley, J. and Widger, D. The Project Gutenberg. Available at: <http://www.gutenberg.org/files/1727/1727-h/1727-h.htm>.
- HRIT, J., LI, C., MARTIN, E. A., GOLL, M. & PANNING, B. 2017. OGT binds a conserved C-terminal domain of TET1 to regulate TET1 activity and function in development. *bioRxiv*.
- HSU, W. W., WU, B. & LIU, W. R. 2016. Sirtuins 1 and 2 Are Universal Histone Deacetylases. *ACS Chem Biol*, 11, 792-9.
- HU, G. & WADE, P. A. 2012. NuRD and pluripotency: a complex balancing act. *Cell Stem Cell*, 10, 497-503.
- HUANG, H., JIANG, X., LI, Z., LI, Y., SONG, C. X., HE, C., SUN, M., CHEN, P., GURBUXANI, S., WANG, J., HONG, G. M., ELKAHLOUN, A. G., ARNOVITZ, S., WANG, J., SZULWACH, K., LIN, L., STREET, C., WUNDERLICH, M., DAWLATY, M., NEILLY, M. B., JAENISCH, R., YANG, F. C., MULLOY, J. C., JIN, P., LIU, P. P., ROWLEY, J. D., XU, M., HE, C. & CHEN, J. 2013. TET1 plays an essential oncogenic role in MLL-rearranged leukemia. *Proc Natl Acad Sci U S A*, 110, 11994-9.
- HUANG, H., LIN, S., GARCIA, B. A. & ZHAO, Y. 2015. Quantitative proteomic analysis of histone modifications. *Chem Rev*, 115, 2376-418.
- HUANG, Y., CHAVEZ, L., CHANG, X., WANG, X., PASTOR, W. A., KANG, J., ZEPEDA-MARTINEZ, J. A., PAPE, U. J., JACOBSEN, S. E., PETERS, B. & RAO, A. 2014. Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc Natl Acad Sci U S A*, 111, 1361-6.
- ICARDI, L., MORI, R., GESELLCHEN, V., EYCKERMAN, S., DE CAUWER, L., VERHELST, J., VERCAUTEREN, K., SAELENS, X., MEULEMAN, P., LEROUX-ROELS, G., DE BOSSCHER, K., BOUTROS, M. & TAVERNIER, J. 2012. The Sin3a repressor complex is a master regulator of STAT

- transcriptional activity. *Proceedings of the National Academy of Sciences*, 109, 12058-12063.
- IQBAL, K., JIN, S.-G., PFEIFER, G. P. & SZABÓ, P. E. 2011. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 3642-3647.
- ISIDRO-LLOBET, A., ALVAREZ, M. & ALBERICIO, F. 2009. Amino acid-protecting groups. *Chem Rev*, 109, 2455-504.
- ITO, R., KATSURA, S., SHIMADA, H., TSUCHIYA, H., HADA, M., OKUMURA, T., SUGAWARA, A. & YOKOYAMA, A. 2014. TET3-OGT interaction increases the stability and the presence of OGT in chromatin. *Genes Cells*, 19, 52-65.
- ITO, S., D'ALESSIO, A. C., TARANOVA, O. V., HONG, K., SOWERS, L. C. & ZHANG, Y. 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES cell self-renewal, and ICM specification. *Nature*, 466, 1129-1133.
- ITO, S., SHEN, L., DAI, Q., WU, S. C., COLLINS, L. B., SWENBERG, J. A., HE, C. & ZHANG, Y. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science (New York, N.Y.)*, 333, 1300-1303.
- IYER, L. M., TAHILIANI, M., RAO, A. & ARAVIND, L. 2009. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle*, 8, 1698-710.
- JAMALADDIN, S., KELLY, R. D. W., O'REGAN, L., DOVEY, O. M., HODSON, G. E., MILLARD, C. J., PORTOLANO, N., FRY, A. M., SCHWABE, J. W. R. & COWLEY, S. M. 2014. Histone deacetylase (HDAC) 1 and 2 are essential for accurate cell division and the pluripotency of embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 9840-9845.
- JEONG, M., SUN, D., LUO, M., HUANG, Y., CHALLEN, G. A., RODRIGUEZ, B., ZHANG, X., CHAVEZ, L., WANG, H., HANNAH, R., KIM, S. B., YANG, L., KO, M., CHEN, R., GOTTGENS, B., LEE, J. S., GUNARATNE, P., GODLEY, L. A., DARLINGTON, G. J., RAO, A., LI, W. & GOODELL, M. A. 2014.

- Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat Genet*, 46, 17-23.
- JEZIORSKA, D.M., MURRAY, R.J.S., DE GOBBI, M., GAENTZSCH, R., GARRICK, D., AYYUB, H., CHEN, T., ET AL. 2017. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *Proceedings of the National Academy of Sciences*, 114, 7526–35.
- JIANG, H., KHAN, S., WANG, Y., CHARRON, G., HE, B., SEBASTIAN, C., DU, J., KIM, R., GE, E., MOSTOSLAVSKY, R., HANG, H. C., HAO, Q. & LIN, H. 2013. SIRT6 regulates TNF-[agr] secretion through hydrolysis of long-chain fatty acyl lysine. *Nature*, 496, 110-113.
- JIN, B., LI, Y. & ROBERTSON, K. D. 2011. DNA Methylation: Superior or Subordinate in the Epigenetic Hierarchy? *Genes & Cancer*, 2, 607-617.
- JIN, C., LU, Y., JELINEK, J., LIANG, S., ESTECIO, M. R. H., BARTON, M. C. & ISSA, J.-P. J. 2014. TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. *Nucleic Acids Research*, 42, 6956-6971.
- KACZMARSKA, Z., ORTEGA, E., GOUDARZI, A., HUANG, H., KIM, S., MARQUEZ, J. A., ZHAO, Y., KHOCHBIN, S. & PANNE, D. 2017. Structure of p300 in complex with acyl-CoA variants. *Nat Chem Biol*, 13, 21-29.
- KADAMB, R., MITTAL, S., BANSAL, N. & SALUJA, D. 2015. Stress-mediated Sin3B activation leads to negative regulation of subset of p53 target genes. *Biosci Rep*, 35.
- KAJI, K., CABALLERO, I. M., MACLEOD, R., NICHOLS, J., WILSON, V. A. & HENDRICH, B. 2006. The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nat Cell Biol*, 8, 285-292.
- KATO, S., ISHII, T. & KOUZMENKO, A. 2015. Point mutations in an epigenetic factor lead to multiple types of bone tumors: role of H3.3 histone variant in bone development and disease. *Bonekey Rep*, 4, 715.
- KELLY, R. D. & COWLEY, S. M. 2013. The physiological roles of histone deacetylase (HDAC) 1 and 2: complex co-stars with multiple leading parts. *Biochem Soc Trans*, 41, 741-9.
-

- KHOUEIRY, R., SOHNI, A., THIENPONT, B., LUO, X., VELDE, J. V., BARTOCETTI, M., BOECKX, B., ZWIJSEN, A., RAO, A., LAMBRECHTS, D. & KOH, K. P. 2017. Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nat Genet*, 49, 1061-1072.
- KIM, J., DANIEL, J., ESPEJO, A., LAKE, A., KRISHNA, M., XIA, L., ZHANG, Y. & BEDFORD, M. T. 2006. Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep*, 7, 397-403.
- KIM, S. T., LIM, D. S., CANMAN, C. E. & KASTAN, M. B. 1999. Substrate specificities and identification of putative substrates of ATM kinase family members. *J Biol Chem*, 274, 37538-43.
- KING, M. T. & REISS, P. D. 1985. Separation and measurement of short-chain coenzyme-A compounds in rat liver by reversed-phase high-performance liquid chromatography. *Analytical Biochemistry*, 146, 173-179.
- KINNER, A., WU, W., STAUDT, C. AND ILIAKIS, G. 2008. Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. *Nucleic Acids Research*, 36, 5678-94.
- KO, C. H., BUCKLEY, A. M. & GABER, R. F. 1990. Trk2 Is Required for Low Affinity K(+) Transport in *Saccharomyces Cerevisiae*. *Genetics*, 125, 305-312.
- KO, M., AN, J., BANDUKWALA, H. S., CHAVEZ, L., AIJO, T., PASTOR, W. A., SEGAL, M. F., LI, H., KOH, K. P., LAHDESMAKI, H., HOGAN, P. G., ARAVIND, L. & RAO, A. 2013. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature*, 497, 122-6.
- KOH, K. P., YABUUCHI, A., RAO, S., HUANG, Y., CUNNIFF, K., NARDONE, J., LAIHO, A., TAHILIANI, M., SOMMER, C. A., MOSTOSLAVSKY, G., LAHESMAA, R., ORKIN, S. H., RODIG, S. J., DALEY, G. Q. & RAO, A. 2011. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell*, 8, 200-13.
- KOROLEV, N., LYUBARTSEV, A. P. & NORDENSKIÖLD, L. 2006. Computer Modeling Demonstrates that Electrostatic Attraction of Nucleosomal DNA is Mediated by Histone Tails. *Biophysical Journal*, 90, 4305-4316.
-

- KOSTRHON, S., KONTAXIS, G., KAUFMANN, T., SCHIRGHUBER, E., KUBICEK, S., KONRAT, R. AND SLADE, D. 2017. A histone-mimicking interdomain linker in a multidomain protein modulates multivalent histone binding. *The Journal of Biological Chemistry*, 292, 17643–57.
- KRIUKIENE, E., LIUTKEVICIUTE, Z. & KLIMASAUSKAS, S. 2012. 5-Hydroxymethylcytosine--the elusive epigenetic mark in mammalian DNA. *Chem Soc Rev*, 41, 6916-30.
- KUBOSAKI, A., TOMARU, Y., FURUHATA, E., SUZUKI, T., SHIN, J. W., SIMON, C., ANDO, Y., HASEGAWA, R., HAYASHIZAKI, Y. & SUZUKI, H. 2012. CpG site-specific alteration of hydroxymethylcytosine to methylcytosine beyond DNA replication. *Biochem Biophys Res Commun*, 426, 141-7.
- KULIS, M. & ESTELLER, M. 2010. 2 - DNA Methylation and Cancer. *In: HERCEG, Z. & USHIJIMA, T. (eds.) Advances in Genetics*. Academic Press.
- KUMAR, G. S., XIE, T., ZHANG, Y. & RADHAKRISHNAN, I. 2011. Solution structure of the mSin3A PAH2-Pfl SID1 complex: a Mad1/Mxd1-like interaction disrupted by MRG15 in the Rpd3S/Sin3S complex. *J Mol Biol*, 408, 987-1000.
- KUNTIMADDI, A., ACHILLE, NICHOLAS J., THORPE, J., LOKKEN, ALYSON A., SINGH, R., HEMENWAY, CHARLES S., ADLI, M., ZELEZNIK-LE, NANCY J. & BUSHWELLER, JOHN H. 2015. Degree of Recruitment of DOT1L to MLL-AF9 Defines Level of H3K79 Di- and Tri-methylation on Target Genes and Transformation Potential. *Cell Reports*, 11, 808-820.
- KWON, H. S., KIM, D.-R., YANG, E. G., PARK, Y. K., AHN, H.-C., MIN, S.-J. & AHN, D.-R. 2012. Inhibition of VEGF transcription through blockade of the hypoxia inducible factor-1 α -p300 interaction by a small molecule. *Bioorganic & Medicinal Chemistry Letters*, 22, 5249-5252.
- LAHERTY, C. D., YANG, W. M., SUN, J. M., DAVIE, J. R., SETO, E. & EISENMAN, R. N. 1997. Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell*, 89, 349-56.
- LAHM, A., PAOLINI, C., PALLAORO, M., NARDI, M. C., JONES, P., NEDDERMANN, P., SAMBUCINI, S., BOTTOMLEY, M. J., LO SURDO, P., CARFI, A., KOCH, U., DE FRANCESCO, R., STEINKUHLER, C. &

- GALLINARI, P. 2007. Unraveling the hidden catalytic activity of vertebrate class IIa histone deacetylases. *Proc Natl Acad Sci U S A*, 104, 17335-40.
- LAI, A., KENNEDY, B. K., BARBIE, D. A., BERTOS, N. R., YANG, X. J., THEBERGE, M.-C., TSAI, S.-C., SETO, E., ZHANG, Y., KUZMICHEV, A., LANE, W. S., REINBERG, D., HARLOW, E. & BRANTON, P. E. 2001. RBP1 Recruits the mSIN3-Histone Deacetylase Complex to the Pocket of Retinoblastoma Tumor Suppressor Family Proteins Found in Limited Discrete Regions of the Nucleus at Growth Arrest. *Molecular and Cellular Biology*, 21, 2918-2932.
- LAUKKA, T., MARIANI, C. J., IHANTOLA, T., CAO, J. Z., HOKKANEN, J., KAELIN, W. G., JR., GODLEY, L. A. & KOIVUNEN, P. 2016. Fumarate and Succinate Regulate Expression of Hypoxia-inducible Genes via TET Enzymes. *J Biol Chem*, 291, 4256-65.
- LECHNER, T., CARROZZA, M. J., YU, Y., GRANT, P. A., EBERHARTER, A., VANNIER, D., BROSCHE, G., STILLMAN, D. J., SHORE, D. & WORKMAN, J. L. 2000. Sds3 (Suppressor of Defective Silencing 3) Is an Integral Component of the Yeast Sin3·Rpd3 Histone Deacetylase Complex and Is Required for Histone Deacetylase Activity. *Journal of Biological Chemistry*, 275, 40961-40966.
- LEROY, G., WESTON, J. T., ZEE, B. M., YOUNG, N. L., PLAZAS-MAYORCA, M. D. & GARCIA, B. A. 2009. Heterochromatin Protein 1 Is Extensively Decorated with Histone Code-like Post-translational Modifications. *Molecular & Cellular Proteomics : MCP*, 8, 2432-2442.
- LI, H., HUI, X., LI, P., XU, A., LI, S., JIN, S. & WU, D. 2015. Expression and efficient purification of tag-cleaved active recombinant human insulin-like growth factor-II from Escherichia coli. *Biotechnology and Bioprocess Engineering*, 20, 234-241.
- LI, Y., SABARI, B. R., PANCHENKO, T., WEN, H., ZHAO, D., GUAN, H., WAN, L., HUANG, H., TANG, Z., ZHAO, Y., ROEDER, R. G., SHI, X., ALLIS, C. D. & LI, H. 2016. Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Mol Cell*, 62, 181-93.
-

- LI, Y., WEN, H., XI, Y., TANAKA, K., WANG, H., PENG, D., REN, Y., JIN, Q., DENT, SHARON Y. R., LI, W., LI, H. & SHI, X. 2014. AF9 YEATS Domain Links Histone Acetylation to DOT1L-Mediated H3K79 Methylation. *Cell*, 159, 558-571.
- LI, Y., ZHAO, D., CHEN, Z. & LI, H. 2017. YEATS domain: Linking histone crotonylation to gene regulation. *Transcription*, 8, 9-14.
- LI, Z., CAI, X., CAI, C. L., WANG, J., ZHANG, W., PETERSEN, B. E., YANG, F. C. & XU, M. 2011. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood*, 118, 4509-18.
- LIANG, G., CHAN, M.F., TOMIGAHARA, Y., TSAI, Y.C., GONZALES, F.A., LI, E., LAIRD, P.W., et al. 2002. Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Molecular and Cellular Biology*, 22, 480-91.
- LIANG, J., WAN, M., ZHANG, Y., GU, P., XIN, H., JUNG, S. Y., QIN, J., WONG, J., COONEY, A. J., LIU, D. & SONGYANG, Z. 2008. Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nat Cell Biol*, 10, 731-9.
- LIN, C., GARRETT, A. S., DE KUMAR, B., SMITH, E. R., GOGOL, M., SEIDEL, C., KRUMLAUF, R. & SHILATIFARD, A. 2011. Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC). *Genes & Development*, 25, 1486-1498.
- LIN, H., SU, X. & HE, B. 2012. Protein lysine acylation and cysteine succination by intermediates of energy metabolism. *ACS Chem Biol*, 7, 947-60.
- LIU, B., LIN, Y., DARWANTO, A., SONG, X., XU, G. & ZHANG, K. 2009. *Identification and Characterization of Propionylation at Histone H3 Lysine 23 in Mammalian Cells.*
- LU, H., PISE-MASISON, C. A., FLETCHER, T. M., SCHILTZ, R. L., NAGAICH, A. K., RADONOVICH, M., HAGER, G., COLE, P. A. & BRADY, J. N. 2002. Acetylation of nucleosomal histones by p300 facilitates transcription from tax-responsive human T-cell leukemia virus type 1 chromatin template. *Mol Cell Biol*, 22, 4450-62.
-

- LUDWIG, A. K., ZHANG, P., HASTERT, F. D., MEYER, S., RAUSCH, C., HERCE, H. D., MULLER, U., LEHMKUHL, A., HELLMANN, I., TRUMMER, C., STORM, C., LEONHARDT, H. & CARDOSO, M. C. 2017. Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise. *Nucleic Acids Res*, 45, 2438-2457.
- LUGER, K., MADER, A. W., RICHMOND, R. K., SARGENT, D. F. & RICHMOND, T. J. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389, 251-60.
- LUO, J., LI, M., TANG, Y., LASZKOWSKA, M., ROEDER, R. G. & GU, W. 2004. Acetylation of p53 augments its site-specific DNA binding both in vitro and in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 2259-2264.
- LUO, J., NIKOLAEV, A. Y., IMAI, S., CHEN, D., SU, F., SHILOH, A., GUARENTE, L. & GU, W. 2001. Negative control of p53 by Sir2alpha promotes cell survival under stress. *Cell*, 107, 137-48.
- LUO, J., SU, F., CHEN, D., SHILOH, A. & GU, W. 2000. Deacetylation of p53 modulates its effect on cell growth and apoptosis. *Nature*, 408, 377-81.
- LUO, Y., JIAN, W., STAVREVA, D., FU, X., HAGER, G., BUNGERT, J., HUANG, S. & QIU, Y. 2009. Trans-regulation of Histone Deacetylase Activities through Acetylation. *The Journal of Biological Chemistry*, 284, 34901-34910.
- MAITI, A. & DROHAT, A. C. 2011. Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine: POTENTIAL IMPLICATIONS FOR ACTIVE DEMETHYLATION OF CpG SITES. *The Journal of Biological Chemistry*, 286, 35334-35338.
- MARKS, P. A. & XU, W. S. 2009. Histone deacetylase inhibitors: Potential in cancer therapy. *J Cell Biochem*, 107, 600-8.
- MARMORSTEIN, R. 2001. Structure of histone acetyltransferases. *J Mol Biol*, 311, 433-44.
- MEEHAN, R. R., LEWIS, J. D., MCKAY, S., KLEINER, E. L. & BIRD, A. P. 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, 58, 499-507.
-

- MESELSON, M. & STAHL, F. W. 1958. THE REPLICATION OF DNA IN ESCHERICHIA COLI. *Proceedings of the National Academy of Sciences of the United States of America*, 44, 671-682.
- MICHAN, S. & SINCLAIR, D. 2007. Sirtuins in mammals: insights into their biological function. *The Biochemical journal*, 404, 1-13.
- MIZZEN, C. A., BROWNELL, J. E., COOK, R. G. & ALLIS, C. D. 1999. Histone acetyltransferases: preparation of substrates and assay procedures. *Methods Enzymol*, 304, 675-96.
- MODIS, Y. & WIERENGA, R. 1998. Two crystal structures of N-acetyltransferases reveal a new fold for CoA-dependent enzymes. *Structure*, 6, 1345-1350.
- MOEHREN, U., DRESSEL, U., REEB, C. A., VÄISÄNEN, S., DUNLOP, T. W., CARLBERG, C. & BANIAHMAD, A. 2004b. The highly conserved region of the co-repressor Sin3A functionally interacts with the co-repressor Alien. *Nucleic Acids Research*, 32, 2995-3004.
- MOELLERING, R. E. & CRAVATT, B. F. 2013. Functional lysine modification by an intrinsically reactive primary glycolytic metabolite. *Science*, 341, 549-53.
- MOLARO, A., FALCIATORI, I., HODGES, E., ARAVIN, A. A., MARRAN, K., RAFII, S., MCCOMBIE, W. R., SMITH, A. D. & HANNON, G. J. 2014. Two waves of de novo methylation during mouse germ cell development. *Genes & Development*, 28, 1544-1549.
- MONTELLIER, E., ROUSSEAU, S., ZHAO, Y. & KHOCHBIN, S. 2012. Histone crotonylation specifically marks the haploid male germ cell gene expression program: post-meiotic male-specific gene expression. *Bioessays*, 34, 187-93.
- MOORE, L. D., LE, T. & FAN, G. 2013. DNA Methylation and Its Basic Function. *Neuropsychopharmacology*, 38, 23-38.
- MULLER, S., FILIPPAKOPOULOS, P. & KNAPP, S. 2011. Bromodomains as therapeutic targets. *Expert Reviews in Molecular Medicine*, 13, e29.
- NERI, F., INCARNATO, D., KREPELOVA, A., RAPELLI, S., PAGNANI, A., ZECCHINA, R., PARLATO, C. & OLIVIERO, S. 2013. Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome Biol*, 14, R91.

- NEUWALD, A. F. & LANDSMAN, D. 1997. GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem Sci*, 22, 154-5.
- NILSSON, I., JOHNSON, A. E. & VON HEIJNE, G. 2003. How Hydrophobic Is Alanine? *Journal of Biological Chemistry*, 278, 29389-29393.
- NOMURA, M., UDA-TOCHIO, H., MURAI, K., MORI, N. & NISHIMURA, Y. 2005. The Neural Repressor NRSF/REST Binds the PAH1 Domain of the Sin3 Corepressor by Using its Distinct Short Hydrophobic Helix. *Journal of Molecular Biology*, 354, 903-915.
- OKANO, M., BELL, D. W., HABER, D. A. & LI, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99, 247-57.
- ONO, R., TAKI, T., TAKETANI, T., TANIWAKI, M., KOBAYASHI, H. & HAYASHI, Y. 2002. LCX, Leukemia-associated Protein with a CXXC Domain, Is Fused to MLL in Acute Myeloid Leukemia with Trilineage Dysplasia Having t(10;11)(q22;q23). *Cancer Research*, 62, 4075-4080.
- OOI, S. K., QIU, C., BERNSTEIN, E., LI, K., JIA, D., YANG, Z., ERDJUMENT-BROMAGE, H., TEMPST, P., LIN, S. P., ALLIS, C. D., CHENG, X. & BESTOR, T. H. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, 448, 714-7.
- PAIK, W. K., PAIK, D. C. & KIM, S. 2007. Historical review: the field of protein methylation. *Trends Biochem Sci*, 32, 146-52.
- PAPIZAN, J.B., SINGER, R.A., TSCHEN, S.-I., DHAWAN, S., FRIEL, J.M., HIPKENS, S.B., MAGNUSON, M.A., et al. 2011. Nkx2.2 repressor complex regulates islet β -cell specification and prevents β -to- α -cell reprogramming. *Genes & Development*, 25, 2291-305.
- PASTOR, W. A., ARAVIND, L. & RAO, A. 2013. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol*, 14, 341-56.
- PEDEUX, R., SENGUPTA, S., SHEN, J. C., DEMIDOV, O. N., SAITO, S. I., ONOGI, H., KUMAMOTO, K., WINCOVITCH, S., GARFIELD, S. H., MCMENAMIN, M., NAGASHIMA, M., GROSSMAN, S. R., APPELLA, E. & HARRIS, C. C.

2005. ING2 Regulates the Onset of Replicative Senescence by Induction of p300-Dependent p53 Acetylation. *Molecular and Cellular Biology*, 25, 6639-6648.
- PEDONE, P. V., PIKAART, M.J., CERRATO, F., VERNUCCI, M., UNGARO, P., BRUNI, C.B. AND RICCIO, A. 1999. Role of histone acetylation and DNA methylation in the maintenance of the imprinted expression of the H19 and Igf2 genes. *FEBS Letters*, 458, 45–50.
- PONTING, C. P., BLAKE, D. J., DAVIES, K. E., KENDRICK-JONES, J. & WINDER, S. J. 1996. ZZ and TAZ: new putative zinc fingers in dystrophin and other proteins. *Trends Biochem Sci*, 21, 11-13.
- QIU, Y., LIU, L., ZHAO, C., HAN, C., LI, F., ZHANG, J., WANG, Y., LI, G., MEI, Y., WU, M., WU, J. & SHI, Y. 2012. Combinatorial readout of unmodified H3R2 and acetylated H3K14 by the tandem PHD finger of MOZ reveals a regulatory mechanism for HOXA9 transcription. *Genes & Development*, 26, 1376-1391.
- QIU, Y., ZHAO, Y., BECKER, M., JOHN, S., PAREKH, B. S., HUANG, S., HENDARWANTO, A., MARTINEZ, E. D., CHEN, Y., LU, H., ADKINS, N. L., STAVREVA, D. A., WIENCH, M., GEORGEL, P. T., SCHILTZ, R. L. & HAGER, G. L. 2006. HDAC1 Acetylation Is Linked to Progressive Modulation of Steroid Receptor-Induced Gene Transcription. *Molecular Cell*, 22, 669-679.
- RAVELLI, A. C., VAN DER MEULEN, J. H., MICHELS, R. P., OSMOND, C., BARKER, D. J., HALES, C. N. & BLEKER, O. P. 1998. Glucose tolerance in adults after prenatal exposure to famine. *Lancet*, 351, 173-7.
- RAVELLI, G. P., STEIN, Z. A. & SUSSER, M. W. 1976. Obesity in young men after famine exposure in utero and early infancy. *N Engl J Med*, 295, 349-53.
- RAZ, R., LEE, C. K., CANNIZZARO, L. A., D'EUSTACHIO, P. & LEVY, D. E. 1999. Essential role of STAT3 for embryonic stem cell pluripotency. *Proc Natl Acad Sci U S A*, 96, 2846-51.
- REGO, A., SINCLAIR, P.B., TAO, W., KIREEV, I. AND BELMONT, A.S. 2008. The facultative heterochromatin of the inactive X chromosome has a distinctive condensed ultrastructure. *Journal of Cell*, 121, 1119–27.
-

- RICHMAN, R., CHICOINE, L. G., COLLINI, M. P., COOK, R. G. & ALLIS, C. D. 1988. Micronuclei and the cytoplasm of growing *Tetrahymena* contain a histone acetylase activity which is highly specific for free histone H4. *J Cell Biol*, 106, 1017-26.
- RINGEL, A. E., CIENIEWICZ, A. M., TAVERNA, S. D. & WOLBERGER, C. 2015. Nucleosome competition reveals processive acetylation by the SAGA HAT module. *Proceedings of the National Academy of Sciences of the United States of America*, 112, E5461-E5470.
- RODRIGUEZ, M. S., DESTERRO, J. M., LAIN, S., LANE, D. P. & HAY, R. T. 2000. Multiple C-terminal lysine residues target p53 for ubiquitin-proteasome-mediated degradation. *Mol Cell Biol*, 20, 8458-67.
- ROSEBOOM, T. J., VAN DER MEULEN, J. H., RAVELLI, A. C., OSMOND, C., BARKER, D. J. & BLEKER, O. P. 2001. Effects of prenatal exposure to the Dutch famine on adult disease in later life: an overview. *Mol Cell Endocrinol*, 185, 93-8.
- ROSSETTO, D., AVVAKUMOV, N. AND CÔTÉ, J. 2012. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics*, 7, 1098-108.
- ROUSSEAUX, S. & KHOCHBIN, S. 2015. Histone Acylation beyond Acetylation: Terra Incognita in Chromatin Biology. *Cell J*, 17, 1-6.
- RUDOLPH, T., BEUCH, S. & REUTER, G. 2013. Lysine-specific histone demethylase LSD1 and the dynamic control of chromatin. *Biol Chem*, 394, 1019-28.
- RUIZ-ANDRES, O., SANCHEZ-NINO, M. D., CANNATA-ORTIZ, P., RUIZ-ORTEGA, M., EGIDO, J., ORTIZ, A. & SANZ, A. B. 2016. Histone lysine crotonylation during acute kidney injury in mice. *Dis Model Mech*, 9, 633-45.
- SABARI, B. R., TANG, Z., HUANG, H., YONG-GONZALEZ, V., MOLINA, H., KONG, H. E., DAI, L., SHIMADA, M., CROSS, J. R., ZHAO, Y., ROEDER, R. G. & ALLIS, C. D. 2015. Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol Cell*, 58, 203-15.
- SABARI, B. R., ZHANG, D., ALLIS, C. D. & ZHAO, Y. 2017. Metabolic regulation of gene expression through histone acylations. *Nat Rev Mol Cell Biol*, 18, 90-101.

- SAHU, S. C., SWANSON, K. A., KANG, R. S., HUANG, K., BRUBAKER, K., RATCLIFF, K. & RADHAKRISHNAN, I. 2008. Conserved themes in target recognition by the PAH1 and PAH2 domains of the Sin3 transcriptional corepressor. *J Mol Biol*, 375, 1444-56.
- SAKSOUK, N., SIMBOECK, E. AND DÉJARDIN, J. 2015. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin*, 8, 3.
- SANCHEZ, R. & ZHOU, M. M. 2009. The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Devel*, 12, 659-65.
- SANTOS, F., HENDRICH, B., REIK, W. & DEAN, W. 2002. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol*, 241, 172-82.
- SAUNDERS, A., HUANG, X., FIDALGO, M., REIMER, M. H., JR., FAIOLA, F., DING, J., SANCHEZ-PRIEGO, C., GUALLAR, D., SAENZ, C., LI, D. & WANG, J. 2017. The SIN3A/HDAC Corepressor Complex Functionally Cooperates with NANOG to Promote Pluripotency. *Cell Rep*, 18, 1713-1726.
- SCHALCH, T., DUDA, S., SARGENT, D.F. AND RICHMOND, T.J. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature*, 436 No. 7047, 138-41.
- SCHMITGES, FRANK W., PRUSTY, ARCHANA B., FATY, M., STÜTZER, A., LINGARAJU, GONDICHATNAHALLI M., AIWAZIAN, J., SACK, R., HESS, D., LI, L., ZHOU, S., BUNKER, RICHARD D., WIRTH, U., BOUWMEESTER, T., BAUER, A., LY-HARTIG, N., ZHAO, K., CHAN, H., GU, J., GUT, H., FISCHLE, W., MÜLLER, J. & THOMÄ, NICOLAS H. 2011. Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Molecular Cell*, 42, 330-341.
- SCHULZE, J. M., WANG, A. Y. & KOBOR, M. S. 2009. YEATS domain proteins: a diverse family with many links to chromatin modification and transcription. *Biochem Cell Biol*, 87, 65-75.
- SEALY, L. & CHALKLEY, R. 1978. The effect of sodium butyrate on histone modification. *Cell*, 14, 115-21.

- SETO, E. & YOSHIDA, M. 2014. Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb Perspect Biol*, 6, a018713.
- SHEEBA, C. J., PALMEIRIM, I. & ANDRADE, R. P. 2007. Chick Hairy1 protein interacts with Sap18, a component of the Sin3/HDAC transcriptional repressor complex. *BMC Dev Biol*, 7, 83.
- SHI, Y., LAN, F., MATSON, C., MULLIGAN, P., WHETSTINE, J. R., COLE, P. A., CASERO, R. A. & SHI, Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, 119, 941-53.
- SHI, Y. J., MATSON, C., LAN, F., IWASE, S., BABA, T. & SHI, Y. 2005. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell*, 19, 857-64.
- SHIIO, Y., ROSE, D. W., AUR, R., DONOHOE, S., AEBERSOLD, R. & EISENMAN, R. N. 2006. Identification and characterization of SAP25, a novel component of the mSin3 corepressor complex. *Mol Cell Biol*, 26, 1386-97.
- SHOGREN-KNAAK, M., ISHII, H., SUN, J. M., PAZIN, M. J., DAVIE, J. R. & PETERSON, C. L. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science*, 311, 844-7.
- SINGER-SAM, J. & RIGGS, A. D. 1993. X chromosome inactivation and DNA methylation. In: JOST, J.-P. & SALUZ, H.-P. (eds.) *DNA Methylation: Molecular Biology and Biological Significance*. Basel: Birkhäuser Basel.
- SMITH, J. S. & BOEKE, J. D. 1997. An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev*, 11, 241-54.
- SONG, C. X., SZULWACH, K. E., DAI, Q., FU, Y., MAO, S. Q., LIN, L., STREET, C., LI, Y., POIDEVIN, M., WU, H., GAO, J., LIU, P., LI, L., XU, G. L., JIN, P. & HE, C. 2013. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell*, 153, 678-91.
- SORMANNI, P., CAMILLONI, C., FARISELLI, P. & VENDRUSCOLO, M. 2015. The s2D method: simultaneous sequence-based prediction of the statistical populations of ordered and disordered regions in proteins. *J Mol Biol*, 427, 982-96.

- STEIN, R., GRUENBAUM, Y., POLLACK, Y., RAZIN, A. & CEDAR, H. 1982. Clonal inheritance of the pattern of DNA methylation in mouse cells. *Proc Natl Acad Sci U S A*, 79, 61-5.
- STERNBERG, P. W., STERN, M. J., CLARK, I. & HERSKOWITZ, I. 1986. Activation of the yeast HO gene by release from multiple negative controls. *Cell*, 48, 567-577.
- STERNER, D. E. & BERGER, S. L. 2000. Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev*, 64, 435-59.
- STREUBEL, G., FITZPATRICK, D. J., OLIVIERO, G., SCELFO, A., MORAN, B., DAS, S., MUNAWAR, N., WATSON, A., WYNNE, K., NEGRI, G. L., DILLON, E. T., JAMMULA, S., HOKAMP, K., O'CONNOR, D. P., PASINI, D., CAGNEY, G. & BRACKEN, A. P. 2017. Fam60a defines a variant Sin3a-Hdac complex in embryonic stem cells required for self-renewal. *EMBO J*, 36, 2216-2232.
- SUN, H., LIU, X., LI, F., LI, W., ZHANG, J., XIAO, Z., SHEN, L., LI, Y., WANG, F. & YANG, J. 2017. First comprehensive proteome analysis of lysine crotonylation in seedling leaves of *Nicotiana tabacum*. *Scientific Reports*, 7, 3013.
- SWANSON, K. A., KNOEPFLER, P. S., HUANG, K., KANG, R. S., COWLEY, S. M., LAHERTY, C. D., EISENMAN, R. N. & RADHAKRISHNAN, I. 2004. HBP1 and Mad1 repressors bind the Sin3 corepressor PAH2 domain with opposite helical orientations. *Nat Struct Mol Biol*, 11, 738-46.
- TAN, L. & SHI, Y. G. 2012. Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development*, 139, 1895-902.
- TAN, M., LUO, H., LEE, S., JIN, F., YANG, JEONG S., MONTELLIER, E., BUCHOU, T., CHENG, Z., ROUSSEAU, S., RAJAGOPAL, N., LU, Z., YE, Z., ZHU, Q., WYSOCKA, J., YE, Y., KHOCHBIN, S., REN, B. & ZHAO, Y. 2011. Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification. *Cell*, 146, 1016-1028.
- TAN, M., PENG, C., ANDERSON, KRISTIN A., CHHOY, P., XIE, Z., DAI, L., PARK, J., CHEN, Y., HUANG, H., ZHANG, Y., RO, J., WAGNER, GREGORY R., GREEN, MICHELLE F., MADSEN, ANDREAS S.,

- SCHMIESING, J., PETERSON, BRETT S., XU, G., ILKAYEVA, OLGA R., MUEHLBAUER, MICHAEL J., BRAULKE, T., MÜHLHAUSEN, C., BACKOS, DONALD S., OLSEN, CHRISTIAN A., MCGUIRE, PETER J., PLETCHER, SCOTT D., LOMBARD, DAVID B., HIRSCHEY, MATTHEW D. & ZHAO, Y. 2014. Lysine Glutarylation Is a Protein Posttranslational Modification Regulated by SIRT5. *Cell Metabolism*, 19, 605-617.
- TANIGUCHI, Y. 2016. The Bromodomain and Extra-Terminal Domain (BET) Family: Functional Anatomy of BET Paralogous Proteins. *International Journal of Molecular Sciences*, 17, 1849.
- TAUNTON, J., HASSIG, C. A. & SCHREIBER, S. L. 1996. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science*, 272, 408-11.
- THOMAS, J. O. 1999. Histone H1: location and role. *Curr Opin Cell Biol*, 11, 312-7.
- TINI, M., BENECKE, A., UM, S. J., TORCHIA, J., EVANS, R. M. & CHAMBON, P. 2002. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. *Mol Cell*, 9, 265-77.
- VALENTE, S., LIU, Y., SCHNEKENBURGER, M., ZWERGEL, C., COSCONATI, S., GROS, C., TARDUGNO, M., LABELLA, D., FLOREAN, C., MINDEN, S., HASHIMOTO, H., CHANG, Y., ZHANG, X., KIRSCH, G., NOVELLINO, E., ARIMONDO, P. B., MIELE, E., FERRETTI, E., GULINO, A., DIEDERICH, M., CHENG, X. & MAI, A. 2014. Selective non-nucleoside inhibitors of human DNA methyltransferases active in cancer including in cancer stem cells. *J Med Chem*, 57, 701-13.
- VAN INGEN, H., BALTUSSEN, M. A. H., AELEN, J. & VUISTER, G. W. 2006. Role of Structural and Dynamical Plasticity in Sin3: The Free PAH2 Domain is a Folded Module in mSin3B. *Journal of Molecular Biology*, 358, 485-497.
- VAN INGEN, H., LASONDER, E., JANSEN, J. F., KAAAN, A. M., SPRONK, C. A., STUNNENBERG, H. G. & VUISTER, G. W. 2004. Extension of the binding motif of the Sin3 interacting domain of the Mad family proteins. *Biochemistry*, 43, 46-54.

- VEENENDAAL, M. V., PAINTER, R. C., DE ROOIJ, S. R., BOSSUYT, P. M., VAN DER POST, J. A., GLUCKMAN, P. D., HANSON, M. A. & ROSEBOOM, T. J. 2013. Transgenerational effects of prenatal exposure to the 1944-45 Dutch famine. *BJOG*, 120, 548-53.
- VELLA, P., SCELFO, A., JAMMULA, S., CHIACCHIERA, F., WILLIAMS, K., CUOMO, A., ROBERTO, A., CHRISTENSEN, J., BONALDI, T., HELIN, K. & PASINI, D. 2013. Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. *Mol Cell*, 49, 645-56.
- VERSCHURE, P. J., VAN DER KRAAN, I., DE LEEUW, W., VAN DER VLAG, J., CARPENTER, A. E., BELMONT, A. S. & VAN DRIEL, R. 2005. In Vivo HP1 Targeting Causes Large-Scale Chromatin Condensation and Enhanced Histone Lysine Methylation. *Molecular and Cellular Biology*, 25, 4552-4564.
- VIDAL, M., BUCKLEY, A. M., HILGER, F. & GABER, R. F. 1990. Direct Selection for Mutants with Increased K(+) Transport in *Saccharomyces Cerevisiae*. *Genetics*, 125, 313-320.
- VIDAL, M. & GABER, R. F. 1991. RPD3 encodes a second factor required to achieve maximum positive and negative transcriptional states in *Saccharomyces cerevisiae*. *Mol Cell Biol*, 11, 6317-27.
- VOLLMUTH, F. & GEYER, M. 2010. Interaction of Propionylated and Butyrylated Histone H3 Lysine Marks with Brd4 Bromodomains. *Angewandte Chemie*, 122, 6920-6924.
- WAGNER, C., DIETZ, M., WITTMANN, J., ALBRECHT, A. & SCHÜLLER, H.-J. 2001. The negative regulator Opi1 of phospholipid biosynthesis in yeast contacts the pleiotropic repressor Sin3 and the transcriptional activator Ino2. *Molecular Microbiology*, 41, 155-166.
- WANG, H., YANG, H., SHIVALILA, C. S., DAWLATY, M. M., CHENG, A. W., ZHANG, F. & JAENISCH, R. 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, 153, 910-8.
- WANG, R., CHERUKURI, P. & LUO, J. 2005. Activation of Stat3 sequence-specific DNA binding and transcription by p300/CREB-binding protein-mediated acetylation. *J Biol Chem*, 280, 11528-34.

- WANG, Y.-L., FAIOLA, F., XU, M., PAN, S. & MARTINEZ, E. 2008. Human ATAC Is a GCN5/PCAF-containing Acetylase Complex with a Novel NC2-like Histone Fold Module That Interacts with the TATA-binding Protein. *The Journal of Biological Chemistry*, 283, 33808-33815.
- WEI, W., MAO, A., TANG, B., ZENG, Q., GAO, S., LIU, X., LU, L., LI, W., DU, J. X., LI, J., WONG, J. & LIAO, L. 2017. Large-Scale Identification of Protein Crotonylation Reveals Its Role in Multiple Cellular Functions. *Journal of Proteome Research*, 16, 1743-1752.
- WIEHLE, L., RADDATZ, G., MUSCH, T., DAWLATY, M. M., JAENISCH, R., LYKO, F. & BREILING, A. 2015. Tet1 and Tet2 Protect DNA Methylation Canyons against Hypermethylation. *Mol Cell Biol*, 36, 452-61.
- WILLIAMS, K., CHRISTENSEN, J., PEDERSEN, M. T., JOHANSEN, J. V., CLOOS, P. A., RAPPSILBER, J. & HELIN, K. 2011. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*, 473, 343-8.
- WIŚNIEWSKI, J. R., ZOUGMAN, A. & MANN, M. 2008. N(ε)-Formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. *Nucleic Acids Research*, 36, 570-577.
- WOOD, A.J. AND OAKEY, R.J. 2006. Genomic imprinting in mammals: emerging themes and established theories. *PLoS Genetics*, 2, e147.
- WU, H., CHEN, X., XIONG, J., LI, Y., LI, H., DING, X., LIU, S., CHEN, S., GAO, S. & ZHU, B. 2011a. Histone methyltransferase G9a contributes to H3K27 methylation in vivo. *Cell Res*, 21, 365-7.
- WU, H., D'ALESSIO, A. C., ITO, S., XIA, K., WANG, Z., CUI, K., ZHAO, K., SUN, Y. E. & ZHANG, Y. 2011b. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature*, 473, 389-93.
- WU, H. & ZHANG, Y. 2011. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev*, 25, 2436-52.
- WU, Q., LI, W., WANG, C., FAN, P., CAO, L., WU, Z. & WANG, F. 2017. Ultradeep Lysine Crotonylome Reveals the Crotonylation Enhancement on Both Histones and Nonhistone Proteins by SAHA Treatment. *Journal of Proteome Research*.
-

- WU, T. P., WANG, T., SEETIN, M. G., LAI, Y., ZHU, S., LIN, K., LIU, Y., BYRUM, S. D., MACKINTOSH, S. G., ZHONG, M., TACKETT, A., WANG, G., HON, L. S., FANG, G., SWENBERG, J. A. & XIAO, A. Z. 2016. DNA methylation on N(6)-adenine in mammalian embryonic stem cells. *Nature*, 532, 329-33.
- XIAO, H., XUAN, W., SHAO, S., LIU, T. & SCHULTZ, P. G. 2015. Genetic Incorporation of ϵ -N-2-Hydroxyisobutyryl-lysine into Recombinant Histones. *ACS Chemical Biology*, 10, 1599-1603.
- XIE, T., HE, Y., KORKEAMAKI, H., ZHANG, Y., IMHOFF, R., LOHI, O. & RADHAKRISHNAN, I. 2011. Structure of the 30-kDa Sin3-associated protein (SAP30) in complex with the mammalian Sin3A corepressor and its role in nucleic acid binding. *J Biol Chem*, 286, 27814-24.
- XIE, Z., DAI, J., DAI, L., TAN, M., CHENG, Z., WU, Y., BOEKE, J. D. & ZHAO, Y. 2012. Lysine succinylation and lysine malonylation in histones. *Mol Cell Proteomics*, 11, 100-7.
- XIONG, J., ZHANG, Z., CHEN, J., HUANG, H., XU, Y., DING, X., ZHENG, Y., NISHINAKAMURA, R., XU, G.-L., WANG, H., CHEN, S., GAO, S. & ZHU, B. 2016. Cooperative Action between SALL4A and TET Proteins in Stepwise Oxidation of 5-Methylcytosine. *Molecular Cell*, 64, 913-925.
- XU, W., WAN, J., ZHAN, J., LI, X., HE, H., SHI, Z. & ZHANG, H. 2017. Global profiling of crotonylation on non-histone proteins. *Cell Res*, 27, 946-949.
- YANG, C., WU, J. & ZHENG, Y. G. 2012. Function of the active site lysine autoacetylation in Tip60 catalysis. *PLoS One*, 7, e32886.
- YANG, P., WANG, Y., CHEN, J., LI, H., KANG, L., ZHANG, Y., CHEN, S., ZHU, B. & GAO, S. 2011. RCOR2 Is a Subunit of the LSD1 Complex That Regulates ESC Property and Substitutes for SOX2 in Reprogramming Somatic Cells to Pluripotency. *STEM CELLS*, 29, 791-801.
- YANG, X. 2017. *Protein Expression Laboratory (Protex), Cloning Service - Available Vectors* [Online]. University of Leicester: Department of Biochemistry. Available: <https://www2.le.ac.uk/departments/molcellbiol/facilities/protex> [Accessed 22-10-2017 2017].
-

- YANG, X., ZHANG, F. & KUDLOW, J. E. 2002. Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein O-GlcNAcylation to transcriptional repression. *Cell*, 110, 69-80.
- YANG, X. J. & GREGOIRE, S. 2005. Class II histone deacetylases: from sequence to function, regulation, and clinical implication. *Mol Cell Biol*, 25, 2873-84.
- YODER, J. A., WALSH, C. P. & BESTOR, T. H. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, 13, 335-340.
- YTHIER, D., LARRIEU, D., BINET, R., BINDA, O., BRAMBILLA, C., GAZZERI, S. & PEDEUX, R. 2010. Sumoylation of ING2 regulates the transcription mediated by Sin3A. *Oncogene*, 29, 5946-5956.
- YU, J. & AUWERX, J. 2009. The role of sirtuins in the control of metabolic homeostasis. *Ann N Y Acad Sci*, 1173 Suppl 1, E10-9.
- YUAN, H., ROSSETTO, D., MELLERT, H., DANG, W., SRINIVASAN, M., JOHNSON, J., HODAWADEKAR, S., DING, E. C., SPEICHER, K., ABSHIRU, N., PERRY, R., WU, J., YANG, C., ZHENG, Y. G., SPEICHER, D. W., THIBAUT, P., VERREAULT, A., JOHNSON, F. B., BERGER, S. L., STERNGLANZ, R., MCMAHON, S. B., COTE, J. & MARMORSTEIN, R. 2012. MYST protein acetyltransferase activity requires active site lysine autoacetylation. *EMBO J*, 31, 58-70.
- YUAN, L. W. & GIORDANO, A. 2002. Acetyltransferase machinery conserved in p300/CBP-family proteins. *Oncogene*, 21, 2253-60.
- YUAN, Z. L., GUAN, Y. J., CHATTERJEE, D. & CHIN, Y. E. 2005. Stat3 dimerization regulated by reversible acetylation of a single lysine residue. *Science*, 307, 269-73.
- ZALENSKY, A.O., SIINO, J.S., GINEITIS, A.A., ZALENSKAYA, I.A., TOMILIN, N. V, YAU, P. AND BRADBURY, E.M. 2002. Human testis/sperm-specific histone H2B (hTSH2B). Molecular cloning and characterization. *The Journal of Biological Chemistry*, 277, 43474-80.
- ZENG, L., ZHANG, Q., LI, S., PLOTNIKOV, A. N., WALSH, M. J. & ZHOU, M. M. 2010. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature*, 466, 258-62.

- ZHANG, K., CHEN, Y., ZHANG, Z. & ZHAO, Y. 2009. Identification and Verification of Lysine Propionylation and Butyrylation in Yeast Core Histones Using PTMap Software. *Journal of proteome research*, 8, 900-906.
- ZHANG, Y., NG, H. H., ERDJUMENT-BROMAGE, H., TEMPST, P., BIRD, A. & REINBERG, D. 1999. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev*, 13, 1924-35.
- ZHANG, Y. & REINBERG, D. 2001. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & Development*, 15, 2343-2360.
- ZHANG, Y., SUN, Z.-W., IRATNI, R., ERDJUMENT-BROMAGE, H., TEMPST, P., HAMPSEY, M. & REINBERG, D. 1998. SAP30, a Novel Protein Conserved between Human and Yeast, Is a Component of a Histone Deacetylase Complex. *Molecular Cell*, 1, 1021-1031.
- ZHANG, Y. W., WANG, Z., XIE, W., CAI, Y., XIA, L., EASWARAN, H., LUO, J., YEN, R. C., LI, Y. & BAYLIN, S. B. 2017. Acetylation Enhances TET2 Function in Protecting against Abnormal DNA Methylation during Oxidative Stress. *Mol Cell*, 65, 323-335.
- ZHAO, B., YE, X., YU, J., LI, L., LI, W., LI, S., YU, J., LIN, J. D., WANG, C. Y., CHINNAIYAN, A. M., LAI, Z. C. & GUAN, K. L. 2008. TEAD mediates YAP-dependent gene induction and growth control. *Genes Dev*, 22, 1962-71.
- ZHENG, C. & HAYES, J. J. 2003. Structures and interactions of the core histone tail domains. *Biopolymers*, 68, 539-46.
- ZHONG, J., LI, X., CAI, W., WANG, Y., DONG, S., YANG, J., ZHANG, J., WU, N., LI, Y., MAO, F., ZENG, C., WU, J., XU, X. & SUN, Z. S. 2017. TET1 modulates H4K16 acetylation by controlling auto-acetylation of hMOF to affect gene regulation and DNA repair function. *Nucleic Acids Res*, 45, 672-684.

Thank you.



Yang, Xiaowen (Dr.) sent an automatic reply.



Yang, Xiaowen (Dr.)



Thank you very much for your email. I am off office and I will reply you when I come back office. If you have any issues regarding Protex, please contact Dipti Vashi @ dv47@leicester.ac.uk.



Chandru, Aditya

Today, 14:18

Yang, Xiaowen (Dr.)



Reply



I am grateful that providence arranged for us to meet at this singular point, throughout the vastness of both space and time, so that I could share in your friendship for this brief moment.