The Pennsylvania State University

The Graduate School

Eberly College of Science

# STRUCTURAL AND BIOCHEMICAL CHARACTERIZATION OF LSD1/COREST HISTONE DEMETHYLASE COMPLEX ON THE NUCLEOSOME

A Dissertation in

Biochemistry, Microbiology, and Molecular Biology

by

Sang-Ah Kim

© 2016 Sang-Ah Kim

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

December 2016

The dissertation of Sang-Ah Kim was reviewed and approved\* by the following:

Song Tan Professor of Biochemistry and Molecular Biology Dissertation Advisor Chair of Committee

Lu Bai Assistant Professor of Biochemistry and Molecular Biology Assistant Professor of Physics

David Gilmour Professor of Molecular and Cell Biology Graduate Education Co-Director

Andrey Krasilnikov Associate Professor of Biochemistry and Molecular Biology

Scott A. Showalter Associate Professor of Chemistry Associate Professor of Biochemistry and Molecular Biology

Scott B. Selleck Professor and Head, Department of Biochemistry and Molecular Biology

\*Signatures are on file in the Graduate School

# ABSTRACT

Chromatin, a complex of DNA and proteins, is a dynamic structure that can respond to external cues to regulate genes. Posttranslational modification of histones plays a key role in this regulation through modulation of chromatin structure and function. One such modification, methylation of histone H3 lysine 4 (H3K4), has been correlated with transcriptional activation and is enriched at the promoters and enhancers of active genes. Lysine residues can be mono-, dior trimethylated, and the specific form of methylation is regulated by a synergistic effort of histone methyltransferases and demethylases. The LSD1 (lysine-specific demethylase 1) protein can alone demethylate mono- and dimethylated H3K4 in peptide substrates, but requires the corepressor protein, CoREST, to demethylate nucleosome substrates. The crystal structure of the LSD1/CoREST in complex with the H3-peptide has been solved and a plethora of information exists on how the LSD1/CoREST works on the peptide substrates. However, these studies do not provide insight into how LSD1/CoREST interacts with the nucleosome, the physiological substrate. As such, full elucidation of LSD1/CoREST's function on the nucleosome is critical to understanding its roles in development, cellular differentiation, embryonic pluripotency, and cancer.

In this dissertation, I provide insights into how the LSD1/CoREST complex interacts with the nucleosome via biochemical and structural approaches. My studies of LSD1/CoREST's enzymatic activity and nucleosome binding show that extranucleosomal DNA dramatically enhances the activity of LSD1/CoREST and that LSD1/CoREST requires DNA beyond the nucleosome core particle for optimal interaction. These results are consistent with the photocrosslinking experiments performed by Nilanjana Chatterjee in collaboration with Blaine Bartholomew's group, which indicate both LSD1 and CoREST subunits are in close contact with extranucleosomal DNA. I have also determined the molecular replacement solutions of

LSD1/CoREST on 161 and 189 bp nucleosome crystal structures at 7.2 Å and 5.3 Å resolution respectively. These studies provide novel insights into how LSD1/CoREST productively engages the nucleosomes with extranucleosomal DNA. Both models suggest that the catalytic domain of LSD1 and the N-terminus of CoREST makes contact with extranucleosomal DNA, while the CoREST SANT2 domain recognizes the nucleosomal disk face. Current work is aimed toward determining the LSD1/CoREST/nucleosome crystal structure at a higher resolution.

# **TABLE OF CONTENTS**

List of Figures	viii
List of Tables	xii
List of Abbreviations	xiii
Acknowledgements	XV
Chapter 1 Introduction	1
1.1 Structure and Function of Chromatin	1
1.1.2 Destteorelational modifications of historica	1 1
1.1.2 Posturalistational modifications of histories	4
1.2 Histone Mathylation Mathyltransfarasas, and Demathylasas	0 0
1.2 Histone Methylation, Methylatisterases, and Demethylases	ð Q
1.2.1 Instone methylation and demethylation	
1.2.2 Lysine methylation and demethylation	
1 3 I SD1/CoREST Complex	13
1 3 1 Biological roles of LSD1	13
1 3 2 Activity of LSD1/CoREST complex	13
1.3.3 Structure of LSD1/CoREST complex	
1.3.4 Substrate binding site of LSD1/CoREST complex	
1.3.5 LSD1 as a therapeutic target	
1.4 References	20
Chapter 2 Materials and Methods	
2.1 Preparation of LSD1/CoREST complexes	29
2.1.1 Cloning and site-directed mutagenesis	29
2.1.2 Co-expression of LSD1/CoREST complexes	32
2.1.3 Purification of LSD1/CoREST complexes	32
2.2 Nucleosome Preparation	35
2.2.1 Expression of recombinant histones	35
2.2.2 Purification of recombinant histones	35
2.2.3 Alkylation of histone H3	
2.2.4 Fluorescent labeling of recombinant histones	
2.2.5 Preparation of nucleosomal DNAs	
2.2.6 Nucleosome reconstitution and purification	41
2.3 Biochemical Characterization of LSD1/CoREST/Nucleosome Complex	
2.3.1 Amplex UltraRed demethylase assay	
2.3.2 Demethylase assay by Western Blot	
2.3.3 HI-FI nucleosome binding assay	
2.3.4 HI-FI stoichiometry assay	
2.3.5 Photocrosslinking assay	
2.5.6 Ubiquityiation assay	
2.4 Structural Characterization of LSD1/CoREST/Nucleosome Complex	50

2.4.1 Reconstitution of LSD1/CoREST/nucleosome complex	50
2.4.2 Crystallization of LSD1/CoREST/nucleosome complex	51
2.4.3 Postcrystallization treatment	51
2.4.4 Data collection and structural determination	52
2.5 References	53
Chapter 3 Extranucleosomal DNA enhances the activity of the LSD1/CoREST histone	
demethylase complex	55
	~ ~
3.1 Abstract	
3.2 Introduction	
5.5 Kesulls	
3.3.1 Steady state kinetic activity suggest a role for extranucleosomal DNA	
3.3.2 LSD1/COREST binds more tightly to nucleosomes with extranucleosomal	(1
DINA	01
3.3.3 LSD1/COREST prefers nucleosomes with extranucleosomal DNA on both	<b>7</b>
sides of the nucleosome.	6/
3.3.4 Stoichiometry of LSDI/CoREST binding to the nucleosome	68
3.3.5 Photocrosslinking studies of LSDI/CoREST on nucleosome	69
3.3.6 LSD1/CoREST does not employ an arginine anchor to interact with the	71
nucleosome acidic patch	/1
3.4 Discussion	/4
3.5 Supplementary Figures.	//
3.6 Acknowledgements	/8
3.7 References	79
Chapter 4 Structural Characterization of I SD1/CoPEST/Nucleosome Complex by X ray	
Crystallography	82
Crystanography	02
4.1 Abstract	82
4.2 Introduction	84
4.3 Results	87
4.3.1 Reconstitution of LSD1/CoREST/nucleosome complex	87
4.3.2 Crystallization and structural determination of	
LSD1/CoREST/nucleosome complex	94
4.3.3 The Dinucleosomes substrate model of LSD1/CoREST/nucleosome	98
4.3.4 The mononucleosome substrate model of LSD1/CoREST/nucleosome	108
4.4 Discussion	
4.5 Supplementary Figures	128
4.6 Supplementary Table	130
4.7 Acknowledgements	131
4.8 References	132
Chapter 5 Biochemical Characterization of LSD1/CoREST on nucleosomes	135
5.1 Abstract	125
5.2 Introduction	.133
5.2 Deculte	120
J.J RESUIS	138

5.3.1 Demethylase activity of LSD1/CoREST complex on mono- and	
dinucleosomes	138
5.3.2 Binding studies of LSD1/CoREST complex on mono- and dinucleosome.	141
5.3.3 Binding studies of LSD1/CoREST mutant complexes	145
5.4 Discussion	156
5.5 Acknowledgements	161
5.6 References	162
Chapter 6 Summary and Future Experiments	164
6.1 Summary	164
6.1.1 Importance of extranucleosomal DNA for LSD1/CoREST	164
6.1.2 Low resolution structure of LSD1/CoREST on nucleosomes	165
6.2 Future Experiments	167
6.2.1 Crystallization of LSD1/CoREST/nucleosomes for higher a resolution	
structure	167
6.2.2. Further validation of current models of LSD1/CoREST/nucleosome	
structures	169
6.2.3 Structural studies of LSD1/CoREST/nucleosome complex using electron	
cryomicroscopy	170
6.3 Supplementary Data	171
6.4 Acknowledgements	172
6.5 References	173

# LIST OF FIGURES

Figure 1-1. Secondary structure of histone proteins and histone-fold of H32
Figure 1-2. Overview of the nucleosome core particle structure
Figure 1-3. Histone-fold heterodimers and octamer
Figure 1-4. Nucleosome core particle surface and interactions7
Figure 1-5. Recognition of nucleosomes by chromatin factors using the acidic patch arginine anchor
Figure 1-6. The distinct regulatory levels of chromatin formed by histone lysine methylation
Figure 1-7. Features of the LSD family
Figure 1-8. Mechanism, domain organization, and structure of LSD1/CoREST complex15
Figure 1-9. Molecular dynamics simulation reveals molecular flexibility of the LSD1/CoRSET complex in solution
Figure 1-10. Histone peptide-binding site in LSD119
Figure 2-1. Schematic representation of pST44 polycistronic expression system30
Figure 2-2. Purification of the LSD1/CoREST complex
Figure 3-1. Kinetic analysis of LSD1/CoREST demethylase on nucleosome substrates60
Figure 3-2. Fluorescence positions on the nucleosome affected by LSD1/CoREST binding
Figure 3-3. Fluorescence change for labels on H2A T10C, H2B K20C, H3 A21C and H3 K27C at different salt concentration
Figure 3-4. Extranucleosomal DNA increases nucleosome binding affinity of LSD1/CoREST
Figure 3-5. Effect of extranucleosomal DNA on one side of the nucleosome on binding affinity to LSD1/CoREST and stoichiometry of LSD1/CoREST binding to the nucleosome
Figure 3-6. Photocrosslinking of LSD1/CoREST to site-specifically labeled nucleosomes71
Figure 3-7. Photocrosslinking results mapped onto the structure of the nucleosome core particle
Supplementary Figure 3-1. Dissociation constants for nucleosomes labeled on H3 A21C77

Figure 4-1. Currently existing hypothesized models of LSD1/CoREST on the nucleosome
Figure 4-2. Effect of deleting the residues in N- or C-terminus of CoREST on the binding affinity for nucleosomes
Figure 4-3. Effect of mutating Lys4 to Met4 in histone H3 in the nucleosome on the binding affinity to LSD1/CoREST
Figure 4-4. Reconstitution of LSD1/CoREST/nucleosome at 100 mM and 75 mM NaCl91
Figure 4-5. Initial and optimized crystals of LSD1/CoREST complexed with different nucleosomes
Figure 4-6. Crystals of LSD1/CoREST/nucleosome and diffraction images
Figure 4-7. Dinucleosome model based on the molecular replacement solution of 161 bp (8N8) nucleosomes in the crystal structure of LSD1/CoREST/nucleosome complex98
Figure 4-8. Overview of the LSD1/CoREST/nucleosome structure in the dinculosome substrate model from the molecular replacement solution
Figure 4-9. Two LSD1/CoREST molecules homodimerize using three contact points101
Figure 4-10. The AOD of LSD1 binds DNA near the end of nucleosomal DNA of the nucleosome 1
Figure 4-11. The N-terminus of CoRESTA1 binds the linker DNA102
Figure 4-12. The 'basic stripes' of LSD1/CoREST homodimer are aligned with end of the nucleosomal DNA and linker DNA
Figure 4-13. The CoREST SANT2 is in close proximity to histone H4 surface and nucleosomal DNA
Figure 4-14. Distance between LSD1 substrate binding site and its target in the dinucleosome substrate model
Figure 4-15. Overview of the LSD1/CoREST/nucleosome structure in the mononucleosome substrate model from the molecular replacement solution
Figure 4-16. Model for 2:1 LSD1/CoREST:nucleosome complex110
Figure 4-17. Comparison of the di- and mononucleosome substrate models111
Figure 4-18. LSD1/CoREST undergoes a conformational change when binding to the nucleosome
Figure 4-19. The CoREST SANT domain is in close proximity to the histone face and nucleosomal DNA

Figure 4-20. Sequence alignment of CoREST SANT domains from four orthologs: human, zebrafish, <i>Drosophila</i> , and <i>Xenopus</i>	115
Figure 4-21. LSD1/CoREST complex makes extensive interactions with extranucleosomal DNA.	116
Figure 4-22. Distance between LSD1 substrate binding site and its target in the mononucleosome substrate model .	118
Figure 4-23. Alignment of v-Myb and CoREST SANT domains.	122
Figure 4-24. Alignment of SANT2 domain of LSD1/CoREST complex with SANT domains in ISW1a and Chd1	123
Figure 4-25. Fluorescence positions on the nucleosome affected by LSD1/CoREST binding mapped onto the mononucleosome substrate model	125
Figure 4-26. Photocrosslinking results mapped onto the mononucleosome substrate model	127
Supplementary Figure 4-1. Binding curves for HI-FI experiments performed with the site-specifically labeled LSD1/CoREST complex	128
Supplementary Figure 4-2. Electron density maps for LSD1/CoREST/161M nucleosome and LSD1/CoREST/189M nucleosome data	129
Figure 5-1. Demethylase activity of LSD1/CoREST on mono- and dinucleosome substrates using Western Blot with an antibody against H3K4me2.	139
Figure 5-2. Ubiquitylation assay on mono- and dinucleosomes.	140
Figure 5-3. Binding affinity of LSD1/CoREST towards dinucleosomes.	142
Figure 5-4. Effect of the dimerization mutations on the binding affinity of LSD1/CoREST towards mono- and dinucleosomes.	145
Figure 5-5. Effect of mutations in the CoREST SANT2 on the binding affinity of LSD1/CoREST towards nucleosomes.	147
Figure 5-6. Effect of N-terminus truncation in CoREST on the binding affinity of LSD1/CoREST towards nucleosomes with different length linker DNA	150
Figure 5-7. Effect of mutations in the LSD1 AOD on the binding affinity of LSD1/CoREST towards nucleosomes with different length linker DNA	152
Figure 5-8. Sequences of histone H3 representing the deletions in the H3 tail.	155
Figure 5-9. Demethylase activity of LSD1/CoREST on 197 bp nucleosomes containing H3 tail truncations.	155

Figure 5-10. The LSD1/CoREST complex from the mononucleosome substrate model superimposed onto the dinucleosome substrate model.	158
Figure 6-1. Recent crystals of LSD1/CoREST/nucleosomes	168
Supplementary Figure 6-1. Binding curves for LSD1/CoREST interacting with nucleosomes containing tailless dimers .	171

# LIST OF TABLES

Table 3-1. Kinetic constants for LSD1 complex using nucleosome substrates containing different linker DNA lengths	.61
Table 3-2. Dissociation constants for LSD1 complex binding to nucleosomes with different length linker DNA.	.66
Table 4-1. Dissociation constants for different LSD1/CoREST complex constructs binding to 157 bp (6N6) and 181 bp (18N18) nucleosomes.	88
Table 4-2. Nucleosome substrates used in crystallization with human LSD1/CoREST and their crystallization and diffraction results	93
Table 4-3. Nucleosome substrates used in crystallization with human LSD1/CoREST∆3 and their crystallization and diffraction results	93
Table 4-4. Nucleosome substrates used in crystallization with zebrafish LSD1/CoREST and their crystallization and diffraction results.	93
Supplementary Table 4-1. Data collection statistics determined by XDS, and molecular replacement statistics calculated by Phaser and REFMAC	130
Table 5-1. Dissociation constants and Hill coefficients for LSD1/CoREST complex binding to mono- and dinucleosomes with different length linker DNA	142
Table 5-2. Dissociation constants for dimerization mutant LSD1/CoREST complexes binding to nucleosomes at 100 mM NaCl	146
Table 5-3. Dissociation constants for SANT2 mutant LSD1/CoREST complexes binding to nucleosomes with different length linker DNA.	148
Table 5-4. Dissociation constants for LSD1/CoREST∆4 complex binding to nucleosomes with different length linker DNA	.151
Table 5-5. Dissociation constants for LSD1/CoREST complex containing LSD1 K355E/K357E/K359E mutations binding to nucleosomes with different length linker DNA.	153
Table 5-6. Dissociation constants for LSD1/CoREST complex containing LSD1 K372A         mutation binding to nucleosomes with different length linker DNA	153
Table 6-1. Components of LSD1/CoREST and nucleosome substrates used in recent crystallization attempts and their crystallization results.	169
Supplementary Table 6-1. Dissociation constants for LSD1/CoREST complex binding to 147, 157, and 181 bp nucleosomes containing tailless dimers	171

# LIST OF ABBREVIATIONS

AOD	Amine oxidase domain
APB	<i>p</i> -azido phenacyl bromide
BSA	Bovine serum albumin
CENP-C	Centromere protein C
Chd1	Chromodomain helices DNA-binding protein 1
CoREST	REST co-repressor protein
DTT	Dithiothretol
DUB	Deubiquitylase
EDTA	Ethylenediamine tetraacetic acid
EM	Electron microscopy
FAD	Flavin adenine dinucleotide
НАТ	Histone acetyltransferase
HDAC	Histone deacetylase complex
HI-FI	High-throughput Interactions by Fluorescence Intensity
HPLC	High-perormance liquid chromatography
HRP	Horseradish peroxidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ISWI	Imitation Switch
LSD	Lysine-specific histone demethylase
LANA	Latency-associated nuclear antigen
NCP	Nucleosome core particle
PTM	Posttranslational modification

PRC	Polycomb repressive complex
PRMT	protein arginine methyltransferase
ТСР	tranylcypromine
SAGA	<u>S</u> pt- <u>A</u> da- <u>G</u> cn5 <u>A</u> cetyltransferase
SAXS	Small-angle X-ray scattering
SET	$\underline{S}u(var)$ 3-9, $\underline{E}$ nhancer of zeste, $\underline{T}$ rithorax
SWIRM	<u>Swi</u> 3, <u>R</u> SC8, and <u>M</u> oira
TEV	Tobacco etch virus

# ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor and mentor, Dr. Song Tan. This dissertation would not have been possible without his continuous support, encouragement, and patience. He has taught me, both consciously and unconsciously, how a great scientist designs and conducts research. His words, "curiosity and attention to detail!" during my first semester in the lab, got stuck with me, and my curiosity and intellectual engagement in research grew progressively over the course of my Ph.D program. It was an honor to complete my Ph.D training under his guidance. I would also like thank the members of my committee, Dr. Lu Bai, Dr. David Gilmour, Dr. Andrey Krasilnikov, and Dr. Scott Showalter for their thoughtful suggestions and guidance throughout my thesis work.

I would like to thank all members of Tan laboratory, past and present, for helpful discussions and constructive criticism for my research. I would like to extend a special acknowledgement to Dr. Robert McGinty. He mentored me from the very first day of my rotation to my fifth year in Tan lab, and guided me through cloning, protein expression and purification, the HI-FI experiments, and crystallization. Beyond that, his ideas and advice made significant contributions to my research and intellectual growth, and I am very grateful to have worked alongside a great scientist. I would like to thank Dr. T.S. Girish, Dr. Jiehuan Huang, and Dr. Matthew Jennings for all the help and support during my research. I also thank our current and former members of the "nucleosome factory", Michael Moore, Maxwell Kruse, Kevin Thyne, Mike Doyle, Bryan Tornabene, and Lauren McCarl, for the continuous supply of high-quality nucleosomes. I cannot emphasize enough the importance of their work in my research. I also express appreciation for the undergraduate students in Tan lab, especially Turner Pecen, James Johnstone, and Victoria Spadafora, for their hard work on the LSD1 project. Also special thank you to Matt Baird, for playing the "Fantastic Baby" song at times in the lab, entertaining me both

inside and outside the lab, discussing science, and proof-reading the dissertation. I also express appreciation to the members of the Center of Eukaryotic Gene Regulation for helpful discussions.

Many thanks to my friends who cheered me on throughout my graduate school despite the distance: Khine Thit Lwin, Christina Hui Min Tun, Alex Thet Paing Kyaw, and Sujin Lee. Even though I got to see you guys only once a year or less, it was always fun catching up with each other's life and seeing you guys doing great things in Myanmar, Atlanta, and Korea. I was horrible at staying in touch during my grad school, but you guys were always in my mind. I am also grateful to meet new friend during my journey here at Penn State: Mina Kim, Eunjin Choi, Soomin Koo, Hee Joo Kim, and Ye Jin Park. My life here was enriched by your presence and friendship. You have all helped me grow mentally, emotionally, and spiritually, and I am forever indebted to you for your prayers and encouragements throughout my journey as a Ph.D student. I would also like to thank Big Bang for making all the experiments more enjoyable with awesome music, and mentally healing me with fantastic concerts in New York.

Finally, I would like to thank my parents, Eung Seob Kim and Jungmi Myung, and my oppa, Sangmin Kim. It was a privilege to have parents who provided me with the best education they can for 22 years, regardless of the circumstances. I will never be able to repay all the sacrifices, the endless prayers, and unconditional love you have shown me. 감사하고 사랑합니다. My close friends know how much I brag about my one and only brother, Sangmin oppa. He was actually the one who first recommended me to get a Ph.D, and helped me build a roadmap for my career. He was always there to listen to my frustrations, support me mentally and financially, and encourage me with his faith in me. 오빠가 내 오빠라서 너무 감사해.

Last but not least, I thank God. 미비하고 부족한 저를 언제나 좋은 길로 인도해주신 신실하신 하나님은 짱입니다요. Studying what God has created amazed me to another level.

Gly-Glu-Asn-Glu-Ser-Ile-Ser 1:1. God makes no mistakes.

# Chapter 1

# Introduction

# **1.1 Structure and Function of Chromatin**

### 1.1.1 Nucleosome core particle

The three billion base pairs of the human genome are packaged into the nucleus of a cell, with a diameter of 10 microns (Alberts B, 1994). The ability of the cell to condense the genome while still maintaining access to specific segments of the genome and reacting promptly to the ever-changing environment requires a highly coordinated effort by a number of factors. This coordination and compaction of the eukaryotic genome are achieved in a polymeric complex called chromatin. The fundamental repeating unit of chromatin is the nucleosome, which consists of a nucleosome core, a linker histone (H1 or H5), and DNA ranging from 160 to 240 base pairs (bp) (McGhee and Felsenfeld, 1980). The nucleosome core contains two copies of four core histone proteins, H2A, H2B, H3, and H4, wrapped around with 145-147 bp of DNA. Each nucleosome core is connected to an adjacent nucleosome core by linker DNA, which associates with linker histone. Nucleosome repeat lengths (NRL), a measure of nucleosomal DNA (145-147 bp) plus the linker DNA, from a wide variety of organisms range from  $\sim 150$  to  $\sim 260$  bp (Bradbury, 1989). The genome-wide sequencing of nucleosomal DNA by the ChIP-Seq technique has estimated the NRL of human to be  $\sim 200$  bp (Schones et al., 2008). However, in addition to species variation, the NRLs vary among tissues, cell types, and even within a single cellular genome (Bradbury, 1989).



**Figure 1-1.** Secondary structure of histone proteins and histone-fold of H3. (a) Schemes representing the secondary structure elements of four core histones. (b) Histone-fold motif of H3 shown in cartoon representation (PDB ID: 1KX5). (Figure adapted from McGinty, R.K. and Tan, S. (2014) In Workman, L. J. and Abmayr, M.S. (eds.), *Fundamentals of Chromatin.* Springer New York, New York, NY, pp. 1-28.)



**Figure 1-2.** Overview of the nucleosome core particle structure (PDB ID: 1KX5). (a) Histones are depicted in cartoon representation and DNA is depicted in stick representation with the dyad marked by an arrow. (b) Nucleosome core particle shown in space-filling representation. (Figure from McGinty, R.K. and Tan, S. (2014) In Workman, L. J. and Abmayr, M.S. (eds.), *Fundamentals of Chromatin.* Springer New York, New York, NY, pp. 1-28.)

The core histones of H2A, H2B, H3, and H4 are basic and relatively small, ranging from

102-135 amino acids, and contain alpha-helical regions that form a histone-fold motif, flanked by

N- and C-terminal extensions (Figure 1-1a). These flexible extensions, also called histone tails, are the major sites of posttranslational modifications which act synergistically to regulate gene expression. The histone-fold motif consists of three  $\alpha$ -helices, connected by two loops ( $\alpha$ 1–L1– $\alpha$ 2–L2– $\alpha$ 3) and is contained in all four core histones (Arents and Moudrianakis, 1995) (Figure 1-1b).

Crystal structures of the nucleosome core particle, solved at 2.8 Å and subsequently at 1.9 Å, show 146 and 147 bp respectively, of human  $\alpha$ -satellite DNA wrapped 1.65 turns around the histone octamer in a left-handed superhelix (Luger et al., 1997, Davey et al., 2002) (Figure 1-2). In the central histone octamer, H3 pairs with H4, and H2A pairs with H2B through the histone-fold motif in an antiparallel arrangement as 'handshake motif' to form a heterodimer (Arents and Moudrianakis, 1995) (Figure 1-3a, b). Two H3/H4 dimers forms a (H3/H4)<sub>2</sub> tetramer through a four-helix bundle that is mediated by H3  $\alpha$ 2 and  $\alpha$ 3 helices. Similarly, H4 of tetramer interacts with H2B of H2A/H2B dimer through a four-helix bundle on each side, completing the octamer (Luger et al., 1997) (Figure 1-3c). The strongly positive surface of resulting octamer contributes to binding to nucleosomal DNA, which is highly negative due to its phosphate backbone. Consequently, the nucleosome core particle has a pseudo-symmetry centered on the dyad, with a diameter of 100 Å and width of 60 Å (Figure 1-2). The structure shows that each histone-fold organizes about 30 bp of DNA, with H3/H4 tetramer interacting with about 60-70 bp of DNA, and each H2A/H2B dimer interacting with about 30 bp of DNA. H4 and H2A Nterminal tails exit the nucleosome core particle on top of the minor groove of the DNA, and H3 and H2B tails exit through the DNA gyres. N-terminal tails of H3 protrude near the entry/exit site near the dyad, whereas N-terminal tails of H2A and H2B exit from the opposite side of the nucleosome (Davey et al., 2002) (Figure 1-2).

After the first crystal structure of the *Xenopus* nucleosome core particle was solved at 2.8 Å and 1.9 Å resolutions by Richmond and colleagues (Luger *et al.*, 1997, Davey *et al.*, 2002),

structures of yeast, human, and *Drosophila* nucleosome core particles were solved (White *et al.*, 2001, Tsunaka *et al.*, 2005, Clapier *et al.*, 2008), showing strikingly similar structures. Subsequently, the structure of the nucleosome core particle containing Widom 601 positioning sequence was solved by Davey and our laboratory independently (Vasudevan *et al.*, 2010, Makde *et al.*, 2010). In our structure, 147 bp Widom 601 DNA fragment was overwound compared to the  $\alpha$ -satellite DNA, resulting in the appearance of a 145 bp nucleosome core particle with one additional base pair on either end (Makde *et al.*, 2010).



**Figure 1-3.** Histone-fold heterodimers and octamer (PDB ID: 1KX5). Heterodimeric histone-fold pairs in 'handshake motif' formed by (a) H3/H4 and (b) H2A/H2B shown in cartoon representation. (c) The histone-fold octamer shown in in cartoon representation. (Figure adapted from McGinty, R.K. and Tan, S. (2014) In Workman, L. J. and Abmayr, M.S. (eds.), *Fundamentals of Chromatin*. Springer New York, New York, NY, pp. 1-28.)

# 1.1.2 Posttranslational modifications of histones

Chromatin is not at an inert state, but is quite dynamic as its architecture and positioning,

as well as the chemical composition of the histones are regulated by chromatin remodelers and

modifiers. Histones can be modified at specific amino acids by acetylation, methylation,

phosphorylation, and ubiquitylation, which act synergistically to alter chromatin contacts or to facilitate recruitment of chromatin remodeling complexes (Kouzarides, 2007a). The posttranslational modifications predominantly occur on histone tails, and various patterns of these modifications on tails elicit specific outcomes (Luger *et al.*, 2012).

Acetylation occurs at lysine residues and is catalyzed by histone acetyltransferases (HATs). This modification, which neutralizes the positive charge of histone tails and decreases their affinity for DNA, is associated with transcriptional activation (Hong *et al.*, 1993). Consequently, HAT complexes are dedicated as transcriptional activators and co-activators, whereas the histone deacetylases (HDACs) are linked to repressors (Kuo and Allis, 1998, Khochbin et al., 2001). Phosphorylation, which adds a negative charge, can occur on serine, threonine, and tyrosine residues of all four histone tails and is regulated by protein kinases and phosphatases. A number of studies suggest that histone phosphorylation is involved in DNA damage repair, transcription regulation of proliferative genes, and chromatin compaction (Lau et al., 2011, van Attikum and Gasser, 2005, Wei et al., 1998). Ubiquitylation is a unique epigenetic mark that attaches a small protein called ubiquitin and is catalyzed by a cascade of ubiquitinactivating E1, ubiquitin-conjugating E2, and ubiquitin-ligating E3, including Polycomb group protein RING1B (Wang et al., 2004). Lys119 on H2A and Lys120 on H2B can be monoubiquitylated, and other lysine residues can be poly-ubiquitylated (Goldknopf et al., 1975, West and Bonner, 1980, Cao and Yan, 2012). The function of ubiquitylation is less well understood than other posttranslational modifications but is known to be associated with transcription, maintenance of chromatin structure, and DNA repair (Cao and Yan, 2012). Histone methylation occurs at lysine and arginine residues, and does not alter the charge of the histone protein unlike acetylation and phosphorylation. Histone methylation can either activate or repress transcription depending on its location and degree. Details of histone methylation are discussed later in the chapter.

### 1.1.3 Recognition of nucleosomes by chromatin factors

Many chromatin factors require nucleosomes for binding to chromatin to perform their respective activity. Even though the structures of numerous chromatin remodeling and histonemodifying enzymes have been solved, a holistic view of their mechanism remains elusive without information of how they act on their physiological substrates.

The crystal structure of RCC1 (regulator of chromosome condensation) bound to the nucleosome at 2.9 Å resolution provided the first atomic view of how a chromatin factor interacts with the nucleosome (Makde *et al.*, 2010). In this structure, RCC1 interacts with the nucleosome via nucleosomal DNA and negatively charged H2A/H2B dimer surface referred to as the acidic patch (Figure 1-4). This acidic patch is formed by negatively charged residues in H2A (Glu56, Glu61, Glu64, Asp90, Glu91, and Glu92) and H2B (Glu102 and Glu110).

Subsequently, crystal structures of silencing protein Sir3, PRC1 ubiquitylation module, and SAGA deubiquitinating (DUB) module bound to nucleosome core particle have been solved (Armache *et al.*, 2011, McGinty *et al.*, 2014, Morgan *et al.*, 2016). These structures show that the chromatin factors bind to nucleosomes through interaction with histone tails, the disk face of histone octamer, and nucleosomal DNA (McGinty and Tan, 2014). In addition to these structures of chromatin proteins on nucleosomes, the crystal structures of Kaposi's sarcoma-associated herpesvirus LANA (Latency-Associated Nuclear Antigen) peptide and centromeric protein CENP-C peptide bound to the nucleosome were solved (Barbera *et al.*, 2006, Kato *et al.*, 2013). All six crystal structures of chromatin proteins and peptides bound to the nucleosome share a common interaction motif, an arginine bound to the H2A/H2B acidic patch (Figure 1-5). In this motif, an arginine side chain is inserted into the acidic patch, representing a hot spot for recognition of nucleosome (McGinty and Tan, 2016). In addition, the acidic patch of one nucleosome can bind to H4 tails of the adjacent nucleosome, mediating higher order folding of chromatin fiber (Kalashnikova et al., 2013).



**Figure 1-4**. Nucleosome core particle surface and interactions. (a) Electrostatic potential and (b) van der Waals surface representations of the nucleosome core particle. The H2A/H2B acidic patch is labeled. (Figure adapted from McGinty, R.K. and Tan, S. (2014) In Workman, L. J. and Abmayr, M.S. (eds.), *Fundamentals of Chromatin*. Springer New York, New York, NY, pp. 1-28.)



**Figure 1-5.** Recognition of nucleosomes by chromatin factors using the acidic patch arginine anchor. From left to right, structures of RCC1 (PDB ID: 3MVD) (Makde *et al.*, 2010), PRC1 (PDB ID: 4R8P) (McGinty *et al.*, 2014), Sir3 (PDB ID: 3TU4) (Armache *et al.*, 2011), and SAGA DUB (PDB ID: 4ZUX) (Morgan *et al.*, 2016) bound to the nucleosome core particle. Overview of structures as viewed from the dyad (top) and zoomed view of the acidic patch (bottom) with arginine-anchor in space-filling representation and key H2A and H2B residues shown as sticks. Histones H2A, H2B, H3, H4 and DNA are shown in cartoon representation and colored wheat, pink, cornflower blue, light green, and light blue respectively. Figure prepared using PyMOL molecular graphics software (Schrodinger, 2011).

#### 1.2 Histone Methylation, Methyltransferases, and Demethylases

### **1.2.1 Histone methylation**

Histone methylation plays a key role in the regulation of important cellular processes such as cell cycle regulation, DNA damage and stress response, development, and differentiation (Kouzarides, 2007a, Pedersen and Helin, 2010, Nottke *et al.*, 2009). Numerous methyltransferases have been identified that mediate the addition of methyl groups to different lysine and arginine residues on histones. Lysine residues can be mono-, di-, or trimethylated (me1, me2, or me3), and arginines can be monomethylated (me1), symmetrically dimethylated (me2s), or asymmetrically dimethylated (me2a) (Kouzarides, 2007b). The SET domain family and lysine N-methyltransferase DOT1L methylate lysines, and the PRMT (protein arginine methyltransferase) family methylates arginines. Accumulating information shows that many of these enzymes are implicated in diseases such as cancers and neurological disorders, accentuating the importance of proper histone methylation maintenance in gene regulation (Hui and Ye, 2015).

## 1.2.2 Lysine methylation and demethylation

Lysine methylation is a key post-translational histone modification because it is associated with both transcriptional activation and repression, and the specific form of methylation at specific lysine residue can be functionally important. For example, the promoter regions of transcriptionally active genes are marked with histone H3 Lys4 (H3K4) trimethylation, while the enhancer regions are marked with H3K4 monomethylation (Bernstein *et al.*, 2002, Heintzman *et al.*, 2007, Barski *et al.*, 2007).

Chromatin can be regulated by lysine methylation at two levels: chromosome and gene levels (Figure 1-6). At 'chromosome level' of regulation, specific histone methylation can contribute to the formation of the two general states of chromatin, heterochromatin and euchromatin (Wozniak and Strahl, 2014). For example, methylation at H3K4 generally appears to counteract silencing mechanisms by specifically impairing Suv39h1-mediated methylation at H3K9, thereby preventing a major pathway of heterochromatin formation (Shilatifard, 2008). In addition, methylation of H3K4 appears to be a global epigenetic mark in euchromatic regions and trimethylation of H3K4 correlates with active transcription (Wozniak and Strahl, 2014).

Unlike histone acetylation, histone methylation does not alter the overall charge of the nucleosomes. Instead, at 'gene level', it seems likely that histone methylation serves as marks or binding sites that can be identified by specific proteins to facilitate downstream events resulting in altered gene expression (Sims et al., 2003). The position of the methylated lysine within histones H3 or H4 marks a gene to be either activated or repressed. While H3K4 methylation is associated with transcriptionally active genes, H3K9 methylation, a mere five amino acids away, is associated with transcriptional repression (Figure 1-6). Usually H3K4me2 and H3K4me3 both peak at the transcription start site. It also appears that H3K4 methylation does not directly affect transcription, but requires downstream effectors (Sims et al., 2003).

Several factors have been reported to recognize and interact with H3K4me2 and H3K4me3. CHD1, an ATPase-dependent chromatin remodeler, recognizes H3K4me2 and H3K4me3 with equal affinities (Flanagan et al. 2005; Sims et al. 2005). NURF, the chromatinremodeler complex, which has shown to stimulate transcription by facilitating preinitiation complex formation, also selectively recognizes H3K4me2/me3 (Mizuguchi et al. 1997). However, the mechanism of how methylation of H3K4 facilitates transcriptional activation is not fully elucidated. It has been suggested that CHD1 and NURF may recruit additional downstream effectors to activate transcriptional activation (Shilatifard, 2008). In addition, studies also demonstrated that TFIID, the basal transcription factor, binds to trimethylated H3K4me3 via PHD finger of TAF3 subunit (Vermeulen et al., 2007). Therefore, H3K4me3 marks play importantant role in defining core promoter by creating a binding site for TFIID.

Previously, lysine methylation was thought to be an irreversible process until the discovery of two enzymes capable of demethylating histone tails: lysine-specific demethylase (LSD1) (Shi *et al.*, 2004) and JmjC-domain containing demethylase (Tsukada *et al.*, 2006), confirming that lysine methylation can be dynamically regulated. Since then, a plethora of lysine demethylases has been discovered, each one catalyzing residue- and degree-specific demethylation of lysine residues at a specific region of the genome. For example, the Jumanji domain protein, JARID1B, demethylates H3K4me2 and H3K4me3, while the LSD1 protein demethylates H3K4me1 and H3K4me2 in promoter regions (Hou and Yu, 2010).





#### **1.2.3 LSD family of histone demethylase**

The LSD family, composed of two members, LSD1 (also known as KDM1A, AOF2, and BHC110) and LSD2 (also known as KDM1B and AOF1), demethylates H3K4me1/me2. LSD1 is enriched at promoter regions while LSD2 associates with gene bodies of actively transcribed genes (Chen et al., 2013). The LSD family is characterized by the presence of both an amine oxidase domain (AOD; also known as AOL) and a SWIRM (SWI3, RSC8, and Moira) domain (Figure 1-7). The SWIRM domain is conserved in chromatin-associated proteins and is thought to be involved in protein and complex stability and nucleosome targeting (Da et al., 2006). The presence of the Tower domain in LSD1 and the zinc finger domain in LSD2 distinguishes the two proteins from each other (Chen et al., 2013) (Figure 1-7). Due to the presence of different domain architectures, each protein is able to associate with different cofactors or additional proteins that positively stimulate the demethylase activity (Fang et al., 2013, Boyer et al., 2002). The LSD1 protein is usually found in complex with REST co-repressor proteins (CoREST1-3), which interact with LSD1 through the Tower domain, and also with the histone deacetylases 1 and 2 (HDAC1-2) (Upadhyay et al., 2014, Yang et al., 2011, Vasilatos et al., 2013). On the other hand, the LSD2 protein, which lacks the Tower domain for interaction with the CoREST proteins, binds NPAC, a cofactor that reads the H3K36me3 mark and enhances LSD2's activity on the nucleosomes (Fang et al., 2013).



**Figure 1-7**. Features of the LSD family. LSD1 and LSD2 have an amine oxidase-like domain (AOL also known as AOD) and a SWIRM domain. The Tower domain is unique to LSD1, whereas the zinc finger (ZF) domain is unique to LSD2. (Figure from Hojfeldt, J.W., Agger, K. and Helin, K. (2013) Histone lysine demethylases as targets for anticancer therapy. *Nature reviews. Drug discovery*, **12**, 917-930.)

#### 1.3 LSD1/CoREST Complex

## 1.3.1 Biological roles of LSD1

LSD1 is an essential protein that plays many different specific roles in biological processes including development, cellular differentiation, embryonic pluripotency, and cancer (Amente et al., 2013). The LSD1 protein occupies the promoters of a subset of developmental genes, where it controls the levels of H3K4 methylation through its demethylase activity. LSD1 has a role in maintaining the silencing of several developmental genes in human embryonic stem cells by regulating the critical balance between H3K4 and H3K27 methylation at their regulatory regions (Adamo *et al.*, 2011). Additionally, the LSD1 protein is highly expressed in undifferentiated human embryonic stems cells and is progressly downregulated during differentiation, suggesting its significant role in development (Whyte et al., 2012). LSD1 is also involved in hematopoietic differentiation by controlling hemangioblast formation and generation of hematopoietic stem cells (Kerenyi et al., 2013). Moreover, neural stem cell proliferation and neuronal development are also regulated by LSD1. The LSD1 protein associates with CoREST to mediate the repression of neuronal-specific genes, and the disassembly of LSD1/CoREST is required for differentiation into mature neurons (Fuentes et al., 2012). Furthermore, LSD1 is overexpressed in many human cancers, and its inhibition reduces cancer cell growth, migration and invasion (Lynch et al., 2012). As the importance of LSD1 in oncogenesis is increasingly appreciated, LSD1 has emerged as a novel therapeutic target for cancer.

#### **1.3.2 Activity of LSD1/CoREST complex**

While the LSD1 protein is able to catalyze demethylation of methylated H3K4 peptides, it is not able to catalyze the demethylation on methylated H3K4 incorporated into nucleosomes

on its own (Yang et al., 2006, Lee et al., 2005). The catalytic activity is contained in its FADdependent amine oxidase domain, and this domain is sufficient to bind to the N-terminal tail of histone H3 (Yang et al., 2006, Chen et al., 2006, Stavropoulos et al., 2006, Forneris et al., 2007, Yang et al., 2007). Although the catalytic site is large enough to accommodate a trimethylated H3K4, the LSD enzymes can remove methyl groups from mono- and dimethylated, but not trimethylated lysines (Stavropoulos et al., 2006). This is because the catalytic mechanism of the LSD family and amine oxidases requires a protonated nitrogen in the substrates, restricting the substrates to mono- or dimethylated lysines that have a lone electron pair on the lysine  $\varepsilon$ -nitrogen atom (Yang et al., 2006, Stavropoulos et al., 2006) (Figure 1-8a). The nucleosome specific enzymatic activity requires CoREST, which heterodimerizes with LSD1 (Lee et al., 2005, Yang et al., 2006, Forneris et al., 2007). Interestingly, the specificity of LSD1 can be altered by its binding partners, including CoREST, androgen receptor, and BHC80 (Lee et al., 2005, Metzger et al. 2005, Shi et al. 2005). The LSD1 acts as a repressor when in association with CoREST by demethylating H3K4me1/me2. In contrast, the LSD1 protein works as a coactivator when in association with androgen receptor by demethylating H3K9me1/me2 (Metzger et al. 2005). While both the CoREST and the androgen receptor positively regulate LSD1's function, BHC80 inhibits the LSD1-mediated demethylation in vitro, conferring negative regulation (Shi et al., 2005).

#### 1.3.3 Structure of LSD1/CoREST complex

The crystal structure of LSD1(171-852) and CoREST(286-482) has been solved 2.6 Å resolution by Yang *et al.* in 2006 (Figure 1-8c). The LSD1/CoREST complex forms an elongated structure of 150 Å in length, with the LSD1 amine oxidase and SWIRM domains on one end, a CoREST SANT2 domain on the other, and a helical linker in between formed by both LSD1 and

CoREST (Yang *et al.*, 2006). CoREST contains two SANT domains, which are highly conserved in many chromatin remodeling proteins and implicated in DNA binding and recognition of histone substrates (Boyer *et al.*, 2002). However, deletion analysis has shown that only one of the SANT domains (SANT2) is necessary and sufficient to stimulate the LSD1's activity towards the nucleosome (Yang *et al.*, 2006).



**Figure 1-8.** Mechanism, domain organization, and structure of the LSD1/CoREST complex. (a) Mechanism of LSD1-catalyzed demethylation of H3K4. (b) Domain organization of human LSD1 and CoREST with boundaries or proteins used in crystallization in Yang *et al.*, 2006 as indicated with the following color schemes: SWIRM, Orange; AOD yellow; LSD1 Tower, ivory; CoREST linker, red; and CoREST SANT2, purple. The regions not shown in the structure are colored in gray. (PDB ID: 2VID) (c) Overall crystal structure of LSD1/CoREST shown in cartoon representation with the same color schemes in (b). The FAD cofactor is shown in stick representation and the active site is indicated with an arrow. Figure prepared using PyMOL molecular graphics software (Schrodinger, 2011).

The amine oxidase domain (AOD) of LSD1 contains a substrate binding lobe and a FAD

binding lobe that is packed against the SWIRM domain. The LSD1 Tower domain that protrudes

from the amine oxidase domain consists of two long helices that pack against each other in

antiparallel orientation. The CoREST linker, which consists of a long helix and a short helix, binds to LSD1 through this Tower domain, forming a parallel coiled coil using the long helix. On the other hand, the short helix of the CoREST linker, which is perpendicular to the long helix, packs against the substrate binding lobe of the AOD. The SANT2 domain of CoREST sits at the tip of the tower domain (Figure 1-8c).

A direct interaction between the peptide substrate and the CoREST protein is not apparent in the structure of LSD1/CoREST/H3-petide complex (Yang *et al.*, 2007, Forneris *et al.*, 2007). Thus, the altered specificity of the LSD1 protein by association either with the CoREST protein or androgen receptor must come from an allosteric mechanism. Although the crystal structures of the LSD1 protein by itself and LSD1 complexed with CoREST suggest that there is no significant conformational change in LSD1 upon association with the CoREST protein, recent studies using molecular dynamics simulation suggest for molecular flexibility of LSD1/CoREST in solution (Baron and Vellore, 2012a). Their studies indicate that LSD1/CoREST in solution is flexible and assumes configurations that are substantially more open or closed than the crystal structures of LSD1/CoREST (Figure 1-9). In particular, the distance between the LSD1 SWIRM and the CoREST SANT2 domain oscillates significantly, from ~100 Å to ~129 Å (distance calculated between the center of mass of the domains). Additionally, the substrate binding pocket is also highly dynamic, but upon H3-peptide binding, the breathing of the LSD1 pocket is decreased, reducing its flexibility (Baron and Vellore, 2012b).

The crystal structures of the LSD1/CoREST complex bound to the H3-peptide provided ample information on how LSD1/CoREST works on the peptide substrates (Forneris *et al.*, 2007, Yang *et al.*, 2007). However, these studies do not provide insight into how LSD1/CoREST interacts with the nucleosome, the physiological substrate. As such, biochemical and structural studies of LSD1/CoREST on the nucleosomes are critical to fully understand LSD1/CoREST's function and its downstream effects and implications in diseases.



**Figure 1-9**. Molecular dynamics simulation reveals molecular flexibility of the LSD1/CoREST complex in solution. The enzyme assumes configurations substantially more open or closed than the crystal structures of LSD1/CoREST, with the distance between the LSD1 SWIRM and the CoREST SANT2 domain fluctuating significantly. (Figure from Baron, R. and Vellore, N.A. (2012) LSD1/CoREST reversible opening-closing dynamics: discovery of a nanoscale clamp for chromatin and protein binding. *Biochemistry*, **51**, 3151-3153.)

## 1.3.4 Substrate binding site of LSD1/CoREST complex

Although the AOD of LSD1 is structurally highly similar to other amine oxidases, its binding pocket is distinctly larger and contains an acidic surface near the entrance to the active site, which complements the basic histone tail (Chen *et al.*, 2006, Yang *et al.*, 2006). The crystal structures of LSD1/CoREST complex bound to a 21-amino acid H3-peptide in which Lys4 is substituted to Met or propargyl unit (Forneris *et al.*, 2007, Yang *et al.*, 2007) show an intricate network of interactions in substrate binding site (Figure 1-10). The first 16 amino acids of H3 peptide all specifically interact with LSD1 but the electron density after Pro16 is too poorly defined to allow model building, suggesting that the flexibility of amino acids 17-21 may be required to allow the tail to reach the binding pocket (Forneris *et al.*, 2007). Furthermore, the LSD1 protein is not active toward peptides shorter than 16 amino acids, and an optimal activity

occurs with 21-residue H3-peptide (Forneris *et al.*, 2005). There are specific "pockets" in the binding site of LSD1 that interact with H3 side chains including Thr6, Arg8, Lys9, and Thr11, and N-terminal amino group of Ala1 (Forneris *et al.*, 2008) (Figure 1-10). The H3 residue Arg2 is essential in establishing the folded conformation of H3-peptide as it is at the center of a hydrogen-bond network. The side chain of Arg2 forms hydrogen-bonds with carbonyl oxygens of Gly12 and Gly13, and the side chain of Ser10. These interactions specifically position the H3 Lys4 side chain in proximity to FAD for demethylation (Forneris *et al.*, 2007).

The intricate network also explains the ability of LSD1 to read the epigenetic modifications on H3. Phosphorylation of Ser10 abolishes the LSD1's activity, as it disrupts the binding geometry. Acetylation on Lys9 also decreases the binding affinity of the substrate to LSD1 whereas methylation on Lys9 shows no effect, indicating that electrostatic interactions are crucial in enzyme-substrate recognition (Forneris *et al.*, 2005). Moreover, the N-terminal amino group of Ala1 interacts with C-terminus of LSD1 helix 524-540, and residues Lys14-Pro16 of H3 are in contact with LSD1 helix 372-395. Even though CoREST does not take part in substrate binding, its short helix makes extensive interactions with LSD1 helix 372-395. By maintaining the overall architecture of the substrate binding site and interacting with nucleosome, CoREST fine tunes LSD1's specificity for H3K4 in nucleosomal substrates.

### 1.3.5 LSD1 as a therapeutic target

The LSD1 protein plays vital roles in cellular processes including neural stem cell proliferation, DNA damage response, and cancer proliferation and progression. Several studies indicate that LSD1 is overexpressed in variety of tumors, and recent findings demonstrate that LSD1 plays an important role in development and maintenance of acute myeloid leukemia (AML). Although the relationship between histone demethylation and human carcinogenesis still remains unclear, it has been shown that overexpression of LSD1 can cause dysregulation of chromatin structure and gene transcription and contribute to malignant transformation of the cells (Hayami et al., 2011). As LSD1 has emerged as an attractive therapeutic target in human cancer, pharmacological inhibitors of LSD1 have been developed. The first irreversible inhibitor identified was tranylcypromine (TCP) which inhibits LSD1 by binding covalently to FAD (Schmidt and McCafferty, 2007), followed by development of a number of TCP analogs. Recently, selective reversible inhibitors of LSD1 have been developed that differ in mode of action from TCP derivatives (Mould *et al.*, 2015). With a wide range of pharmacological inhibitors in development, understanding the biology and activity of LSD1 in the context of nucleosome, its physiological substrate, appears to be the key to further advance in therapeutics.



**Figure 1-10**. Histone peptide-binding site in LSD1. Specific "pockets" important for interaction between LSD1 and H3 peptide are emphasized in bold, and network of hydrogen bonds are depicted in dashed lines. Histone H3 is shown in green, and "short helix" of CoREST is shown in red. (Figure from Forneris, F., Binda, C., Battaglioli, E. and Mattevi, A. (2008) LSD1: oxidative chemistry for multifaceted functions in chromatin regulation. *Trends in biochemical sciences*, **33**, 181-189.)

### **1.4 References**

- Adamo, A., B. Sese, S. Boue, J. Castano, I. Paramonov, M. J. Barrero and J. C. Izpisua Belmonte.
  "LSD1 Regulates the Balance between Self-Renewal and Differentiation in Human Embryonic Stem Cells." *Nat Cell Biol* 13, no. 6 (2011): 652-9.
- Alberts B, Bray D, Lewis J, Raff M, Roberts K and Watson J D. *Molecular Biology of the Cell*. 3rd ed. New York: Garland, 1994.
- Amente, S., L. Lania and B. Majello. "The Histone LSD1 Demethylase in Stemness and Cancer Transcription Programs." *Biochim Biophys Acta* 1829, no. 10 (2013): 981-6.
- Arents, G. and E. N. Moudrianakis. "The Histone Fold: A Ubiquitous Architectural Motif Utilized in DNA Compaction and Protein Dimerization." *Proc Natl Acad Sci U S A* 92, no. 24 (1995): 11170-4.
- Armache, K. J., J. D. Garlick, D. Canzio, G. J. Narlikar and R. E. Kingston. "Structural Basis of Silencing: Sir3 BAH Domain in Complex with a Nucleosome at 3.0 a Resolution." *Science* 334, no. 6058 (2011): 977-82.
- Barbera, A. J., J. V. Chodaparambil, B. Kelley-Clarke, V. Joukov, J. C. Walter, K. Luger and K.
  M. Kaye. "The Nucleosomal Surface as a Docking Station for Kaposi's Sarcoma Herpesvirus LANA." *Science* 311, no. 5762 (2006): 856-61.
- Baron, R. and N. A. Vellore. "LSD1/CoREST Is an Allosteric Nanoscale Clamp Regulated by H3-Histone-Tail Molecular Recognition." *Proc Natl Acad Sci U S A* 109, no. 31 (2012): 12509-14.
- Baron, R. and N. A. Vellore. "LSD1/CoREST Reversible Opening-Closing Dynamics: Discovery of a Nanoscale Clamp for Chromatin and Protein Binding." *Biochemistry* 51, no. 15 (2012): 3151-3.
- Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev and K. Zhao. "High-Resolution Profiling of Histone Methylations in the Human Genome." *Cell* 129, no. 4 (2007): 823-37.
- Bernstein, B. E., E. L. Humphrey, R. L. Erlich, R. Schneider, P. Bouman, J. S. Liu, T. Kouzarides and S. L. Schreiber. "Methylation of Histone H3 Lys 4 in Coding Regions of Active Genes." *Proc Natl Acad Sci U S A* 99, no. 13 (2002): 8695-700.
- Boyer, L. A., M. R. Langer, K. A. Crowley, S. Tan, J. M. Denu and C. L. Peterson. "Essential Role for the SANT Domain in the Functioning of Multiple Chromatin Remodeling Enzymes." *Mol Cell* 10, no. 4 (2002): 935-42.
- Bradbury, E. Morton. "K. E. Van Holde. Chromatin. Series in Molecular Biology. Springer-Verlag, New York. 1988. 530 Pp. \$98.00." *Journal of Molecular Recognition* 2, no. 3 (1989): i-i.
- Cao, J. and Q. Yan. "Histone Ubiquitination and Deubiquitination in Transcription, DNA Damage Response, and Cancer." *Front Oncol* 2, (2012): 26.
- Chen, F., H. Yang, Z. Dong, J. Fang, P. Wang, T. Zhu, W. Gong, R. Fang, Y. G. Shi, Z. Li and Y. Xu. "Structural Insight into Substrate Recognition by Histone Demethylase LSD2/KDM1B." *Cell Res* 23, no. 2 (2013): 306-9.
- Chen, Y., Y. Yang, F. Wang, K. Wan, K. Yamane, Y. Zhang and M. Lei. "Crystal Structure of Human Histone Lysine-Specific Demethylase 1 (LSD1)." *Proc Natl Acad Sci U S A* 103, no. 38 (2006): 13956-61.
- Clapier, C. R., S. Chakravarthy, C. Petosa, C. Fernandez-Tornero, K. Luger and C. W. Muller.
  "Structure of the Drosophila Nucleosome Core Particle Highlights Evolutionary Constraints on the H2A-H2B Histone Dimer." *Proteins* 71, no. 1 (2008): 1-7.

- Da, G., J. Lenkart, K. Zhao, R. Shiekhattar, B. R. Cairns and R. Marmorstein. "Structure and Function of the SWIRM Domain, a Conserved Protein Module Found in Chromatin Regulatory Complexes." *Proc Natl Acad Sci U S A* 103, no. 7 (2006): 2057-62.
- Davey, C. A., D. F. Sargent, K. Luger, A. W. Maeder and T. J. Richmond. "Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9 Å Resolution." *J Mol Biol* 319, no. 5 (2002): 1097-113.
- Fang, R., F. Chen, Z. Dong, D. Hu, A. J. Barbera, E. A. Clark, J. Fang, Y. Yang, P. Mei, M.
  Rutenberg, Z. Li, Y. Zhang, Y. Xu, H. Yang, P. Wang, M. D. Simon, Q. Zhou, J. Li, M.
  P. Marynick, X. Li, H. Lu, U. B. Kaiser, R. E. Kingston, Y. Xu and Y. G. Shi.
  "LSD2/KDM1B and Its Cofactor NPAC/GLYR1 Endow a Structural and Molecular Model for Regulation of H3K4 Demethylation." *Mol Cell* 49, no. 3 (2013): 558-70.
- Forneris, F., C. Binda, A. Adamo, E. Battaglioli and A. Mattevi. "Structural Basis of LSD1-CoREST Selectivity in Histone H3 Recognition." *J Biol Chem* 282, no. 28 (2007): 20070-4.
- Forneris, F., C. Binda, E. Battaglioli and A. Mattevi. "LSD1: Oxidative Chemistry for Multifaceted Functions in Chromatin Regulation." *Trends Biochem Sci* 33, no. 4 (2008): 181-9.
- Forneris, F., C. Binda, M. A. Vanoni, E. Battaglioli and A. Mattevi. "Human Histone Demethylase LSD1 Reads the Histone Code." J Biol Chem 280, no. 50 (2005): 41360-5.
- Fuentes, P., J. Canovas, F. A. Berndt, S. C. Noctor and M. Kukuljan. "CoREST/LSD1 Control the Development of Pyramidal Cortical Neurons." *Cereb Cortex* 22, no. 6 (2012): 1431-41.
- Goldknopf, I. L., C. W. Taylor, R. M. Baum, L. C. Yeoman, M. O. Olson, A. W. Prestayko and
  H. Busch. "Isolation and Characterization of Protein A24, a "Histone-Like" Non-Histone
  Chromosomal Protein." *J Biol Chem* 250, no. 18 (1975): 7182-7.

- Hayami, S., J. D. Kelly, H. S. Cho, M. Yoshimatsu, M. Unoki, T. Tsunoda, H. I. Field, D. E. Neal, H. Yamaue, B. A. Ponder, Y. Nakamura and R. Hamamoto. "Overexpression of LSD1 Contributes to Human Carcinogenesis through Chromatin Regulation in Various Cancers." Int J Cancer 128, no. 3 (2011): 574-86.
- Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford and B. Ren. "Distinct and Predictive Chromatin Signatures of Transcriptional Promoters and Enhancers in the Human Genome." *Nat Genet* 39, no. 3 (2007): 311-8.
- Hong, L., G. P. Schroth, H. R. Matthews, P. Yau and E. M. Bradbury. "Studies of the DNA Binding Properties of Histone H4 Amino Terminus. Thermal Denaturation Studies Reveal That Acetylation Markedly Reduces the Binding Constant of the H4 "Tail" to DNA." J Biol Chem 268, no. 1 (1993): 305-14.
- Hou, H. and H. Yu. "Structural Insights into Histone Lysine Demethylation." *Curr Opin Struct Biol* 20, no. 6 (2010): 739-48.
- Hui, C. and T. Ye. "Synthesis of Lysine Methyltransferase Inhibitors." Front Chem 3, (2015): 44.
- Kalashnikova, A. A., M. E. Porter-Goff, U. M. Muthurajan, K. Luger and J. C. Hansen. "The Role of the Nucleosome Acidic Patch in Modulating Higher Order Chromatin Structure." *J R Soc Interface* 10, no. 82 (2013): 20121022.
- Kato, H., J. Jiang, B. R. Zhou, M. Rozendaal, H. Feng, R. Ghirlando, T. S. Xiao, A. F. Straight and Y. Bai. "A Conserved Mechanism for Centromeric Nucleosome Recognition by Centromere Protein CENP-C." *Science* 340, no. 6136 (2013): 1110-3.
- Kerenyi, M. A., Z. Shao, Y. J. Hsu, G. Guo, S. Luc, K. O'Brien, Y. Fujiwara, C. Peng, M.
  Nguyen and S. H. Orkin. "Histone Demethylase LSD1 Represses Hematopoietic Stem and Progenitor Cell Signatures During Blood Cell Maturation." *Elife* 2, (2013): e00633.

- Khochbin, S., A. Verdel, C. Lemercier and D. Seigneurin-Berny. "Functional Significance of Histone Deacetylase Diversity." *Curr Opin Genet Dev* 11, no. 2 (2001): 162-6.
- Kouzarides, T. "Chromatin Modifications and Their Function." Cell 128, no. 4 (2007): 693-705.
- Kouzarides, T. "Snapshot: Histone-Modifying Enzymes." Cell 131, no. 4 (2007): 822.
- Kuo, M. H. and C. D. Allis. "Roles of Histone Acetyltransferases and Deacetylases in Gene Regulation." *Bioessays* 20, no. 8 (1998): 615-26.
- Lau, A. T., S. Y. Lee, Y. M. Xu, D. Zheng, Y. Y. Cho, F. Zhu, H. G. Kim, S. Q. Li, Z. Zhang, A.
  M. Bode and Z. Dong. "Phosphorylation of Histone H2B Serine 32 Is Linked to Cell Transformation." *J Biol Chem* 286, no. 30 (2011): 26628-37.
- Lee, M. G., C. Wynder, N. Cooch and R. Shiekhattar. "An Essential Role for CoREST in Nucleosomal Histone 3 Lysine 4 Demethylation." *Nature* 437, no. 7057 (2005): 432-5.
- Luger, K., M. L. Dechassa and D. J. Tremethick. "New Insights into Nucleosome and Chromatin Structure: An Ordered State or a Disordered Affair?" *Nat Rev Mol Cell Biol* 13, no. 7 (2012): 436-47.
- Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond. "Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution." *Nature* 389, no. 6648 (1997): 251-60.
- Lynch, J. T., W. J. Harris and T. C. Somervaille. "LSD1 Inhibition: A Therapeutic Strategy in Cancer?" *Expert Opin Ther Targets* 16, no. 12 (2012): 1239-49.
- Makde, R. D., J. R. England, H. P. Yennawar and S. Tan. "Structure of RCC1 Chromatin Factor Bound to the Nucleosome Core Particle." *Nature* 467, no. 7315 (2010): 562-6.
- Makde, R. D. and S. Tan. "Strategies for Crystallizing a Chromatin Protein in Complex with the Nucleosome Core Particle." *Anal Biochem* 442, no. 2 (2013): 138-45.
- McGhee, J. D. and G. Felsenfeld. "Nucleosome Structure." *Annu Rev Biochem* 49, (1980): 1115-56.

- McGinty, R. K., R. C. Henrici and S. Tan. "Crystal Structure of the PRC1 Ubiquitylation Module Bound to the Nucleosome." *Nature* 514, no. 7524 (2014): 591-6.
- McGinty, R. K. and S. Tan. "Recognition of the Nucleosome by Chromatin Factors and Enzymes." *Curr Opin Struct Biol* 37, (2016): 54-61.
- McGinty, Robert K. and Song Tan. "Histone, Nucleosome, and Chromatin Structure." In *Fundamentals of Chromatin*, edited by L. Jerry Workman and M. Susan Abmayr, 1-28. New York, NY: Springer New York, 2014.
- Metzger, E., M. Wissmann, N. Yin, J. M. Muller, R. Schneider, A. H. Peters, T. Gunther, R. Buettner and R. Schule. "LSD1 Demethylates Repressive Histone Marks to Promote Androgen-Receptor-Dependent Transcription." *Nature* 437, no. 7057 (2005): 436-9.
- Morgan, M. T., M. Haj-Yahya, A. E. Ringel, P. Bandi, A. Brik and C. Wolberger. "Structural Basis for Histone H2B Deubiquitination by the SAGA DUB Module." *Science* 351, no. 6274 (2016): 725-8.
- Mould, D. P., A. E. McGonagle, D. H. Wiseman, E. L. Williams and A. M. Jordan. "Reversible Inhibitors of LSD1 as Therapeutic Agents in Acute Myeloid Leukemia: Clinical Significance and Progress to Date." *Med Res Rev* 35, no. 3 (2015): 586-618.
- Nottke, A., M. P. Colaiacovo and Y. Shi. "Developmental Roles of the Histone Lysine Demethylases." *Development* 136, no. 6 (2009): 879-89.
- Pedersen, M. T. and K. Helin. "Histone Demethylases in Development and Disease." *Trends Cell Biol* 20, no. 11 (2010): 662-71.
- Schmidt, D. M. and D. G. McCafferty. "Trans-2-Phenylcyclopropylamine Is a Mechanism-Based Inactivator of the Histone Demethylase LSD1." *Biochemistry* 46, no. 14 (2007): 4408-16.
- Schones, D. E., K. Cui, S. Cuddapah, T. Y. Roh, A. Barski, Z. Wang, G. Wei and K. Zhao.
  "Dynamic Regulation of Nucleosome Positioning in the Human Genome." *Cell* 132, no. 5 (2008): 887-98.

Schrodinger, LLC. "The Pymol Molecular Graphics System." 1.4 (2011).

- Shi, Y. J., C. Matson, F. Lan, S. Iwase, T. Baba and Y. Shi. "Regulation of LSD1 Histone Demethylase Activity by Its Associated Factors." *Mol Cell* 19, no. 6 (2005): 857-64.
- Shi, Y., F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, R. A. Casero and Y. Shi.
  "Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1." *Cell* 119, no. 7 (2004): 941-53.
- Shilatifard, A. "Molecular Implementation and Physiological Roles for Histone H3 Lysine 4 (H3K4) Methylation." Curr Opin Cell Biol 20, no. 3 (2008): 341-8.
- Sims, R. J., 3rd, K. Nishioka and D. Reinberg. "Histone Lysine Methylation: A Signature for Chromatin Function." Trends Genet 19, no. 11 (2003): 629-39.
- Stavropoulos, P., G. Blobel and A. Hoelz. "Crystal Structure and Mechanism of Human Lysine-Specific Demethylase-1." *Nat Struct Mol Biol* 13, no. 7 (2006): 626-32.
- Tsukada, Y., J. Fang, H. Erdjument-Bromage, M. E. Warren, C. H. Borchers, P. Tempst and Y. Zhang. "Histone Demethylation by a Family of Jmjc Domain-Containing Proteins." *Nature* 439, no. 7078 (2006): 811-6.
- Tsunaka, Y., N. Kajimura, S. Tate and K. Morikawa. "Alteration of the Nucleosomal DNA Path in the Crystal Structure of a Human Nucleosome Core Particle." *Nucleic Acids Res* 33, no. 10 (2005): 3424-34.
- Upadhyay, G., A. H. Chowdhury, B. Vaidyanathan, D. Kim and S. Saleque. "Antagonistic Actions of Rcor Proteins Regulate LSD1 Activity and Cellular Differentiation." *Proc Natl Acad Sci U S A* 111, no. 22 (2014): 8071-6.
- van Attikum, H. and S. M. Gasser. "The Histone Code at DNA Breaks: A Guide to Repair?" *Nat Rev Mol Cell Biol* 6, no. 10 (2005): 757-65.
- Vasilatos, S. N., T. A. Katz, S. Oesterreich, Y. Wan, N. E. Davidson and Y. Huang. "Crosstalk between Lysine-Specific Demethylase 1 (LSD1) and Histone Deacetylases Mediates

Antineoplastic Efficacy of HDAC Inhibitors in Human Breast Cancer Cells." *Carcinogenesis* 34, no. 6 (2013): 1196-207.

- Vasudevan, D., E. Y. Chua and C. A. Davey. "Crystal Structures of Nucleosome Core Particles Containing the '601' Strong Positioning Sequence." *J Mol Biol* 403, no. 1 (2010): 1-10.
- Vermeulen, M., K. W. Mulder, S. Denissov, W. W. Pijnappel, F. M. van Schaik, R. A. Varier, M.
  P. Baltissen, H. G. Stunnenberg, M. Mann and H. T. Timmers. "Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4." Cell 131, no. 1 (2007): 58-69.
- Wang, H., L. Wang, H. Erdjument-Bromage, M. Vidal, P. Tempst, R. S. Jones and Y. Zhang.
  "Role of Histone H2A Ubiquitination in Polycomb Silencing." *Nature* 431, no. 7010 (2004): 873-8.
- Wei, Y., C. A. Mizzen, R. G. Cook, M. A. Gorovsky and C. D. Allis. "Phosphorylation of Histone H3 at Serine 10 Is Correlated with Chromosome Condensation During Mitosis and Meiosis in Tetrahymena." *Proc Natl Acad Sci U S A* 95, no. 13 (1998): 7480-4.
- West, M. H. and W. M. Bonner. "Histone 2B Can Be Modified by the Attachment of Ubiquitin." *Nucleic Acids Res* 8, no. 20 (1980): 4671-80.
- White, C. L., R. K. Suto and K. Luger. "Structure of the Yeast Nucleosome Core Particle Reveals Fundamental Changes in Internucleosome Interactions." *EMBO J* 20, no. 18 (2001): 5207-18.
- Wozniak, G. G. and B. D. Strahl. "Hitting the 'Mark': Interpreting Lysine Methylation in the Context of Active Transcription." Biochim Biophys Acta 1839, no. 12 (2014): 1353-61.
- Yang, M., J. C. Culhane, L. M. Szewczuk, C. B. Gocke, C. A. Brautigam, D. R. Tomchick, M.
  Machius, P. A. Cole and H. Yu. "Structural Basis of Histone Demethylation by LSD1
  Revealed by Suicide Inactivation." *Nat Struct Mol Biol* 14, no. 6 (2007): 535-9.

- Yang, M., C. B. Gocke, X. Luo, D. Borek, D. R. Tomchick, M. Machius, Z. Otwinowski and H. Yu. "Structural Basis for CoREST-Dependent Demethylation of Nucleosomes by the Human LSD1 Histone Demethylase." *Mol Cell* 23, no. 3 (2006): 377-87.
- Yang, P., Y. Wang, J. Chen, H. Li, L. Kang, Y. Zhang, S. Chen, B. Zhu and S. Gao. "RCOR2 Is a Subunit of the LSD11 Complex That Regulates ESC Property and Substitutes for SOX2 in Reprogramming Somatic Cells to Pluripotency." *Stem Cells* 29, no. 5 (2011): 791-801.

# Chapter 2

## **Materials and Methods**

## 2.1 Preparation of LSD1/CoREST complexes

## 2.1.1 Cloning and site-directed mutagenesis

The coding regions for human LSD1 $\Delta$ 1 = LSD1(171–852) and CoREST $\Delta$ 1 = CoREST(286–482) were amplified from HeLa cDNA (referred to as LSD1 and CoREST in the text) and cloned into the pST50Trc transfer vectors by Song Tan. For LSD1, the modification was made to include N-terminal StrepII-hexahistidine and transferred into the pST44 polycistronic expression vector. The pST44 polycistronic expression system, developed in our lab, enables the protein subunits to be coexpressed and purified as a complex in contrast to the traditional methods where the proteins are expressed and purified individually for subsequent reconstitution (Tan *et al.*, 2005). The pST50Trc1-4 transfer plasmids are available for each of the four positions in the pST44 polycistronic plasmids which harbor the corresponding restriction sites (Figure 2-1).

For mutant LSD1 or CoREST, a site-directed mutagenesis was performed on the pST50Trc transfer vector carrying either the LSD1 $\Delta$ 1 or CoREST $\Delta$ 1. Five ng of template plasmid was amplified for typically 16-18 cycles in Q5 Reaction buffer supplied by New England Biolabs containing 25 mM TAPS-HCl (pH 9.3 @ 25°C), 50 mM KCl, 2 mM MgCl<sub>2</sub>, and 1 mM  $\beta$ -mercaptoethanol. DNA was added to a mix of the buffer, 2 pmol of oligonucleotide primer, 200  $\mu$ M dNTPs, and 1 unit of Q5 polymerase. Following the amplification, the PCR product was exposed to 10 units of DpnI restriction enzyme at 37°C for 1 hour. Two  $\mu$ l of the digested product was transformed into TG1 *E. coli* competent cells by standard transformation protocol. The transformed cells were plated on TYE plate (1.0% bactotryptone, 0.5% yeast extract, 0.8% NaCl,

and 1.5% agar) with 100  $\mu$ g/ml ampicillin and grown at 37°C overnight. The next day, the transformation colonies were screened by PCR screening followed by restriction digestion, which relies on the introduction or removal of a translationally silent restriction site incorporated into the mutagenesis oligonucleotides. Therefore, the appearance or disappearance of the particular restriction site can be used as a marker for successful mutagenesis.



**Figure 2-1**. Schematic representation of pST44 polycistronic expression system. The pST50Trc1–4 plasmids provide the translation cassettes for the 4 possible positions in the pST44 polycistronic expression vector. Specific pairs of restriction enzymes are used to subclone the translation cassettes from the pST50Trc1– 4 plasmids into pST44 (XbaI-BgIII for cassette 1, EcoRI-HindIII for cassette 2, SacI-KpnI for cassette 3, BspEI-MluI for cassette 4). The XbaI-BgIII translation cassette from pST50Trc1 plasmid is detailed to highlight the translational enhancer ( $\epsilon$ ) and the Shine-Dalgarno sequence (SD) which precede the coding region (black bar bracketed by NdeI and BsrGI sites). Restriction sites and coding regions for translation cassettes 1, 2, 3, 4 are shown in blue, green, yellow and red respectively. Transcriptional promoter and termination signals on pST50Trc1–4 are present but not shown. All restriction sites shown are unique. (Figure from Selleck, W. and Tan, S. (2008) Recombinant protein complex expression in E. coli. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*, **Chapter 5**, Unit 5 21.)

The positive colonies were grown in 100 ml of 2xTY media (1.6% bacto tryptone, 1.0%

yeast extract, and 0.5% NaCl) with 50 µg/ml ampicillin overnight. Plasmid DNA was isolated via

the alkaline lysis method. The cell cultures were harvested by centrifugation at 3,220 x g for 5 minutes, and the cell pellet was resuspended in 5 ml of lysis buffer (50 mM glucose, 25 mM Tris-Cl, pH 8.0, and 10 mM EDTA), followed by an addition of 10 ml of NaOH/SDS (0.2 M NaOH, 1% SDS). The mixture was shaken vigorously and incubated on ice for 5 minutes. The prechilled 10 ml of 5 M KAc/2.5 M HAc was added to precipitate chromosomal DNA and neutralize the pH, followed by another vigorous mixing and incubation on ice for 5 minutes. The sample was centrifuged at 3220 x g for 5 minutes, and the supernatant was mixed with 12.5 ml of 100% isopropanol and incubated at room temperature for 5 minutes in the 50 ml round-bottomed polypropylene centrifuge tubes. The samples were spun at 13,000 rpm at 20°C in a Sorvall SS-34 rotor and the supernatant was poured off. The pellet was resuspended in 0.5 ml of 70% ethanol and transferred to Eppendorf tubes, followed by centrifugation at full-speed in a microcentrifuge for 1 minute. The supernatant was aspirated off, and the pellet was resuspended in TE (10, 50) (10 mM Tris-Cl, pH 8.0, 50 mM EDTA), followed by an addition of 1.5 µl of 10 mg/ml RNase A and incubation in 37°C water bath for 15 minutes. The sample was extracted with phenol/chloroform twice, and with chloroform once to remove the remaining cellular proteins and RNase A. The aqueous phase was loaded on to the prepared spun columns with Sephacryl S400 HR resin equilibrated in TE (10, 0.1) and centrifuged at 1,610 x g for 3 minutes. The plasmid eluted in TE (10, 0.1).

The sequence of the plasmid was confirmed by restriction mapping and DNA sequencing. The cassette carrying LSD1 or CoREST was subcloned into a pST44 polycistronic expression vector containing the complementary subunit.

#### 2.1.2 Co-expression of LSD1/CoREST complexes

Hexahistidine-tagged human LSD1 $\Delta$ 1 = LSD1(171–852) and CoREST $\Delta$ 1 = CoREST(286–482) were expressed using the pST44-polycistronic expression vector (Tan et al., 2005) in BL21(DE3)pLysS cells to produce recombinant LSD1 $\Delta$ 1/CoREST $\Delta$ 1 complex (referred in the text as LSD1/CoREST). The transformed *E. coli* cells were plated on the TYE plate containing 100 µg/ml ampicillin and 25 µg/ml chloramphenicol and grown at 37°C overnight. The next day, 3-5 colonies were selected and grown in 100 ml of 2xTY media (1.6% bacto tryptone, 1.0% yeast extract, and 0.5% NaCl) with 50 µg/ml ampicillin and 25 µg/ml chloramphenicol at 37°C as a starter culture, until the OD600 between 0.1 and 1.0 was reached. For large scale expression, typically 6 or 12 L of 2xTY was used, and 5 ml of the starter culture was inoculated into each of 500 ml 2xTY media. Cultures were permitted to grow at 37°C until an optical density of 0.2 was reached at which point the temperature was reduced to 23°C. Five hundred  $\mu$ l of 0.2 M of IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside) was added to induce the expression when the OD600 was at  $\sim 0.8$ . After growing for 5 hours, the cells were centrifuged in the SLA3000 rotor at 7,000 rpm for 5 minutes. The cell pellet was resuspended in P300-EDTA (50 mM sodium phosphate pH 7.0, 300 mM NaCl, 1 mM benzamidine, and 5 mM βmercaptoethanol) at the ratio of 25 ml of P300-EDTA per 1 L of culture. The resuspended cells were flash frozen in liquid nitrogen and stored at -20°C.

#### 2.1.3 Purification of LSD1/CoREST complexes

The frozen cells were thawed in the 30°C water bath. The thawed whole cell extract was sonicated with Branson Digital Sonifier (450D) at 70% power for ~7 repeats of 14 pulses (0.5 second on and 0.5 second off). The extract was incubated on ice between each round of

sonication. The soluble extract was prepared by centrifugation of the sonicated whole cell extract in the SS34 rotor at 18,000 rpm for 20 minutes at 4°C. The soluble extract was loaded on to Talon (Clonthech) cobalt affinity column equilibrated with P300-EDTA (50 mM sodium phosphate pH 7.0, 300 mM NaCl, 1 mM benzamidine, and 5 mM β-mercaptoethanol) buffer at 4°C. The column was washed with P300-EDTA with 15 mM imidazole and eluted with P300-EDTA with 100 mM imidazole. (The LSD1/CoREST complex fractions have yellow tint due to the FAD cofactor in the LSD1 protein. The intensity of the yellow color can also be used as a rough estimate for the concentration of the LSD1/CoREST complex.) Fractions were analyzed by SDS-PAGE, and the desired fractions were pooled and dialyzed against 2L of T100 (20mM Tris-Cl pH 8.0, 100 mM NaCl, and 10 mM  $\beta$ -mercaptoethanol) at room temperature overnight during digestion with TEV protease to remove the hexahistidine tag. The pool was centrifuged at 18,000 rpm for 15 minutes in the SS34 rotor at 10°C before being loaded on to the SourceQ column (GE Healthcare). A simple gradient was used to elute the complex, and the desired fractions were pooled based on the SDS-PAGE results (Figure 2-2). The purity of the complex after SourceQ anion-exchange chromatography was usually > 95%. The pooled fractions were dialyzed into 2L of H100 (10 mM HEPES, 100 mM NaCl, and 5 mM  $\beta$ -mercaptoethanol) and was concentrated with Vivaspin centrifugal concentrator (30,000 MWCO, Sartorius). Dynamic light scattering was used to check for aggregation of the complex, and the polydispersity of the LSD1/CoREST complexes was below 15%, indicating the sample was monodispersed. Glycerol was added to 20% w/v before the complex was flash frozen and stored at -20°C. All mutant LSD1/CoREST complexes were expressed and purified similarly to their wild-type versions.



**Figure 2-2.** Purification of the LSD1/CoREST complex. (a) Chromatogram of LSD1/CoREST purified over by SourceQ anion-exchange chromatography. (b) Corresponding SDS-PAGE gel of chromatography fractions.

#### 2.2 Nucleosome Preparation

## 2.2.1 Expression of recombinant histones

The introduction of mutations into *Xenopus* histones was performed by site-directed mutagenesis on the pST50Trc transfer vectors as described above. Recombinant *Xenopus* core histones were prepared as described previously (Luger *et al.*, 1999).

Expression of histones was performed using the T7 expression system in the pST50Tr vector and grown in BL21(DE3)pLysS cells. The transformed cells were plated on TYE plates supplemented with 100  $\mu$ g/ml ampicillin and 25  $\mu$ g/ml chloramphenicol and grown overnight at 37°C. The following day, 5-7 colonies were selected to inoculate 100 ml 2xTY starter culture flasks. The starter culture was used to inoculate large scale cultures (usually 3-6 L) of 2xTY media, and once the OD600 of 0.6-0.8 was reached, 0.5 ml of the 0.2 M IPTG was used to induce expression. After growing for additional 3 hours, the cells were harvested by centrifugation in the SLA3000 rotor at 7,000 rpm for 5 minutes. The cell pellets were then resuspended in T100 (20mM Tris-Cl pH 8.0, 100 mM NaCl, and 10 mM  $\beta$ -mercaptoethanol) (typically 150 ml of T100 for 6 L of cells), and flash frozen in liquid nitrogen and stored at -20°C.

## 2.2.2 Purification of recombinant histones

The frozen cells were thawed in the 30°C water bath and sonicated in 50 ml aliquots with the Branson sonicator for 4 cycles of 10 pulses at 50% power. The extract was incubated on ice between each round of sonication. The sonicated cell extract was transferred to 50 ml polypropylene tubes and spun in the SS34 rotor at 18,000 rpm for 30 minutes at room temperature. The supernatant was poured off as the pellet contains the histone proteins. The

pellets were then resuspended in 180 ml of TRITON buffer (20 mM Tris-Cl pH 8.0, 0.5 mM EDTA, 100 mM NaCl, 10 mM β-mercaptoethanol, 1 mM benzamidine, and 1% TRITON X-100) to dissolve membranes and membrane-associated proteins. The resuspended sample was centrifuged at 15,000 rpm in the SS34 rotor for 10 minutes at room temperature. The supernatant was transferred to a flask, while the pellets were extracted tree more times with TRITON buffer. The pellets were then resuspended in 50 ml WASH buffer (20 mM Tris-Cl pH 8.0, 0.5 mM EDTA, 100 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, and 1 mM benzamidine) followed by centrifugation at 15,000 rpm in the SS34 rotor for 10 minutes at 20°C. The supernatant was poured and aspirated off as the off-white pellet contained the histone proteins. An SDS-PAGE gel of the samples in each step of the inclusion body prep (crude extract whole, crude extract supernatant, TRITON whole, TRITON supernatants, and WASH whole) was run to confirm that the histones remained in the pellet. One ml of DMSO (dimethyl sulfoxide) was added to the pellet and incubated for 30 minutes at room temperature. After the incubation, 30 ml of TG0 (20 mM Tris-Cl, pH 8.0, 10 mM DTT, and 7 M guanidine-HCl) was added and homogenized using the douncer, followed by an additional addition of 20 ml of TG0. The homogenized sample was centrifuged at 15,000 rpm in the SS34 rotor for 20 minutes at 20°C. The supernatant was loaded on to the S300 column equilibrated with AU200 (20 mM NaAc, pH 5.2, 0.5 mM EDTA, 10 mM β-mercaptoethanol, 150 mM NaCl, 50 mM NH<sub>4</sub>Cl, and 8M urea).

The fractions from the S300 chromatography were pooled based on the SDS-PAGE gel and further purified by SourceS cation-exchange high-performance liquid chromatography (HPLC) (GE Healthcare). The SourceS fractions were analyzed by the SDS-PAGE electrophoresis and pooled accordingly. The concentration of the histone proteins was determined, with typical yield of 100-150 mg for *Xenopus* H2A, H2B, and H3 or 60-80 mg for H4 from 6 liters of cells. The pool was dialyzed against 2 L of 5 mM β-mercaptoethanol with at least 4 changes to remove urea. The dialyzed sample was then lyophilized in a round-bottomed flask and the lyophilized histones were stored in -20°C until subsequent steps.

#### 2.2.3 Alkylation of histone H3

The dimethylated histone H3 at Lys4 with methyl-lysine analog was prepared by chemically modifying a site-directed mutagenized cysteine residue to mimic methylated H3K4 (Simon *et al.*, 2007). In particular, the recombinant H3K4C protein was alkylated with 2-chloro-*N*,*N*-dimethylethylamine hydrochloride (Aldrich) to produce *N*,*N*-dimethylated aminoethylcysteine modified H3 previously shown to be recognized by H3K4 specific antibodies (Simon *et al.*, 2007).

Ten mg of purified and lyophilized histone H3Kc4 was dissolved in 930  $\mu$ l of alkylation buffer (1 M HEPES pH 7.8, 4 M guanidine-HCl, and 10 mM D/L methionine) and 20  $\mu$ l of freshly prepared 1 M DTT (dithiothreitol) was added, followed by an incubation at 37°C for 1 hour. All the subsequent steps were performed in the dark. Fifty  $\mu$ l of fresh 1 M 2-chloroethyldimethylammonium chloride was added to the histone and incubated for 2 hours at room temperature. After the incubation, 10  $\mu$ l of fresh 1 M DTT was added and incubated for 30 minutes at room temperature. Second alkylation of the histone was performed by addition of 50 ul of fresh 1 M 2-chloroethyl-dimethylammonium chloride and incubation for 2 hours in the dark at room temperature. The reaction was quenched with 50  $\mu$ l of 14.4 M  $\beta$ -mercaptoethanol.

The alkylated histone was loaded on to the S300 column pre-equilibrated with 5 mM  $\beta$ mercaptoethanol and the pooled fractions were lyophilized. Alternatively, the alkylated histones were dialyzed into H100 and was purified over SourceS cation-exchange column. The pooled fractions were dialyzed again into 5 mM  $\beta$ -mercaptoethanol, lyophilized, and stored at -80°C. The alkylation of histone H3 was confirmed by LC-MS analysis (data not shown).

### 2.2.4 Fluorescent labeling of recombinant histones

Histones engineered to contain site-specific cysteine mutations were expressed and purified identically to the wild-type histones. Histone H3 mutants were created in the H3 (C110A) background to ensure unique labeling at the engineered position.

The ~30 mg of the lyophilized histone was dissolved in 20 ml of the labeling buffer (20 mM Tris-Cl pH 7.5, 0.25 mM Tris (2-carboxyethyl)phosphine (TCEP), and 7 M guanidine-HCl) and incubated at room temperature for at least 20 minutes. The dissolved histone was treated with 100  $\mu$ l of 10 mM Oregon Green 488 maleimide dye (LifeTech) and incubated at 4°C in the dark. After 2 hours, additional 100  $\mu$ l of 10 mM Oregon Green 488 was added and incubated at 4°C in the dark overnight. The labeling reaction was quenched by adding 50  $\mu$ l of 1M DTT. An equivalent molar amount of the complementary pair histone for either the tetramer or dimer was added to the labeled histone and dialyzed against H100 (10 mM HEPES pH 7.5, 100 mM NaCl, and 10 mM  $\beta$ -mercaptoethanol). The reconstituted tetramer or dimer was further purified by SourceQ anion-exchange chromatography. The pooled fractions were concentrated using Vivaspin 4, (10,000 MWCO, Sartorius) and glycerol was added 20% w/v before the sample was flash frozen and stored in -80°C. Typical labeling efficiencies of 10–30% were observed.

#### 2.2.5 Preparation of nucleosomal DNAs

Plasmid construct with the desired Widom 601 DNA sequence was transformed into HB101 cells, and plated onto the TYE plate containing 100 µg/ml ampicillin. After incubation at

 $37^{\circ}$ C for ~20 hours, four colonies were restreaked onto TYE plate with 100 µg/ml ampicillin and incubated at  $37^{\circ}$ C for ~20 hours again. Starter culture with 11 ml of 2xTY media with 50 µg/ml ampicillin was inoculated with 3-5 colonies from the restreaked plate and grown in the  $37^{\circ}$ C shaking incubator until the OD600 was between 0.1 and 1.0. Large scale expression was typically performed with 12 L of TB (Terrific Broth) media. Two ml of the starter culture was used to inoculate each 450 ml of TB media with 50 ml of salt solution and 50 µg/ml ampicillin. Cells were grown for ~22 hours in  $37^{\circ}$ C shaking incubator until the OD600 rotor at 7,000 rpm at room temperature. The cells were stored as pellets in 6 x 500 ml centrifuge bottles at -20 °C.

The frozen cells were thawed by the addition of 50 ml of LYSIS buffer (25 mM Tris-Cl pH 8.0, 10 mM EDTA, and 50 mM glucose) and incubation in the 30°C water bath. The cells were resuspended, and 120 ml of NaOH/SDS solution (0.2 M NaOH and 1% SDS) was added to each bottle. The cells were mixed by vigorous shaking of the bottles and incubated on ice for 10-20 minutes. Ninety ml of prechilled 5 M KAc/2.5 M HAc was added to each bottle and shaken vigorously, followed by an incubation on ice for 20 minutes. After the incubation, the suspension was centrifuged at 10,000 rpm in the SLA-3000 rotor for 5 minutes at 4°C. Next, the supernatant was filtered through Miracloth and a large scintered glass funnel into a 2 L Erlenmeyer flask. The resulting yellow supernatant was transferred into a 4 L beaker and 850 ml of isopropanol was added and mixed with the supernatant. After 15 minutes of incubation at room temperature, the sample was centrifuged at 10,000 rpm for 30 minutes at 20°C. The supernatant was poured and aspirated off, and the pellets were rinsed with 70% ethanol and air dried for 10 minutes. The pellets were then resuspended in a total of 60 ml TE (10, 50, pH 8.0) and transferred to three 50 ml Falcon tubes. The suspension was treated with a total of  $120 \,\mu$ l of  $10 \,\text{mg/ml}$  RNase A and incubated at 37°C for 45 minutes. Additional 60 µl of RNase A solution was added, followed by incubation at 37°C for 45 minutes. Each suspension was extracted with 20 ml of phenol

(equilibrated with TE) by vortexing for 30 seconds two times with 2-minute interval in between. The suspension was centrifuged at 3,220 x *g* for 15 minutes at 20°C. The phenol extraction was repeated once more, followed by two chloroform extractions. The aqueous phase from the chloroform extraction was distributed between two 50 ml polypropylene centrifuge tubes. To each tube, TE (10, 50, pH 8.0) was added to a final volume of 30 ml and subsequently, 4.5 ml of 5 M NaCl and 3.42 g of PEG 6000 were added. The samples were mixed at 37°C for 5 minutes and incubated on ice for 30 minutes. The samples were centrifuged at 15,000 rpm in the SS34 rotor for 10 minutes at 4°C. The supernatant was poured and aspirated off, and the pellets were dissolved in a total of 30 ml TE (10,1, pH 8.0) and distributed between two 50 ml Falcon tubes. The suspension was extracted with an equal volume of chloroform two times. The aqueous phase was transferred into the 50 ml polypropylene tubes and mixed with 0.1 volumes of 3 M NaAc pH 5.2 and 2.5 volumes of 95% ethanol and centrifuged at 15,000 rpm in the SS34 rotor for 20 minutes at 20°C. The supernatant was poured off and the pellet was resuspended in ~10 ml of TE (10, 0, 1, pH 8.0) to give ~10 mg/ml plasmid DNA.

The plasmid was digested with EcoRV restriction enzyme in buffer containing 50 mM Tris-Cl pH 7.6, 100 mM NaCl, 10 mM MgCl<sub>2</sub>, and 1 mM DTT with 1 mg/ml EcoRV. The digest was performed in the 37°C water bath overnight. The undigested sample was saved prior to EcoRV addition. After confirming that the plasmid was completely digested by running the undigested and digested samples on a native PAGE (10% acrylamide, 40:1 acrylamide: bisacrylamide, 0.5x TBE), the vector-derived DNA was precipitated by addition of 0.154 volume of 50% PEG 6000 and 0.128 volume of 5 M NaCl. After incubation on ice for 1 hour, the samples were centrifuged at 15,000 rpm in the SS34 rotor for 20 minutes at 4°C. The supernatant, which contained the desired nucleosomal DNA, was transferred to a 50 ml Falcon tubes, followed by an addition of 2.5 volumes of absolute ethanol. The sample was incubated at room temperature for 15 minutes and centrifuged at 15,000 rpm in the SS34 rotor for 20 minutes at 20°C. The supernatant was poured and aspirated off and the pellet was air dried until there was no more ethanol visible. The pelleted DNA was resuspended with TE (10, 0.1) to produce about 10 mg/ml DNA, and the concentration of DNA was quantified. The nucleosomal DNA was further purified by SourceQ anion-exchange high-performance liquid chromatography. The fractions were analyzed by a native PAGE. The desired fractions were pooled, added with 2.5 volumes of absolute ethanol, and incubated at room temperature for 15 minutes. After the incubation, the sample was transferred into a 50 ml polypropylene tube and centrifuged at 15,000 rpm in the SS34 rotor for 15 minutes at 20°C. The centrifugation was done multiple times in the same polypropylene tube. After the last centrifugation, the supernatant was removed by aspiration, and the pellet was air dried until there was no more ethanol visible. The DNA pellet was then resuspended in TE (10,0.1) to give a final concentration of about 10 mg/ml and aliquoted into the 1.5 ml Eppendorf tubes for storage at -80°C.

#### 2.2.6 Nucleosome reconstitution and purification

The histone dimer or tetramer was reconstituted and refolded from the individual core histones and then purified over a SourceS cation-exchange column. All DNAs used for nucleosome reconstitution containing the Widon 601 positioning sequence with variable lengths of DNA extensions were expressed and purified as described above (Dyer *et al.*, 2004). Nucleosomes were reconstituted by mixing histone H2A/H2B dimer, the histone H3/H4 tetramer, and Widom 601 nucleosome DNA at the molar ratio of 2:1:1 for mononucleosomes, and 4:2:1 for dinucleosomes in RB-high buffer (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 1 mM DTT, and 2 M KCl). The reconstituted samples were dialyzed against RB-low buffer (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 1 mM DTT, and 250 mM KCl) (350 mM KCl was used instead of 250 mM KCl for dinucleosomes) by a peristaltic pump. The reconstituted nucleosomes were further purified by SourceQ anion-exchange HPLC. The desired fractions were pooled and dialyzed into NCP storage buffer (10 mM potassium cacodylate pH 6.5 and 0.1 mM EDTA) followed by concentration with Vivaspin centrifugal concentrator (10,000 MWCO, Sartorius). The quantitated nucleosomes were stored at 4°C.

#### 2.3 Biochemical Characterization of LSD1/CoREST/Nucleosome Complex

## 2.3.1 Amplex UltraRed demethylase assay

Nucleosome substrates used in this assay contained a methyl-lysine analog of dimethylated H3K4. LSD1/CoREST activity on nucleosomes with various DNA lengths was measured by a peroxidase-coupled assay, which monitors hydrogen peroxide production (Forneris *et al.*, 2005). Amplex<sup>®</sup> UltraRed (Invitrogen), which is colorless and nonfluorescent, produces highly fluorescent resorufin upon oxidation (Zhou *et al.*, 1997). The 100 µl reactions in a 96-well microplate (Greiner) were initiated by the addition of 10 µl of 500 nM LSD1/CoREST solution to reaction mixtures (90 µl) consisting of 20 mM HEPES-NaOH (pH 7.5), 50 mM NaCl, 10 µM Amplex<sup>®</sup> UltraRed, and 0.76 µM horseradish peroxidase. Fluorescence changes were monitored with excitation at 530 nm and emission at 590 nm using a BioTek Synergy H4 Hybrid Multi-Mode Microplate Reader. The initial velocity values were fitted to the Michaelis–Menten equation using GraphPad Prism software to determine  $V_{max}$  and  $k_{cat}$  values. At 50, 150 and 200 nM of the limiting nucleosome substrates, these initial velocity values were determined using fluorescence data up to when 26%, 25% and 13% of 147 bp nucleosomes were consumed respectively. The equivalent percentages were 38%, 31% and 25% for 157 bp nucleosomes, 13%, 19% and 17% for 165 bp nucleosomes and 57%, 34% and 32% for 185 bp nucleosomes.

#### 2.3.2 Demethylase assay by Western Blot

Nucleosome substrates containing the methyl-lysine analog of dimethylated H3K4 were diluted in the demethylase reaction buffer containing 20 mM HEPES pH 7.5 and 50 mM NaCl at 0.2  $\mu$ M concentration unless otherwise stated. An aliquot of the sample was saved prior to addition of LSD1/CoREST enzyme. LSD1/CoREST complex was added to a final concentration

of 0.2  $\mu$ M and mixed by gentle pipetting. Assays were quenched with equivalent volumes of SDS-PAGE gel loading dye at desired time points.

The demethylase activity of the LSD1/CoREST enzyme with various nucleosome substrates was evaluated as follows: samples with different nucleosome substrates at various time points were run on the 18% SDS-PAGE gel. The gel was soaked in 30 ml of the transfer buffer (25 mM Tris-Cl pH ~8.3, 192 mM glycine, 20% methanol, and 1% SDS) for ~3 minutes. The transfer buffer included 1% SDS to assist in the transfer of histones to the nitrocellulose membrane (GE Healthcare). The nitrocellulose membranes were prewetted in the transfer buffer as well. The nitrocellulose membrane was layered with the SDS-PAGE gel, sandwiched between the blotting papers and fiber pads. Blotting was performed at 100 V for 60 minutes in the cold room with constant stirring. After the transfer, the nitrocellulose membrane was equilibrated in 30 ml of TBS buffer (25 mM Tris-Cl pH 8.0, and 150 mM NaCl) at room temperature for 5 minutes. The nitrocellulose membrane was blocked by incubating in the preincubation buffer (25 mM Tris-Cl pH 8.0, 150 mM NaCl, and 2% nonfat dry milk) for 30 minutes at room temperature. The membrane was washed twice in 50 ml of TTBS (25 mM Tris-Cl pH 8.0, 150 mM NaCl, and 0.05% Tween 20) for 5 minutes each to remove excess preincubation buffer. The washed membrane was incubated in 10 ml of TTBS buffer containing 2 µl of rabbit antibody for H3K4me2 on the rocker at room temperature for 1 hour. The membrane was washed three times for 5 minutes in 50 ml of TTBS and incubated in 20 ml of TTBS with 2µl of secondary antibody (donkey anti-rabbit antibody conjugated to horseradish peroxidase) on a rocker at room temperature for 1 hour. The membrane was washed again three times for 5 minutes in 50 ml of TTBS. The ECL detection solution 1 and 2 were mixed in 1:1 ratio and pipetted onto the protein side of the membrane. After one minute, the membrane wrapped in Saranwrap was exposed to film, typically for 10-20 seconds.

### 2.3.3 HI-FI nucleosome binding assay

LSD1/CoREST's interaction with nucleosomes of different DNA lengths was measured by the HI-FI (high-throughput interactions by fluorescence intensity) procedure as described (Hieb *et al.*, 2012, Winkler *et al.*, 2012).

A 384-well microplate (VWR) was passivated by filling all wells with 1% Hellmanex solution using a wash bottle and incubating for 20 minutes at room temperature. The Hellmanex solution was discarded and all wells were thoroughly rinsed with MiliQ water four to five times. Next, 1 M KOH was used to fill all the wells using a wash bottle, and the plate was incubated for 20 minutes. The 1 M KOH solution was discarded from the plate and the plate was rinsed with MiliQ water four to five times. The plate was dried overnight under an exhaust hood. Next, 100 µl of a solution containing 2% 1,7-dichlorooctamethyltetrasiloxane and 98% heptane was added to each well of the microplate using a multichannel pipette and incubated for 1 minute. The plate was processed in sections to ensure consistent incubation times. The organic solution was discarded from the plate and all wells were rinsed with MiliQ water thoroughly four to five times.

Site-specific Oregon Green 488 labeled nucleosomes were diluted from a stock concentration into a single dilution of two-fold over the final concentration used in the assay (typically 0.2-20 nM) in the HI-FI buffer (20 mM Tris–HCl, pH 7.6, 5 mM DTT, 5% glycerol, 0.01% NP40, 0.01% 3-CHAPS ([(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate), 100 µg/ml BSA (bovine serum albumin), 2 mM EDTA (ethylenediaminetetraacetic acid), and either 50 or 75 mM NaCl). A titration series of LSD1/CoREST of 14 different concentrations ranging from 0.5 nM to 40 µM (2x final concentration) were prepared in the same HI-FI buffer. Twenty 20 µl of the LSD1/CoREST titration series was added into each microplate well. Typically, 0 µM LSD1/CoREST samples were included in the first and the last wells. The same amount of the labeled nucleosome was added to each well. The plate was spun down at 500 rpm

for 1 minute in the tabletop centrifuge, vortexed for 2 minutes, and incubated at room temperature for ~15 minutes before the scan. Fluorescence was monitored with excitation at 488 nm and emission at 526 nm by Typhoon 9410 scanner (GE Healthcare). Typically the assays were performed at least two times and the dissociation constants were determined using the single binding isotherm equation 11.1 in Winkler *et al.* Initial experiments indicated Hill coefficients of 1.18, 1.08, 1.01 and 0.98 for 147, 157, 181 and 207 bp nucleosomes, respectively at 50 mM NaCl, and therefore used a Hill coefficient of 1.0 was used for subsequent analysis unless otherwise stated.

I selected nucleosome positions H3 K27C and H3 A21C based on the largest fluorescence changes using near saturating concentrations of LSD1/CoREST at 50, 75 and 100 mM NaCl (Figure 3-3). I observed reduced fluorescence change with increasing extranucleosomal DNA from 147 to 157 to 181 bp for nucleosomes labeled on H2A T10C and H2B K20C, but the reverse for nucleosomes labeled on H3 A21C and H3 K27C. These findings are consistent with the probes on H2A T10C and H2B K20C detecting nonspecific binding of LSD1/CoREST to the nucleosome, binding that is reduced at higher stringency conditions of increased NaCl concentrations.

#### 2.3.4 HI-FI stoichiometry assay

Stoichiometric measurements were performed with 805 nM of 181 bp nucleosome labeled on H3 A21C at 100 mM NaCl. The nucleosomes were titrated with LSD1/CoREST at a ratio of 0.1:1 to 5:1 [LSD1/CoREST]:[nucleosome]. The two linear phases of the plot of normalized fluorescence change versus concentration ratio were each fit with a line. The binding stoichiometry corresponds to the concentration ratio at the intersection between the two linear phases of the plot (Hieb *et al.*, 2012, Winkler *et al.*, 2012).

#### 2.3.5 Photocrosslinking assay

The following protocol was kindly provided by Blaine Bartholomew.

#### 2.3.5.1 Synthesis of photoreactive phosphorothioate DNA probes

A series of 36 photoreactive phosphorothioate DNA probes were synthesized as described earlier (Persinger *et al.*, 1999) to scan LSD1/CoREST interactions with 147 bp of nucleosomal DNA and approximately 35 bp of linker DNA from the edge of the nucleosome. Of these 36 probes, 10 were positioned in the extranucleosomal DNA and the remaining 26 were positioned every 5 bp in the nucleosomal DNA with the photocrosslinker either facing towards or away from the histone octamer. Eighteen to twenty-mer oligonucleotide primers complementary to these 36 positions in the extranucleosomal and nucleosomal DNA were commercially synthesized with a phosphorothioate moiety incorporated between the third and fourth nucleotides from the 5' end. The oligonucleotide primers were incubated at 25°C for 1 hour with APB (*p*-azido phenacyl bromide) in the dark to attach the photocrosslinker to the phosphorothioate moiety through a 7 Å tether. Chemically synthesized phosphorothioate oligonucleotides exist as a racemic mixture and therefore allowed the conjugated photocrosslinker to project towards both major and minor grooves of DNA.

The template DNA for photoreactive probe synthesis was prepared by digestion of the plasmid p159-1-G4 27-Xba1 with NdeI, incorporating biotinylated nucleotides, Bio-14-dATP (Invitrogen) and Bio-11-dUTP (Enzo) with Klenow exo<sup>-</sup> DNA polymerase, and HindIII digestion followed by immobilizing the biotinylated DNA on Dynabeads M-280 Streptavidin (Invitrogen). Non-biotinylated DNA strand was removed by denaturation with 0.1 M sodium hydroxide. The

bead-bound single-stranded template DNA was dephosphorylated with calf intestinal alkaline phosphatase and employed for photoreactive probe synthesis.

The APB modified oligonucleotides generated earlier were 5'-end radiolabeled and annealed to the single stranded template. The annealed oligonucleotides were extended by T4 DNA Polymerase in presence of all four dNTPs, and also T4 DNA ligase and ATP added simultaneously to the reaction to seal the nicks after extension. Two hundred and twenty-three base pairs photoreactive phosphorothioate DNA probes were released from the beads by digestion with EcoRI restriction enzyme.

#### 2.3.5.2 Synthesis of photoreactive nucleosomes

Nucleosomes (34N42) containing the 601 positioning sequence (Lowary and Widom, 1998) flanked by 34 and 42 bp of extranucleosomal DNA were assembled separately in the dark with the 36 different photoreactive DNA probes, salmon sperm carrier DNA and recombinant *Xenopus laevis* histone octamer at 30°C by stepwise salt dilution from 2 M to 280 mM NaCl and analyzed by 4% (35:1 acrylamide to bisacrylamide) native polyacrylamide gel electrophoresis.

## 2.3.5.3 Site-specific DNA photoaffinity crosslinking

Each of the 36 different 34N42 photoreactive and radiolabeled nucleosomes (100 nM) were separately incubated with or without LSD1/CoREST (860 nM) at 30°C for 30 min in a 12.5  $\mu$ l reaction containing 20 mM HEPES–NaOH (pH 7.8), 3 mM MgCl2, 6% (v/v) glycerol, 95 mM NaCl and 0.1  $\mu$ g/ $\mu$ l BSA. Nucleosome binding by LSD1/ CoREST was assessed by loading 2  $\mu$ l of the binding reaction on a 4% (79:1 acrylamide to bis-acrylamide) native polyacrylamide gel.

Remaining binding reactions were then subjected to 2 min UV irradiation (310 nm, 2.65 mW/cm<sup>2</sup>, at the distance of 8 cm) to crosslink protein subunits in close vicinity of the photocrosslinker with radiolabeled DNA followed by digestion with 4.6 units of DNase I (Ambion) at 30°C for 15 min. SDS was added to a final concentration of 0.4% and samples heated at 90°C for 3 min. Heating in the presence of SDS released DNA from the histone octamer. The crosslinked protein–DNA samples were cooled on ice, brought to room temperature, supplemented with 1 mM zinc acetate and further digested with 20 units of S1 nuclease (USB) at 30°C for 15 min. Digestion with DNase I and S1 nuclease removed excess DNA to reduce altering the mobility of crosslinked protein subunits on SDS-PAGE. The nuclease digested samples were analyzed by 4–20% gradient Tris–glycine SDS-PAGE followed by Western transfer on a nitrocellulose membrane at 80 V for 6 hours at 4°C. The LSD1 and CoREST subunits crosslinked at different positions in the nucleosomal and extranucleosomal DNA were visualized by autoradiography by exposing the membrane to a screen and scanning it using a phosphorimager (Fuji).

## 2.3.6 Ubiquitylation assay

Ubiquitylation assays were performed by mixing 1  $\mu$ M nucleosome or dinucleosome, 4  $\mu$ M ubiquitin, 30 nM Uba1 (E1), 375 nM UbcH5c (E2), and 375 nM Ring1B(2–116)–Bmi1(2–109) (E3) in 50 mM HEPES, pH 7.5, 75 mM NaCl, 1 mM DTT, 10 mM ZnSO4, 2mM MgCl2, and 3mM ATP. Assay samples were placed in a 30°C water bath for 12 hours, after which the assays were quenched with equivalent volumes of SDS-PAGE gel loading dye and boiled for 5 minutes. The assay samples were separated by SDS–PAGE using 18% gels and stained with Coomassie brilliant blue for visualization.

#### 2.4 Structural Characterization of LSD1/CoREST/Nucleosome Complex

## 2.4.1 Reconstitution of LSD1/CoREST/nucleosome complex

The LSD1/CoREST complex was reconstituted with nucleosomes prepared with 601 DNA of various lengths at a 2.5-3.4:1 enzyme:substrate ratio in 10 mM Tris-Cl pH 7.5, 75 mM NaCl, and 1 mM DTT. In particular, the LSD1/CoREST complex was added to nucleosome in four separate equimolar additions with 5 minute interval at room temperature. The solution was mixed by gentle pipetting after each addition. The complex was purified by Superdex 200 increase 10/300 column (GE Healthcare) in the reconstitution buffer supplemented with 0.1 mM phenylmethylsulphonyl fluoride (PMSF). In most cases, the chromatographic peak of LSD1/CoREST/nucleosome complex was sharp and symmetric, indicating that the complex was homogenous and apt for crystallization (McGinty *et al.*, 2016). The peak of the unbound LSD1/CoREST was also present behind the LSD1/CoREST/nucleosome peak, which resulted from the excess LSD1/CoREST complex added in reconstitution. The A<sub>260</sub>/A<sub>280</sub> ratio was plotted and a flat line across the LSD1/CoREST/nucleosome complex peak indicated that the complex is homogeneous and saturated (McGinty *et al.*, 2016). In addition, the A<sub>498</sub> was monitored during the run as the FAD cofactor in LSD1 absorbs at 498 nm, and alignment of the A<sub>498</sub> peak could serve as another way to assess the quality of the LSD1/CoREST/nucleosome complex.

Fractions were further analyzed by the SDS-PAGE and the desired fractions were pooled (typically 2-3 fractions of 0.4 ml/fraction). Additional 0.1 mM PMSF was added to the pool to minimize proteolysis of the complex and the pooled fractions were concentrated to ~10 mg/ml by Vivaspin 4 (10,000 MWCO, Sartorius) centrifuge concentration device. The concentration was calculated based the absorption at 260 nm of nucleosomal DNA in 0.2N NaOH. In most cases, the purified LSD1/CoREST/nucleosome complex was used immediately for crystallization trials.

## 2.4.2 Crystallization of LSD1/CoREST/nucleosome complex

Crystallization screens were performed in modified microbatch by combining 1  $\mu$ l of the concentrated LSD1/CoREST/nucleosome complex with 1  $\mu$ l of screening solution, overlayed by 70 ml Al's oil [1:1 mixture of silicon oil (Clearco) and mineral oil (Fisher)] at either 4 or 21°C in 96-well microtitre plates. The screens contain solutions of the buffer with a range of pH (5.0 – 8.5), various monovalent ions of different concentrations (75 mM, 100 mM, and 150 mM), and precipitant of a range between 2–15% (typically PEG2000-MME was used). A fraction of solutions also contained divalent ions at millimolar concentrations.

The crystal screen trials were observed after 3, 7, 14, and 30 days after setup under the microscope. Crystals that grew in three dimensions with birefringence and sharp edges were optimized by varying the pH, salt concentration, and precipitant amount around the condition that produced the initial crystals. Optimization typically resulted in larger crystals.

## 2.4.3 Postcrystallization treatment

Postcrystallization soaks were performed to improve diffraction properties of crystals. First, a crystallization base soaking solution in which the crystal is stable was determined. This solution typically contained the buffer with the same pH, equivalent salt concentration, and 1.25 – 1.5 fold higher concentration of precipitant than the solution of the crystal drop. Crystals were transferred into the base solution with loops and observed for signs of dissolving or cracking, which was an indication that the base solution was inappropriate. Then the crystals were transferred to drops with gradually increasing concentration of the cryoprotectant and dehydration agent. Various cryoprotectants and dehydration agents were explored, including methylpentanediaol (MPD), glycerol, PEG400, PEG550-MME, PEG600, and PEG750-MME. Typically the crystals were soaked to a final concentration of 20% or 24% in 4% increments for 10 minutes either in the 4°C cold room or at room temperature. In some cases, a combination of cryoprotectants and dehydration agents was used (i.e. 20% PEG400 + 10% glycerol or 10% MPD + 20% glycerol). After the final soak, crystals were flashed cooled by plunging into liquid nitrogen.

#### **2.4.4 Data collection and structural determination**

Diffraction data were collected at Advanced Photon Source's NE-CAT beamline 24-ID-E or 24-ID-C. Data were processed using XDS and Scala (Kabsch, 2010, Evans, 2006), and the structure was solved by molecular replacement with Phaser (McCoy *et al.*, 2007). Molecular replacement was performed with combinations of polyalanine versions of search models: one NCP with 145 bp Widom 601 DNA (PDB ID: 3LZ0) and two LSD1/CoREST complex (PDB ID: 2VID) in the case of the dinucleosome substrate model, and one NCP with 145 bp Widom601 DNA (PDB ID: 3LZ0) and one LSD1/CoREST complex (PDB ID: 2VID) as two segments in the case of the mononucleosome substrate model. All molecular graphics were prepared with PyMOL software (Schrodinger, 2011).

#### **2.5 References**

- Dyer, P. N., R. S. Edayathumangalam, C. L. White, Y. Bao, S. Chakravarthy, U. M. Muthurajan and K. Luger. "Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA." *Methods Enzymol* 375, (2004): 23-44.
- Evans, P. "Scaling and Assessment of Data Quality." *Acta Crystallogr D Biol Crystallogr* 62, no. Pt 1 (2006): 72-82.
- Forneris, F., C. Binda, M. A. Vanoni, E. Battaglioli and A. Mattevi. "Human Histone Demethylase LSD1 Reads the Histone Code." J Biol Chem 280, no. 50 (2005): 41360-5.
- Hieb, A. R., S. D'Arcy, M. A. Kramer, A. E. White and K. Luger. "Fluorescence Strategies for High-Throughput Quantification of Protein Interactions." *Nucleic Acids Res* 40, no. 5 (2012): e33.
- Kabsch, W. "Xds." Acta Crystallogr D Biol Crystallogr 66, no. Pt 2 (2010): 125-32.
- Lowary, P. T. and J. Widom. "New DNA Sequence Rules for High Affinity Binding to Histone Octamer and Sequence-Directed Nucleosome Positioning." *J Mol Biol* 276, no. 1 (1998): 19-42.
- Luger, K., T. J. Rechsteiner and T. J. Richmond. "Expression and Purification of Recombinant Histones and Nucleosome Reconstitution." *Methods Mol Biol* 119, (1999): 1-16.
- McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read. "Phaser Crystallographic Software." *J Appl Crystallogr* 40, no. Pt 4 (2007): 658-674.
- McGinty, R. K., R. D. Makde and S. Tan. "Preparation, Crystallization, and Structure Determination of Chromatin Enzyme/Nucleosome Complexes." *Methods Enzymol* 573, (2016): 43-65.
- Persinger, J., S. M. Sengupta and B. Bartholomew. "Spatial Organization of the Core Region of Yeast TFIIB-DNA Complexes." *Mol Cell Biol* 19, no. 7 (1999): 5218-34.

Schrodinger, LLC. "The Pymol Molecular Graphics System." 1.4 (2011).

- Selleck, W. and S. Tan. "Recombinant Protein Complex Expression in E. Coli." *Curr Protoc Protein Sci* Chapter 5, (2008): Unit 5 21.
- Simon, M. D., F. Chu, L. R. Racki, C. C. de la Cruz, A. L. Burlingame, B. Panning, G. J. Narlikar and K. M. Shokat. "The Site-Specific Installation of Methyl-Lysine Analogs into Recombinant Histones." *Cell* 128, no. 5 (2007): 1003-12.
- Tan, S., R. C. Kern and W. Selleck. "The pST44 Polycistronic Expression System for Producing Protein Complexes in Escherichia Coli." *Protein Expr Purif* 40, no. 2 (2005): 385-95.
- Winkler, D. D., K. Luger and A. R. Hieb. "Quantifying Chromatin-Associated Interactions: The HI-FI System." *Methods Enzymol* 512, (2012): 243-74.
- Zhou, M., Z. Diwu, N. Panchuk-Voloshina and R. P. Haugland. "A Stable Nonfluorescent Derivative of Resorufin for the Fluorometric Determination of Trace Hydrogen Peroxide: Applications in Detecting the Activity of Phagocyte Nadph Oxidase and Other Oxidases." *Anal Biochem* 253, no. 2 (1997): 162-8.

# Chapter 3

# Extranucleosomal DNA enhances the activity of the LSD1/CoREST histone demethylase complex

Work presented in this chapter is a part of the published manuscript "Kim, S.A., Chatterjee, N., Jennings, M.J., Bartholomew, B. and Tan, S. (2015) Extranucleosomal DNA enhances the activity of the LSD1/CoREST histone demethylase complex. *Nucleic acids research*, **43**, 4868-4880."

## **3.1 Abstract**

Lysine-specific demethylase 1 (LSD1) is a FAD-dependent amine oxidase that regulates transcription by removing methyl groups from mono- and dimethylated lysine 4 of histone H3 (H3K4). LSD1 plays vital roles in cellular processes including neural stem cell proliferation, DNA damage response, and cancer proliferation and progression. LSD1 itself can only act on peptide substrates, but the association with its corepressor protein CoREST endows LSD1 with the ability to demethylate nucleosomal substrates. Although the previously reported crystal structure of LSD1 in complex with CoREST and H3-peptide substrate shows an intricate network of interactions between the substrate and LSD1, the mechanism by which LSD1/CoREST binds to and acts on nucleosomes remains largely unknown.

I have measured the catalytic activity and binding affinities of LSD1/CoREST on nucleosomes containing extranucleosomal DNA of different lengths. To my knowledge, this is the first time that kinetic analysis and dissociation constant measurements of LSD1/CoREST have been performed on nucleosomes. My studies of LSD1/CoREST's enzyme activity and nucleosome binding show that extranucleosomal DNA dramatically enhances the activity of LSD1/CoREST, and that LSD1/CoREST binds to the nucleosome as a 1:1 complex.

Photocrosslinking experiments performed by Nilanjana Chatterjee in collaboration with Blaine Bartholomew's group further indicate both LSD1 and CoREST subunits are in close contact with DNA around the nucleosome dyad as well as extranucleosomal DNA. These results suggest that the LSD1/CoREST interacts with extranucleosomal DNA when it productively engages its nucleosome substrate.
#### 3.2 Introduction

Lysine-specific demethylase 1 (LSD1) regulates transcription by removing methyl groups from mono- and dimethylated Lys of histone H3. LSD1's catalytic activity is contained in its FAD-dependent amine oxidase domain, and this domain is sufficient to bind to the N-terminal tail of histone H3 (Yang et al., 2006, Chen et al., 2006, Stavropoulos et al., 2006, Yang et al., 2007). While the LSD1 protein is able to catalyze demethylation of methylated H3K4 peptides, it is not able to catalyze the demethylation on methylated H3K4 incorporated into nucleosomes (Yang et al., 2006, Lee et al., 2005). The nucleosome specific enzymatic activity requires an additional protein, CoREST, which heterodimerizes with LSD1. The LSD1/CoREST complex forms an elongated structure with the LSD1 amine oxidase and SWIRM domains on one end, a CoREST SANT domain on the other, and a helical linker in between formed by both LSD1 and CoREST (Yang et al., 2006) (Figure 1-8c). The catalytic activity and H3 peptide binding properties of LSD1 and the LSD1/CoREST complex are relatively well understood due to extensive biochemical and structural studies (Hou and Yu, 2010, Culhane and Cole, 2007). In contrast, we possess little information regarding LSD1/CoREST's interactions with the nucleosome, its physiological substrate. In this chapter, I describe the published results of LSD1/CoREST's catalytic and binding activities on nucleosome substrates. I find that the LSD1/CoREST complex is catalytically more active and binds more tightly to nucleosomes containing extranucleosomal DNA. Furthermore, photocrosslinking studies suggest that LSD1/CoREST may bind to nucleosomal DNA around the nucleosomal dyad and on extranucleosomal DNA one to two helical turns beyond the end of the nucleosome core particle.

## **3.3 Results**

## 3.3.1 Steady state kinetic activity suggest a role for extranucleosomal DNA

My studies of LSD1/CoREST lysine demethylase activity on nucleosome substrates (MW > 200 kDa) required a more sensitive peroxidase assay than used previously with peptide substrates (MW = 2-3 kDa) (Stavropoulos *et al.*, 2006, Forneris *et al.*, 2005) due to the nearly 100 fold difference in substrate molecular weight. I therefore used a fluorescence-based demethylase assay developed by Matt Jennings, my predecessor on the LSD1 project. This assay employs the reagent Amplex UltraRed which is converted to the highly fluorescent resorufin dependent on hydrogen peroxide produced during lysine demethylation (Culhane et al., 2010) (Figure 3-1a). Recombinant nucleosome core particles containing a methyl-lysine (aminoethylcysteine) analog of dimethylated H3 Lys 4 (H3K<sub>c</sub>4) was reconstituted with 147 bp of the Widom 601 nucleosome positioning sequence. Since the 601 nucleosome core particle encompasses 145 bp of DNA (Makde et al., 2010, Vasudevan et al., 2010), this 147 bp 601 nucleosome core particle includes one additional base pair on either end of the central core particle, i.e. the 147 bp sequence constitutes a 1N1 positioning sequence where N represents the central 145 bp 601 nucleosome core particle sequence. My data from this assay appear to follow Michaelis-Menten kinetics. My results showed that LSD1/CoREST demethylates H3Kc4me2 147 bp nucleosomes with a  $k_{cat}$  of 0.315 min<sup>-1</sup> and a  $K_m$  of 339 nM (Figure 3-1b, Table 3-1). This represented a slower turnover than LSD1 protein (without CoREST) on a 21 amino acid H3K4me2 peptide ( $k_{cat} = 8.1 \text{ min}^{-1}$ ) or LSD1/CoREST on a 21 amino acid H3K4me1 peptide  $(k_{cat} = 7.4 \text{ min}^{-1})$ , and a smaller  $K_m$  than LSD1 or LSD1/CoREST (4.2  $\mu$ M for LSD1 with H3K4me2 peptide, 5.1 µM for LSD1/CoREST with H3K4me1 peptide) (Forneris et al., 2007, Forneris *et al.*, 2005). These peptide studies were performed using methylated lysine residues,

whereas my nucleosomes studies employed methylated aminoethylcysteine analogs of lysine. I do not expect the methylated aminoethylcysteine analogs to affect enzyme kinetic values significantly as it is structurally and chemically similar to lysine, only slightly lengthening the side chain by ~0.28 Å (Simon *et al.*, 2007).

Considering the comparatively lower  $k_{cat}$  of LSD1/CoREST on the 147 bp nucleosome substrate, the effect of lengthening the extranucleosomal DNA on the activity of the enzyme was examined. Nucleosomes containing the same preparation of histone H3K4me2 methyl-lysine analog were prepared, now with 157 (6N6), 165 (10N10) and 185 bp (20N20) nucleosomal DNA. In each case, the extranucleosomal DNA was extended symmetrically from the central 145 bp 601 sequence. While the enhanced catalytic activity had been suggested in previous experiments performed by Matt Jennings, it is the first time that the kinetic constants for nucleosomes with different DNA lengths were determined. I observed a slight increase in  $k_{cat}$  to 0.392 min<sup>-1</sup> for the 157 bp nucleosome substrate, and an apparent three-fold decrease of  $K_m$  (107 nM) (Figure 3-1b, Table 3-1).

Unfortunately, I was unable to reliably measure the activity using LSD1/CoREST enzyme concentrations or nucleosome substrate concentrations below 50 nM due to poor signal to noise ratios. Since steady state kinetics may not apply under these conditions due to significant concentrations of enzyme–substrate complex present, the concentration of unbound substrate is likely to be overestimated here. Consequently, this measured  $K_m$  probably represents an upper limit. More significant changes in  $k_{cat}$  were observed using nucleosomes with even longer DNA fragments. The 165 bp (10N10) nucleosomes which extend the DNA by 9 bp on each side compared to the 147 bp substrate increase  $k_{cat}$  almost 2-fold (0.717 min<sup>-1</sup> versus 0.315 min<sup>-1</sup>) while the 185 bp (20N20) nucleosomes increase  $k_{cat}$  more than 5-fold over the 147 bp nucleosomes (1.73 min<sup>-1</sup>) (Figure 3-1b, Table 3-1). The longer DNA also further decreased



**Figure 3-1.** Kinetic analysis of LSD1/CoREST demethylase on nucleosome substrates. (a) Schematic of Amplex Ultrared based fluorescence LSD1/CoREST demethylase assay. The hydrogen peroxide generated in the histone demethylase assay drives the horseradish peroxidase catalyzed conversion of Amplex UltraRed to the highly fluorescent resorufin molecule. (b) Michaelis–Menten analysis of LSD1/CoREST demethylase activity on nucleosomes with symmetrical extranucleosomal DNA. Kinetic data for 147 bp (1N1), 157 bp (6N6), 165 bp (10N10) and 185 bp (20N20) are shown in yellow, blue, green and red, respectively. Each data point was measured at least three times.

DNA length	DNA format	k <sub>cat</sub> (min⁻¹)	K <sub>m</sub> (M)	k <sub>cat</sub> /K <sub>m</sub> (M <sup>-1</sup> s <sup>-1</sup> )
147	1+ 145 +1	0.315 ± 0.022	339 ± 60x 10 <sup>-9</sup>	1.6 x 10 <sup>4</sup>
157	6+ 145 +6	0.392 ± 0.019	107 ± 20x 10 <sup>-9</sup>	6.1 x 10⁴
165	10+ 145 +10	0.717 ± 0.031	36.8 ± 10x 10 <sup>-9</sup>	3.3 x 10⁵
185	20+ 145+20	1.73 ± 0.049	32.3 ± 6.5x 10 <sup>-9</sup>	8.9 x 10⁵

**Table 3-1**. Kinetic constants for LSD1 complex using nucleosome substrates containing different linker DNA lengths.

The concomitant increase in  $k_{cat}$  and decrease in  $K_m$  results in a higher  $k_{cat}/K_m$  specificity constant for nucleosomes with longer extranucleosomal DNA. These results suggest specificity constants of 1.6 x 10<sup>4</sup>,  $6.1 \times 10^4$ ,  $>3 \times 10^5$  and  $>9 \times 10^5$  M<sup>-1</sup> s<sup>-1</sup> for nucleosomes containing 147, 157, 165 and 185 bp, respectively (Table 3-1). It is worth noting that because the same preparation of histone H3 modified with the K4 methyl-lysine analog was used to reconstitute each nucleosome, it is unlikely that these differences in the specific constants reflect differences in amount of methyl-lysine modification. Moreover, since the  $K_m$  determined for the 157, 165 and 185 bp substrates may be overestimates, the specificity constants for these substrates could be even higher. Thus, it appears that LSD1/CoREST is significantly more efficient on nucleosomes containing extranucleosomal DNA.

## 3.3.2 LSD1/CoREST binds more tightly to nucleosomes with extranucleosomal DNA

Since my steady state kinetics studies indicated that LSD1/CoREST was more efficient on nucleosomes with extranucleosomal DNA, I next compared LSD1/CoREST's ability to bind to nucleosomes with different lengths of extranucleosomal DNA. I employed the HI-FI nucleosome binding assay developed by Luger *et al.* (Hieb *et al.*, 2012, Winkler *et al.*, 2012) to measure the affinity of LSD1/CoREST for its nucleosome target. In this assay, one detects the quenching of a fluorescent dye installed on the nucleosome due to the binding of the chromatin protein to the nucleosome. This requires the dye to be located on a nucleosome position that is close enough to the chromatin protein to be detected without interfering with binding.

To scout for suitable nucleosome positions, histones containing Cys substitution at nine histone positions across the nucleosome surface were prepared: H2A Thr10, Asp72, Glu91, Ser113; on H2B Lys20, Ser120; on H3 Ala21, Lys27; and on H4 Gln27 (Figure 3-2). The Oregon Green 488 maleimide was conjugated to the individual Cys mutant histone protein, and each fluorescently labeled histone was then reconstituted into recombinant nucleosomes. I tested these fluorescently labeled nucleosomes in binding experiments with LSD1/CoREST complex at 50 mM NaCl and found little or no (<10%) change in fluorescence for nucleosomes labeled on H2A E91C, H2A S113C and H4 Q27C, suggesting that these positions are either located far away from LSD1/CoREST or alternatively, that fluorescent labeling at these positions inhibits binding of LSD1/CoREST to the nucleosome. Nucleosomes labeled on H2A T10C, H2A D72C, and H2B S120C produced measurable fluorescence changes (between 10 and 15%) when mixed with LSD1/CoREST complex. However, incubation of LSD1/CoREST with nucleosomes labeled at H2B K20C, H3 A21C, and H3 K27C resulted in large changes (>15%) in fluorescence. The fluorescent changes for nucleosomes labeled at H3 A21C and H3 K27C were particularly dramatic, with changes of 30–60% detected (Figure 3-3), more than the minimum of 10% recommended for quantitative studies (Winkler et al., 2012).

I selected the H3 K27C position to use in my LSD1/CoREST nucleosome binding assay. Nucleosomes labeled on H3 K27C (Figure 3-4a) were titrated with 5 nM to 10 μM LSD1/CoREST, and the normalized fluorescence change plotted as a function of the LSD1/CoREST concentration. I performed each titration at least three times in separate experiments using fresh dilutions of the LSD1/CoREST complex. The data in Figure 3-4 and Table 3-2 demonstrate that I was able to obtain reproducible results with typical standard deviations of 15% or less.



**Figure 3-2**. Fluorescence positions on the nucleosome affected by LSD1/CoREST binding. Oregon Green 488 was conjugated site-specifically to histones engineered with unique cysteines, reconstituted with appropriate histones and Widom 601 147 bp DNA and used in fluorescence quenching nucleosome binding experiments. The results are mapped on the space filling representation of the 1.9 Å crystal structure of the nucleosome core particle (PDB code 1KX5). Nucleosomes labeled on H2B K20, H3 A21, and H3 K27 showed >15% fluorescence quenching upon LSD1/CoREST binding (side chains shown in red), nucleosome labeled on H2A T10, H2A D72, and H2B S120 produced 10–15% fluorescence quenching (purple residues) while nucleosomes labeled on H2A E91, H2A S113 and H4 Q27 produced less than 10% fluorescence quenching (blue residues). The four H2A residues in the nucleosome acidic patch, E61, E64, D90, and E92, are shown in yellow. Figure prepared by Song Tan using PyMOL molecular graphics software (Schrodinger, 2011).



**Figure 3-3**. Fluorescence change for 147, 157, and 181 bp nucleosomes labeled on H2A T10C, H2B K20C, H3 A21C, and H3 K27C in the presence of near saturating concentrations of LSD1/CoREST at 50, 75 and 100 mM NaCl.



**Figure 3-4.** Extranucleosomal DNA increases nucleosome binding affinity of LSD1/CoREST. (**a**) Purified recombinant nucleosome core particles used for nucleosome binding assay as separated on native 10% acrylamide gel and visualized by ethidium bromide staining. (**b**) Binding curves for LSD1/CoREST interacting with 147 (yellow), 155 (blue) and 181 bp (pink) nucleosomes fluorescently labeled on H3 K27C. Three or more binding curves for individual experiments were performed for each nucleosome, with typical standard deviations of 15% of less for the calculated dissociation constants. The dashed vertical lines correspond to the calculated dissociation constants for each respective binding curve. (**c**) Equivalent binding curves for LSD1/CoREST interacting with 185 (red) and 207 bp (green) nucleosomes. LSD1/CoREST binds more weakly to the 207 bp nucleosomes compared to the 185 bp nucleosomes despite the longer extranucleosomal DNA. (**d**) Summary of dissociation constants for 13 nucleosomes containing 147–207 bp symmetrically positioned DNA. The standard deviations from three or more measurements are shown as vertical bars for each data point.

DNA length	histones	DNA format	K <sub>d</sub> (nM)	relative to 147 bp
147	WT	1+ 145 +1	1641 ± 211	1.0
151	WT	3+ 145 +3	716 ± 100	2.3
155	WT	5+ 145 +5	419 ± 23	3.9
157	WT	6+ 145 +6	470 ± 64	3.5
159	WT	7+ 145 +7	236 ± 40	7.0
161	WT	8+ 145 +8	232 ± 29	7.1
163	WT	9+ 145 +9	237 ± 22	6.9
165	WT	10+ 145+10	205 ± 26	8.0
169	WT	12+ 145+12	282 ± 68	5.8
177	WT	16+ 145 + 16	214 ± 35	7.7
181	WT	18+ 145 + 18	100 ± 27	16.2
185	WT	20+ 145+20	135 ± 27	12.1
207	WT	31+ 145+31	263 ± 27	6.2
151L	WT	5+ 145 +1	718 ± 31	2.3
161L	WT	15+145+1	406 ± 52	4.0
177L	WT	31+145+1	769 ± 174	2.1
177	acidic	16+ 145 + 16	98 ± 27	16.7
	patch			
	mutation			

**Table 3-2**. Dissociation constants for LSD1 complex binding to nucleosomes with different length linker DNA

I determined that LSD1/CoREST binds 147 bp nucleosomes with a dissociation constant of 1.64  $\mu$ M (Figure 3-4b, Table 3-2). To analyze the effect of extranucleosomal DNA on LSD1/CoREST binding, nucleosomes fluorescently labeled on H3 K27C were prepared using nucleosomal DNA lengths from 151 to 207 bp, each extended symmetrically about the central 601 bp sequence (Figure 3-4a). Extending the 147 bp DNA fragment by just 2 bp on either side to 151 bp (3N3) resulted in a  $K_d$  of 716 nM or ~2× greater affinity (Figure 3-4b, Table 3-2). Adding an additional 2 bp on either side produced a 155 bp nucleosomal DNA (5N5) and almost a 2-fold decrease in  $K_d$  to 419 nM, while the 159 bp DNA (7N7) further reduced the dissociation constant to 236 nM or almost seven times greater affinity compared to the 147 bp nucleosomes. Extending the DNA from 159 to 177 bp did not produce significant changes in the dissociation constant. I did observe a reproducible reduction in the  $K_d$  with 181 bp nucleosomes to 100 ± 27 nM (Figure 3-4b, Table 3-2), while LSD1/CoREST bound 185 and 207 bp nucleosomes with dissociation constants of  $135 \pm 27$  and  $263 \pm 27$  nM, respectively (Figure 3-4c, d, Table 3-2).

I validated the observation that extranucleosomal DNA increases the affinity of LSD1/CoREST using nucleosomes labeled on H3 A21C, the other labeling position that produced large changes in fluorescence upon LSD1/CoREST binding. I observed the same trend of increased binding affinity from 147 to 157 bp to 181 bp, but not the decreased binding affinity with 207 bp nucleosomes (Table 3-2 and Supplementary Figure 3-1). Surprisingly, the measured apparent dissociation constants using nucleosomes labeled on H3 A21C were ~20-fold lower than using nucleosomes labeled on H3 K27C. The simplest explanation is that the fluorescent probe installed on H3 K27C inhibits LSD1/CoREST binding, or the probe installed on H3 A21C enhances LSD1/CoREST binding, or some combination of the two occurs. While the precise dissociation constants presented here should therefore be viewed with some skepticism, consistent trends in my results establish an important role of extranucleosomal DNA for LSD1/CoREST binding to the nucleosome.

# **3.3.3 LSD1/CoREST** prefers nucleosomes with extranucleosomal DNA on both sides of the nucleosome

To examine the possibility that LSD1/CoREST interacts with extranucleosomal DNA on only one side of the nucleosome, I assayed binding to nucleosomes containing extranucleosomal DNA extensions on one side of our standard 147 bp nucleosomes (Figure 3-5a, Table 3-2). The 151L nucleosomes (5N1) extends the DNA by 4 bp on the left side and decreases the  $K_d$  from 1641 to 718 nM, similar to the 716 nM dissociation constant observed for the symmetrical 151 bp nucleosome of the same DNA length. However, adding an additional 10 bp on the left side (161L = 15N1) further decreases the  $K_d$  to 406 nM versus 232 nM for the symmetrical 161 bp nucleosomes. These results suggest that LSD1/CoREST requires extranucleosomal DNA on both sides of the nucleosome for maximal binding.

Surprisingly, extending the nucleosomal DNA an additional 16 bp in the 177L (31N1) nucleosomes increased the  $K_d$  to 769 nM, or slightly weaker affinity than the 151L (5N1) nucleosomes despite a substantially longer extranucleosomal DNA (Figure 3-5a, Table 3-2). This decreased affinity mirrors the result for the symmetrical 207 bp nucleosomes (31N31) which bound LSD1/CoREST less tightly than the symmetrical 181 or 185 bp nucleosomes. A possible interpretation of these results is discussed below.

## 3.3.4 Stoichiometry of LSD1/CoREST binding to the nucleosome

I have determined the stoichiometry of LSD1/CoREST binding to the nucleosome using the HI-FI stoichiometry assay described by Winkler *et al.* (Winkler *et al.*, 2012). LSD1/CoREST was titrated with the high affinity 181 bp symmetrically positioned nucleosomes fluorescently labeled on H3 A21C. I used H3 A21C label position instead of H3 K27C because of the higher binding affinity detected with this position, which allowed me to determine the stoichiometry of LSD1/CoREST binding to 181 bp nucleosomes at 100 mM NaCl where nonspecific binding to the nucleosome is reduced. The concentration of the labeled nucleosome is kept at ~5-fold above the experimentally determined dissociation constant so that any titration of LSD1/CoREST complex results in immediate binding as indicated by the pink data points (Figure 3-5b). Binding continues until all available binding sites on the nucleosome are saturated, resulting in the flattening of the fluorescence. The plot of normalized fluorescence change as a function of the LSD1/CoREST:nucleosome concentration ratio shows data that can be fit with two lines which intersect at a binding ratio of 0.972 or ~1.0 (Figure 3-5b). Therefore I conclude that LSD1/CoREST binds to the 181 bp nucleosome with a 1:1 stoichiometry. This is consistent with the recent determination by SAXS that LSD1/CoREST binds to 146 bp nucleosome core particles as a 1:1 complex (Pilotto *et al.*, 2015).



**Figure 3-5.** Effect of extranucleosomal DNA on one side of the nucleosome on binding affinity to LSD1/CoREST and stoichiometry of LSD1/CoREST binding to the nucleosome. (**a**) Binding curves for LSD1/CoREST interacting with 147 (1N1, yellow), 151L (5N1, blue), 161L (15N1, green) and 177L (31N1, red) nucleosomes fluorescently labeled on H3 K27C. Three or more binding curves for individual experiments were performed for each nucleosome. The dashed vertical lines correspond to the calculated dissociation constants for each respective binding curve. (**b**) Stoichiometry plot for LSD1/CoREST binding to 181 bp nucleosomes fluorescently labeled on H3 A21C. LSD1/CoREST was titrated with a fixed concentration of nucleosomes and the normalized fluorescence change plotted as a function of the ratio of [LSD1/CoREST] to [nucleosomes]. The plot indicates a stoichiometry ratio 0.972 LSD1/CoREST molecules per nucleosome.

#### 3.3.5 Photocrosslinking studies of LSD1/CoREST on nucleosome

The steady state kinetics and nucleosome binding assays suggest that LSD1/CoREST interacts with extranucleosomal DNA. These results were consistent with the photocrosslinking studies performed by Blaine Bartholomew and colleagues prior to my biochemical studies. The photocrosslinking studies provide insights into which subunit of the LSD1/CoREST complex is in close proximity to nucleosomal DNA. In this study, they used a library of 38 nucleosomes each containing a photoactivatable crosslinking group incorporated along the DNA phosphate

backbone close to a radiolabel (Hota and Bartholomew, 2012). Each nucleosome contained the same 221 bp DNA sequence with the 145 bp 601 nucleosome positioning sequence located between a 34 bp extension on one side, and 42 bp on the other (34N42). The library consists of nucleosomes with an aryl azide photoreactive group incorporated approximately every 5 bp within the central 145 bp and every 3 bp in the extranucleosomal DNA. The LSD1/CoREST complex was incubated with each nucleosome, exposed to UV irradiation to crosslink the appropriate protein subunit to the DNA, treated with nucleases to remove excess DNA, separated by SDS-PAGE and visualized by autoradiography (Figure 3-6). As expected, crosslinks to the histone proteins are detected at many positions within the nucleosome core. In addition, crosslinks to both LSD1 and CoREST proteins are detected at three distinct positions. Strong crosslinks to CoREST and weaker crosslinks to LSD1 are observed near the nucleosome dyad, while some crosslinks are observed ~30–40 bp on one side of the dyad corresponding to the opposite edge of the nucleosome core particle relative to the dyad.

The strongest crosslinks to both LSD1 and CoREST occur on the extranucleosomal DNA. CoREST crosslinks strongly to extranucleosomal DNA approximately one helical turn from the end of the nucleosome core particle and more weakly one turn beyond. LSD1 also crosslinks to these locations, but with reversed intensities (stronger to DNA two helical turns from nucleosome core particle end, weaker to DNA one turn from core particle end). When modeled on the three-dimensional structure of the nucleosome, the LSD1 and CoREST crosslinks to the nucleosomal dyad and extranucleosomal DNA align to create a potential contiguous interaction surface or face (Figure 3-7).



**Figure 3-6**. Photocrosslinking of LSD1/CoREST to site-specifically labeled nucleosomes. (**a**) Autoradiograph of SDS-PAGE gel shows LSD1 and CoREST subunits crosslinked to nucleosomes site-specifically modified with an aryl azide photoreactive group from -102 to -17 bp from the nucleosome dyad. The base pair positions from the dyad are shown in gray, while the base pair positions from the 145 bp nucleosome core particle end are shown in black above the autoradiograph. The migration of LSD1, CoREST and histone proteins are shown on the left. (**b**) Equivalent autoradiograph for photocrosslinking of LSD1/CoREST to nucleosomes site-specifically modified with a photoreactive group from -17 to +68 bp from the nucleosome dyad. Experiment performed by Nilanjana Chatterjee, a former member of Blaine Bartholomew laboratory.

# **3.3.6 LSD1/CoREST** does not employ an arginine anchor to interact with the nucleosome acidic patch

Crystal structures of four chromatin proteins (the RCC1 chromatin factor (Makde et al.,

2010), the Sir3 silencing protein (Armache et al., 2011), the Polycomb PRC1 ubiquitylation

complex (McGinty et al., 2014), and SAGA deubiquitylation (DUB) module (Morgan et al.,

2016)) and two peptides (the herpesviral LANA peptide (Barbera et al., 2006) and a CENP-C

centromeric peptide (Kato et al., 2013)) bound to the nucleosome have been determined to date. All six structures share a common 'arginine anchor' motif where the chromatin protein or peptide inserts an arginine side chain into an acidic patch on the histone dimer surface of the nucleosome (Figure 1-5) formed by H2A residues Glu61, Glu64, Asp90 and possibly Glu92 (locations shown in Figure 3-2). Given this common use of the arginine anchor as a nucleosome recognition motif, I asked if LSD1/CoREST would also interact with this nucleosome acidic patch. A recombinant nucleosome containing the H2A E61A, E64A, D90A, E92A quadruple mutation, H3 fluorescently labeled on K27C, and 177 bp of 601 DNA (16N16) was prepared and tested in the HI-FI binding assay. This acidic patch quadruple mutation eliminated binding and catalytic activity of the PRC1 ubiquitylation complex with the nucleosome (McGinty et al., 2014). The results showed that LSD1/CoREST binds the 177 bp nucleosome containing H2A E61A/E64A/D90A/E92A mutations with  $98 \pm 27$  nM dissociation constant, similar to the equivalent non-mutated nucleosome with 181 bp DNA and  $\sim$ 2-fold higher affinity than to equivalent non-mutated nucleosomes with 177 bp DNA (Table 3-2). This result indicates that LSD1/CoREST does not make critical interactions to the nucleosome acidic patch, nor does it employ an arginine anchor. In fact, it appears that the acidic patch partially inhibits binding of LSD1/CoREST to the nucleosome.



**Figure 3-7.** Photocrosslinking results mapped onto the structure of the nucleosome core particle. The molecule shown contains the histone octamer from the 1.9 Å structure of the nucleosome core particle to show the extent of the histone tails (PDB ID: 1KX5), the DNA from the 2.5 Å structure of the 601 nucleosome core particle to show the 145 bp nucleosomal DNA (PDB ID: 3LZ0), and 31 bp of modeled B-form extranucleosomal DNA. Positions along the phosphate backbone crosslinked to LSD1, CoREST, and both LSD1/CoREST are shown in green, blue and red and labeled with the base pair position from the nucleosome dyad, while positions with little or no crosslinking are shown in light blue. Figure prepared by Song Tan using PyMOL molecular graphics software (Schrodinger, 2011).

## **3.4 Discussion**

My results show that extranucleosomal DNA increases both the catalytic activity of the LSD1/CoREST complex on nucleosome substrates as well as nucleosome substrate binding affinity. These findings indicate that LSD1/CoREST interacts with extranucleosomal DNA beyond the nucleosome core particle and that these interactions positively influence the histone demethylase activity of the complex.

My enzymatic and binding studies in addition to Blaine Bartholomew group's photocrosslinking studies suggest that LSD1/CoREST binds productively to the nucleosome near the nucleosome dyad, as well as to approximately 20 bp of extranucleosomal DNA extending from the end of the nucleosome core particle. I observed maximal binding to 181 bp symmetric nucleosomes, containing 18 bp extranucleosomal DNA on each end. Blaine Bartholomew and colleagues' data also detected crosslinking of LSD1 and CoREST to extranucleosomal DNA ~10 and 20 bp from the nucleosome core particle, as well as to DNA around the nucleosome dyad. Altogether these results highlight the role of extranucleosomal DNA in directing LSD1/CoREST activity. Furthermore, I identified only three histone positions where LSD1/CoREST binding significantly quenched fluorescence of the probes incorporated into nucleosomes (H2B K20, H3 A21, and H3 K27). In contrast, modest or no change in fluorescence was detected at the six other histone positions examined, many on the histone face of the nucleosome. I also found that that the histone dimer quadruple acidic patch mutation does not adversely affect LSD1/CoREST binding to the nucleosome. Therefore, it appears likely that LSD1/CoREST interacts with the nucleosome primarily through interactions with nucleosomal DNA around the nucleosome dyad and with extranucleosomal DNA.

However, I also detected the fluorescence changes using nucleosomes labeled at the H2A T10C and at H2B K20C positions at the H2A T10C and at H2B K20C positions located away

from the nucleosome dyad (Figure 3-2) and crosslinking to nucleosomal DNA at -33 and +42 bp from the nucleosome dyad (Figure 3-6). It is likely that these interactions reflect nonspecific binding of LSD1/CoREST to nucleosomal DNA because these interactions are weakened by increased salt concentrations even in the presence of extranucleosomal DNA (Figure 3-3). In the absence of extranucleosomal DNA required for the preferred binding mode of LSD1/CoREST on the nucleosome, LSD1/CoREST might bind nonspecifically to nucleosomal DNA producing a modest change in fluorescence at the H2A T10C label close to nucleosomal DNA. Increasing the salt concentration in the absence of extranucleosomal DNA reduces this nonspecific binding. The decrease is further amplified in the presence of extranucleosomal DNA as the preferred LSD1/CoREST binding mode, to DNA around the nucleosome dyad and extending from the nucleosome core particle, is favored. This interpretation is also consistent with a study which showed that LSD1/CoREST binds DNA nonspecifically possibly through the CoREST SANT domain (Pilotto *et al.*, 2015).

I propose that LSD1/CoREST binds as a 1:1 complex to the nucleosome core particle as well as to extranucleosomal DNA to position the LSD1 amine oxidase domain for catalysis with the histone H3 N-terminal tail. This would explain my nucleosome binding and crosslinking results as well as our enzymatic assays showing that extranucleosomal DNA increase catalytic activity of the LSD1/CoREST complex. Pilotta *et al.* have proposed that LSD1/CoREST may scan nucleosomal DNA via nonspecific DNA binding (Pilotto *et al.*, 2015), and I extend this concept to suggest additional interactions with extranucleosomal DNA may position LSD1/CoREST for appropriate binding of histone H3 tail and subsequent catalysis.

My nucleosome binding studies showed increased binding affinity when the nucleosome core particle is symmetrically extended with extranucleosomal DNA up to 181 bp (18N18), but extensions to 185 bp (20N20) and 207 bp (31N31) led to an unexpected decrease in binding

affinity when nucleosome labeled on H3 K27C were used. Similarly, lengthening nucleosomes on one side from 147 bp (1N1) to 151L (5N1) to 161L (15N1) increased binding affinity, but the further extension to 177L (31N1) decreased binding affinity back to the level of 151L. This decrease in binding affinity occurred despite close contact with extranucleosomal DNA observed in our photocrosslinking experiments 10–20 bp from the central 145 bp nucleosome core particle. Several explanations could account for these observations. It is possible that LSD1/CoREST bends or distorts extranucleosomal DNA beyond 181 bp of symmetrically extended nucleosome DNA and that such DNA distortions expend binding energy. It is also possible that the LSD1/CoREST protein complex undergoes conformational changes to engage undistorted extranucleosomal DNA. Another possible explanation is that a second LSD1/CoREST could bind to the extranucleosomal DNA in such a way that competes with LSD1/CoREST detected by the installed fluorescent probe. I currently cannot distinguish between these possibilities and I do not have an explanation for the essentially unchanged apparent dissociation constant for 207 versus 181 bp nucleosomes when nucleosomes labeled on H3 A21C were used (Supplementary Figure 3-1).

My biochemical analysis provides only very low resolution information for how LSD1/CoREST binds to its nucleosome substrate. Nevertheless, it seems likely that both LSD1 and CoREST subunits interact with the nucleosomal DNA near the nucleosomal dyad as well as with extranucleosomal DNA, presumably to position LSD1's catalytic domain to bind to and act on the H3 N-terminal tail. Further investigations, particularly structural studies, should help determine precisely how the various domains and regions of the LSD1 and CoREST proteins interact with elements of the nucleosome to achieve appropriate positioning for catalysis.

## **3.5 Supplementary Figures**



**Supplementary Figure 3-1.** Summary of dissociation constants for nucleosomes containing 147 bp, 157bp, 181 bp and 207 bp symmetrically positioned DNA labeled on H3A21C. The standard deviations from three or more measurements are shown as vertical bars for each data point.

## 3.6 Acknowledgements

Song Tan has designed and cloned the plasmid constructs for nucleosomal DNA of various lengths. I thank Matt Jennings for his initial work on the LSD1/CoREST project, including the demethylase assays employing the Amplex UltraRed. Michael Moore, Maxwell Kruse, and Kevin Thyne have prepared the recombinant nucleosomes used in the activity and the HI-FI binding assays. Nilanjana Chatterjee, a former member of Blaine Bartholomew laboratory, has performed the photocrosslinking assays. Kevin Thyne has prepared the native gel of the nucleosomes shown in Figure 3-4a, and Song Tan has prepared the figures 3-2 and 3-7. I thank Rob McGinty for his guidance and help in executing the HI-FI binding assays, and Marty Bollinger for helpful advice in kinetic analysis of my enzymatic assays.

## **3.7 References**

- Armache, K. J., J. D. Garlick, D. Canzio, G. J. Narlikar and R. E. Kingston. "Structural Basis of Silencing: Sir3 Bah Domain in Complex with a Nucleosome at 3.0 Å Resolution." *Science* 334, no. 6058 (2011): 977-82.
- Barbera, A. J., J. V. Chodaparambil, B. Kelley-Clarke, V. Joukov, J. C. Walter, K. Luger and K.
  M. Kaye. "The Nucleosomal Surface as a Docking Station for Kaposi's Sarcoma Herpesvirus LANA." *Science* 311, no. 5762 (2006): 856-61.
- Chen, Y., Y. Yang, F. Wang, K. Wan, K. Yamane, Y. Zhang and M. Lei. "Crystal Structure of Human Histone Lysine-Specific Demethylase 1 (LSD1)." *Proc Natl Acad Sci U S A* 103, no. 38 (2006): 13956-61.
- Culhane, J. C. and P. A. Cole. "LSD1 and the Chemistry of Histone Demethylation." *Curr Opin Chem Biol* 11, no. 5 (2007): 561-8.
- Culhane, J. C., D. Wang, P. M. Yen and P. A. Cole. "Comparative Analysis of Small Molecules and Histone Substrate Analogues as LSD1 Lysine Demethylase Inhibitors." *J Am Chem Soc* 132, no. 9 (2010): 3164-76.
- Forneris, F., C. Binda, A. Adamo, E. Battaglioli and A. Mattevi. "Structural Basis of LSD1-CoREST Selectivity in Histone H3 Recognition." *J Biol Chem* 282, no. 28 (2007): 20070-4.
- Forneris, F., C. Binda, M. A. Vanoni, E. Battaglioli and A. Mattevi. "Human Histone Demethylase LSD1 Reads the Histone Code." J Biol Chem 280, no. 50 (2005): 41360-5.
- Hieb, A. R., S. D'Arcy, M. A. Kramer, A. E. White and K. Luger. "Fluorescence Strategies for High-Throughput Quantification of Protein Interactions." *Nucleic Acids Res* 40, no. 5 (2012): e33.

- Hota, S. K. and B. Bartholomew. "Approaches for Studying Nucleosome Movement by ATP-Dependent Chromatin Remodeling Complexes." *Methods Mol Biol* 809, (2012): 367-80.
- Hou, H. and H. Yu. "Structural Insights into Histone Lysine Demethylation." Curr Opin Struct Biol 20, no. 6 (2010): 739-48.
- Kato, H., J. Jiang, B. R. Zhou, M. Rozendaal, H. Feng, R. Ghirlando, T. S. Xiao, A. F. Straight and Y. Bai. "A Conserved Mechanism for Centromeric Nucleosome Recognition by Centromere Protein CENP-C." *Science* 340, no. 6136 (2013): 1110-3.
- Kim, S. A., N. Chatterjee, M. J. Jennings, B. Bartholomew and S. Tan. "Extranucleosomal DNA Enhances the Activity of the LSD1/CoREST Histone Demethylase Complex." *Nucleic Acids Res* 43, no. 10 (2015): 4868-80.
- Lee, M. G., C. Wynder, N. Cooch and R. Shiekhattar. "An Essential Role for CoREST in Nucleosomal Histone 3 Lysine 4 Demethylation." *Nature* 437, no. 7057 (2005): 432-5.
- Makde, R. D., J. R. England, H. P. Yennawar and S. Tan. "Structure of RCC1 Chromatin Factor Bound to the Nucleosome Core Particle." *Nature* 467, no. 7315 (2010): 562-6.
- McGinty, R. K., R. C. Henrici and S. Tan. "Crystal Structure of the PRC1 Ubiquitylation Module Bound to the Nucleosome." *Nature* 514, no. 7524 (2014): 591-6.
- Morgan, M. T., M. Haj-Yahya, A. E. Ringel, P. Bandi, A. Brik and C. Wolberger. "Structural Basis for Histone H2B Deubiquitination by the SAGA DUB Module." *Science* 351, no. 6274 (2016): 725-8.

Pilotto, S., V. Speranzini, M. Tortorici, D. Durand, A. Fish, S. Valente, F. Forneris, A. Mai, T. K. Sixma, P. Vachette and A. Mattevi. "Interplay among Nucleosomal DNA, Histone Tails, and Corepressor CoREST Underlies LSD1-Mediated H3 Demethylation." *Proc Natl Acad Sci U S A* 112, no. 9 (2015): 2752-7.

Schrodinger, LLC. "The Pymol Molecular Graphics System." 1.4 (2011).

- Stavropoulos, P., G. Blobel and A. Hoelz. "Crystal Structure and Mechanism of Human Lysine-Specific Demethylase-1." *Nat Struct Mol Biol* 13, no. 7 (2006): 626-32.
- Vasudevan, D., E. Y. Chua and C. A. Davey. "Crystal Structures of Nucleosome Core Particles Containing the '601' Strong Positioning Sequence." *J Mol Biol* 403, no. 1 (2010): 1-10.
- Winkler, D. D., K. Luger and A. R. Hieb. "Quantifying Chromatin-Associated Interactions: The HI-FI System." *Methods Enzymol* 512, (2012): 243-74.
- Yang, M., J. C. Culhane, L. M. Szewczuk, C. B. Gocke, C. A. Brautigam, D. R. Tomchick, M. Machius, P. A. Cole and H. Yu. "Structural Basis of Histone Demethylation by LSD1 Revealed by Suicide Inactivation." *Nat Struct Mol Biol* 14, no. 6 (2007): 535-9.
- Yang, M., C. B. Gocke, X. Luo, D. Borek, D. R. Tomchick, M. Machius, Z. Otwinowski and H. Yu. "Structural Basis for CoREST-Dependent Demethylation of Nucleosomes by the Human LSD1 Histone Demethylase." *Mol Cell* 23, no. 3 (2006): 377-87.

## Chapter 4

## Structural Characterization of LSD1/CoREST/Nucleosome Complex by Xray Crystallography

## 4.1 Abstract

Lysine-specific demethylase 1 (LSD1) is a FAD-dependent amine oxidase that regulates transcription by removing methyl groups from mono- and dimethylated Lys of histone H3. Its corepressor protein CoREST enhances and fine-tunes LSD1's activity and specificity on the nucleosome substrates to repress transcription. I have shown that extranucleosomal DNA enhances the affinity and catalytic activity of LSD1/CoREST complex towards nucleosomes, but we still lack structural information regarding LSD1/CoREST's interaction on nucleosomes with extranucleosomal DNA.

I have crystallized LSD1/CoREST in complex with the nucleosome and I have used these crystals to determine the molecular replacement solutions of LSD1/CoREST on 161 (8N8) and 189 bp (22N22) nucleosome crystal structures at 7.2 Å and 5.3 Å resolutions respectively. The structure of LSD1/CoREST on the 161 bp nucleosome suggests the possibility that LSD1/CoREST uses a dinucleosome substrate, whereas the structure of LSD1/CoREST on 189 bp nucleosome suggests that the complex uses a mononucleosome with extranucleosomal DNA as the substrate. The LSD1 AOD in the dinucleosome substrate model engages the nucleosomal DNA whereas LSD1 AOD in the mononucleosome substrate model makes contacts with extranucleosomal DNA. However, both models show that the CoREST SANT2 binds on the nucleosome disk face and that N-terminus of CoREST∆1 makes contacts with extranucleosomal

DNA. Even though a higher resolution structure is required to visualize specific residues involved in the interaction, my structures already provide novel insights into how LSD1/CoREST engages the nucleosome with DNA extensions.

## 4.2 Introduction

Currently, there is a wealth of information regarding how the LSD1/CoREST complex binds the H3-peptide. The crystal structures of LSD1/CoREST/H3-peptide ternary complex revealed the various contacts made between the LSD1's active site and its peptide substrate (Forneris *et al.*, 2007, Yang *et al.*, 2007). In addition, previous structural and biochemical studies have shown that the H3 tail has to be stripped of other posttranslational modifications in the first 16 amino acids for its optimal interaction with LSD1 (Forneris *et al.*, 2005, Forneris *et al.*, 2007). However, only hypothesized models exist for how LSD1/CoREST complex engages the nucleosomes (Figure 4-1) (Yang *et al.*, 2006, Pilotto *et al.*, 2015). Yet, none of these models has taken the role extranucleosomal DNA into the account. Their models propose that the CoREST SANT2 domain interacts with the nucleosomal DNA, and the catalytic LSD1 amine oxidase domain is positioned close to DNA gyres where H3 tail exits the nucleosome (Figure 4-1). However, since extranucleosomal DNA dramatically increases the catalytic activity and binding affinity of LSD1/CoREST on nucleosomes (Kim *et al.*, 2015), further investigation, particularly structural studies with nucleosomes containing extranucleosomal DNA, will be necessary to fully elucidate the mechanism of LSD1/CoREST on nucleosomes.

Before the binding studies of LSD1/CoREST complex on nucleosomes with extranucleosomal DNA were performed, attempts to crystallize LSD1/CoREST on 147 bp nucleosome were made by other members of the laboratory without success. My activity and binding studies of LSD1/CoREST on nucleosomes with various DNA lengths established the importance of extranucleosomal DNA for the complex's interaction and function on the nucleosome. Therefore, I pursued crystallization of LSD1/CoREST on nucleosomes with various lengths of extranucleosomal DNA, which are better substrates for LSD1/CoREST. In addition, I explored different truncations and orthologs of LSD1/CoREST for crystallization, in hope of enhancing crystallization and/or diffraction properties.

After extensive crystallization screens of 33 different LSD1/CoREST/nucleosome complexes, I was able to grow crystals of LSD1/CoREST on 159 (7N7), 161 (8N8), 163 (9N9), 189 (22N22), and 193 bp (24N24) nucleosomes either containing WT histones or H3 with K4M mutation. Four of them diffracted between 5.3-8.9 Å resolutions, which allowed me to determine the molecular replacement solutions to the crystal structures. Although side chains and loop regions in my structures cannot be placed with confidence, my models provide insights into how each domain of LSD1/CoREST interacts with different regions of the nucleosome.



**Figure 4-1**. Currently existing hypothesized models of LSD1/CoREST on the nucleosome. (a) Yang's theoretical model of LSD1/CoREST on nucleosome core particle. (Figure 7 from Yang, M., Gocke, C.B., Luo, X., Borek, D., Tomchick, D.R., Machius, M., Otwinowski, Z. and Yu, H. (2006) Structural basis for CoREST-dependent demethylation of nucleosomes by the human LSD1 histone demethylase. *Molecular cell*, **23**, 377-387.) (b) Pilotto's model of LSD1/CoREST on nucleosome containing 146 bp of Widom 601 DNA based on SAXS data. (Figure 5c from Pilotto, S., Speranzini, V., Tortorici, M., Durand, D., Fish, A., Valente, S., Forneris, F., Mai, A., Sixma, T.K., Vachette, P. *et al.* (2015) Interplay among nucleosomal DNA, histone tails, and corepressor CoREST underlies LSD1-mediated H3 demethylation. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 2752-2757.)

## 4.3 Results

## 4.3.1 Reconstitution of LSD1/CoREST/nucleosome complex

#### 4.3.1.1 LSD1/CoREST complex constructs and orthologs

The human LSD1/CoREST construct used in my activity and binding studies in the previous chapter is the same construct that successfully crystallized to 2.6 Å resolution by Yang et al (Yang et al., 2006). Even though this construct contains LSD1(171-852) and CoREST(286-482), the electron density permitted structure determination of LSD1(171-836) and CoREST(308-440). The C-terminal 16 residues of LSD1, and N-terminal 22 residues and C-terminal 42 residues of CoREST were missing in this 2006 crystal structure. These missing regions conceivably represent flexible loops that yielded poor electron density. As the presence of flexible regions can inhibit macromolecular crystallization (Rhodes, 2006), it is worth exploring constructs removing these flexible regions for the crystallographic purpose. In particular, I focused on the CoREST subunit and tested whether deletions in the N-terminus or C-terminus of CoREST $\Delta$ 1 (hereafter referred to CoREST $\Delta$ 2 = CoREST(308-482) and CoREST $\Delta$ 3 = CoREST(286-440) respectively) affect binding of LSD1/CoREST on the nucleosomes. The HI-FI nucleosome binding studies at 75 mM NaCl (the salt concentration at which reconstitution of LSD1/CoREST/nucleosome complex was performed for crystallization) showed that LSD1/CoREST $\Delta$ 2 binds to 157 (6N6) and 181 bp (18N18) nucleosomes with 2-fold lower affinity compared to LSD1/CoREST (Figure 4-2, Table 4-1), suggesting that the CoREST residues 286-307 play a role in binding to nucleosome. On the other hand, the binding affinity of LSD1/CoREST $\Delta$ 3 on the 157 and 181 bp nucleosomes did not change significantly (Figure 4-2,

Table 4-1). Therefore, LSD1/CoREST $\Delta$ 3 was used in crystallization trials, in addition to the LSD1/CoREST construct.



**Figure 4-2.** Effect of deleting residues in the N- or C-terminus of CoREST on the binding affinity for nucleosomes. (a) Binding curves for LSD1/CoREST (orange), LSD1/CoREST $\Delta 2$  (pink), and LSD1/CoREST $\Delta 3$  (blue) interacting with 157 bp nucleosomes fluorescently labeled on H3 K27C at 75 mM NaCl. The dashed vertical lines correspond to the calculated dissociation constants for each respective binding curve. (b) Equivalent binding curves for LSD1/CoREST $\Delta 2$  (red), LSD1/CoREST $\Delta 2$  (purple), and LSD1/CoREST $\Delta 3$  (green) interacting with 181 bp nucleosomes.

LSD1/CoREST	CoREST	DNA length	K <sub>d</sub> (nM)	relative to
		(bp)		LSD1/CoREST
LSD1/CoREST	286-482	157	862 ± 27	1.0
LSD1/CoREST∆2	308-482	157	1681 ± 334	2.2
LSD1/CoREST∆3	286-440	157	712	0.95
LSD1/CoREST	286-482	181	392 ± 31	1.0
LSD1/CoRESTA2	308-482	181	844 ± 179	1.9
LSD1/CoRESTA3	286-440	181	373	0.83

 Table 4-1. Dissociation constants for different LSD1/CoREST complex constructs binding to 157

 bp (6N6) and 181 bp (18N18) nucleosomes

Orthologs can also be used to enhance crystallization of proteins when an organism of

choice fails to crystallize or the crystals are not suitable for structure determination

(Andrykovitch et al., 2003). The amino acid sequences should be similar enough that the

orthologs share the same fold, but different enough so that they may pack differently in crystals

and potentially diffract to higher resolution. In the case of the RCC1/nucleosome complex, crystals with human RCC1 produced poor diffraction and crystals with yeast RCC1 (Srm1) produced nonisotropic diffraction to ~ 7 Å. On the other hand, crystals grown with *Drosophila* RCC1 (Bj1) yielded diffraction data to 2.9 Å resolution (Makde *et al.*, 2010, Makde and Tan, 2013). Our lab has constructed expression plasmids for human, zebrafish, *C. elegans, Drosophila*, and *Xenopus* LSD1/CoREST complexes. However, only the human and the zebrafish orthologs could be expressed and purified in large quantities. I have put most of my efforts into the human LSD1/CoREST complex, but also worked with the zebrafish ortholog with similar extents as the human complex [zLSD1(176-867) and zCoREST(230-458)] in my crystallization studies.

## 4.3.1.2 Nucleosomes with various DNA lengths

My nucleosome binding studies of LSD1/CoREST demonstrated that extranucleosomal DNA enhances the interaction of LSD1/CoREST on the nucleosomes and the optimal binding affinity was observed with the 181 bp (18N18) nucleosomes (Figure 3-4d). I initially decided to set up crystallization trials of the LSD1/CoREST/nucleosome complex with Widom 601 nucleosome positioning sequence that varies in length from 155 bp (5N5) to 177 bp (16N16), which produced high enough binding affinity for stable complex formation. My rationale was that while the 147 bp nucleosome is not the ideal substrate for LSD1/CoREST complex, nucleosomes with longer DNA length would have an extra DNA bp on either or both sides of the nucleosome that is not bound by the protein and can hinder crystallization. I initially began crystallization trials with 155 (5N5), 157 (6N6), 159 (7N7), 161 (8N8), 163 (9N9), 165 (10N10) and 177 bp (16N16) nucleosomes. Later crystallization studies were performed using 185 (20N20), 189 (22N22), and 193 bp (24N24) mononucleosomes, and 305 (N16N), 306 (N16N), 321 (8N16N7), and 322 bp (8N16N8) dinucleosomes. In particular, nucleosomes with DNA lengths varied by 2

bp, or 1 bp on either side of the nucleosome were used because previous crystal structures of the nucleosome bound with protein factors have demonstrated that DNA ends are important for crystal packing. Additionally, DNA ends can also mediate crystal contacts with the chromatin factors and even one base pair of DNA can determine whether the end of the nucleosomal DNA makes a crystal contact or not (Makde and Tan, 2013).

## 4.3.1.3 Nucleosome with H3 K4M mutation increases binding affinity for LSD1/CoREST

Forneris et al. have shown that a 21-amino acid H3-peptide bearing K4M mutation led to a ~30-fold increase in binding affinity for LSD1/CoREST compared to the wild-type peptide (Forneris et al., 2007). To examine if this phenomenon holds true in the context of the nucleosome substrates, a recombinant nucleosome containing the H3 K4M mutation and 181 bp (18N18) of Widom 601 DNA was prepared and tested in the HI-FI binding assay at 75 mM NaCl. Note that this particular HI-FI binding study was performed with the unlabeled nucleosome and the LSD1/CoREST complex that was site-specifically labeled at CoREST V415C. The CoREST residue Val 415 was chosen as the labeling position because its location in the SANT2 domain suggests the likelihood that it is in the close proximity to interaction site. The binding affinities of the labeled LSD1/CoREST towards the nucleosomes of various DNA lengths follow the same trend as when the HI-FI probe is on the nucleosome (Supplementary Figure 4-1). The results showed that LSD1/CoREST binds the 181 bp nucleosome containing H3 K4M mutation with <1.9 nM dissociation constant, ~40-fold higher affinity than to equivalent non-mutated nucleosomes with 181 bp DNA ( $78.1 \pm 0.8$  nM) (Figure 4-3). Previously, the peptide with H3 K4M mutation was successfully used for co-crystallization with LSD1/CoREST and allowed for structure determination of the peptide in the substrate binding site of LSD1/CoREST (Forneris et

*al.*, 2007). Therefore, recombinant nucleosomes containing the H3 K4M mutation and desired lengths of Widom 601 DNA were prepared for crystallization studies.



**Figure 4-3.** Effect of mutating Lys4 to Met4 in histone H3 in the nucleosome on the binding affinity to LSD1/CoREST. Binding curves for LSD1/CoREST interacting with wild-type (yellow) and H3 K4M (blue) nucleosomes at 75 mM NaCl. The dashed vertical lines correspond to the calculated dissociation constants for each respective binding curve.





## 4.3.1.4 Optimization of salt concentration in reconstitution buffer

In Chapter 3, I have suggested that LSD1/CoREST might bind nonspecifically to nucleosomal DNA at 50 mM salt concentration. This nonspecific binding of LSD1/CoREST to DNA was detected by fluorescence change using nucleosomes labeled at H2A T10C and H2B K20C in my studies (Figure 3-2) and also by DNA binding studies by Pilotto *et al* (Pilotto *et al.*, 2015). It is imperative that these interactions are reduced for crystallization as monodispersity and homogeneity of the sample is key to successful crystallization (Rhodes, 2006). I have shown that the nonspecific binding of LSD1/CoREST to the nucleosomes is weakened by an increased salt concentrations (Figure 3-3). Therefore, salt concentrations of 75 and 100 mM were tested for analytical reconstitution and purification of LSD1/CoREST/nucleosome complex by size exclusion chromatography. The results showed that LSD1/CoREST dissociates from the 159 bp (7N7) nucleosome in buffer containing 100 mM NaCl (Figure 4-4a, c). On the other hand, LSD1/CoREST/nucleosome successfully reconstituted and eluted from the size exclusion column as a sharp, symmetric peak in the reconstitution buffer containing 75 mM NaCl (Figure 4-4b, d). I therefore used the buffer containing 75 mM NaCl for reconstitution and crystallization studies.

## 4.3.1.5 LSD1/CoREST is successfully reconstituted on nucleosomes with various DNA lengths

The human LSD1/CoREST complex was reconstituted with 21 different nucleosomes (Table 4-2), the human LSD1/CoREST∆3 was reconstituted with 4 different nucleosomes (Table 4-3), and the zebrafish LSD1/CoREST was reconstituted with 8 different nucleosomes (Table 4-4), totaling 33 different LSD1/CoREST/nucleosome complexes of various constructs and orthologs. Reconstitution was performed in buffer containing 10 mM Tris-Cl pH 7.5, 75 mM NaCl, and 1 mM DTT at room temperature and the reconstituted complexes were purified over
DNA length	Histones	DNA format	Crystal	Diffraction
155	WT	5+145 +5	X	
157	WT	6+ 145 +6	Х	
159	WT	7+ 145 +7	7+ 145 +7 0	
161	WT	8+ 145 +8 o		7.2 Å
163	WT	9+ 145 + 9	0	8.9 Å
165	WT	10+ 145 + 10	х	
177	WT	16+ 145 + 16	Х	
185	WT	20+ 145 + 20	Х	
189	WT	22+ 145 + 22	0	Х
193	WT	24+ 145 + 24	0	х
305	WT	145 + 15 +145	х	
306	WT	145 + 16 +145	х	
321	WT	8 + 145 + 16 +145 + 7	х	
322	WT	8 + 145 + 16 +145 + 8	0	Х
			х	
159	H3 K4M	7+ 145 +7	х	
161	H3 K4M	8+145 +8	0	х
163	H3 K4M	9+145 +9	0	х
185	H3 K4M	20+ 145 + 20	Х	
189	H3 K4M	22+ 145 + 22	0	5.3 Å
193	H3 K4M	24+ 145 + 24	0	X
322	H3 K4M	8 + 145 + 16 +145 + 8	0	X

**Table 4-2.** Nucleosome substrates used in crystallization with human LSD1/CoREST and their crystallization and diffraction results. Human LSD1/CoREST was screened with 21 different nucleosomes.

DNA length	Histones	DNA format	Crystal	Diffraction
161	H3 K4M	8+ 145 +8 x		
163	H3 K4M	9+ 145 +9 x		
189	H3 K4M	22+ 145 + 22	0	5.8 Å
8N16N8	H3 K4M	8 + 145 + 16 +145 + 8	Х	

**Table 4-3**. Nucleosome substrates used in crystallization with human LSD1/CoREST $\Delta$ 3 and their crystallization and diffraction results. Human LSD1/CoREST $\Delta$ 3 was screened with 4 different nucleosomes.

DNA length	Histones	DNA format	Crystal	Diffraction
157	WT	6+145+6	х	
159	WT	7+ 145 +7	Х	
161	WT	8+145 +8	Х	
163	WT	9+145 +9	Х	
8N16N8	WT	8 + 145 + 16 +145 + 8	Х	
159	H3 K4M	7+ 145 +7	Х	
161	H3 K4M	8+145 +8	X	
163	H3 K4M	9+145 +9	Х	

**Table 4-4**. Nucleosome substrates used in crystallization with zebrafish LSD1/CoREST and their crystallization and diffraction results. Zebrafish LSD1/CoREST was screened with 8 different nucleosomes.

size exclusion column. In most cases, the LSD1/CoREST/nucleosome complexes eluted as a sharp, symmetric peak (Figure 4-4b). A shift in chromatographic peak to a decreased elution time was observed for LSD1/CoREST on nucleosome with longer DNA lengths (185–193 bp) and dinucleosomes. Slightly higher (~ 10 % increase) purification yield was realized when nucleosomes containing H3 K4M mutation was used. Nonetheless, the shape of the chromatographic peak was similar in all LSD1/CoREST/nucleosome complexes. The purified complexes were concentrated and used immediately to set up crystal trials.

# 4.3.2 Crystallization and structural determination of LSD1/CoREST/nucleosome complex

For each LSD1/CoREST/nucleosome complex, three or more crystallization sparse matrix screens, developed by our lab, were set up in modified microbatch with 8-10 mg/ml of the complex and placed at 4°C refrigerator and/or 21°C room. Each screen contained 96 different solutions, with various monovalent ions either at 75, 100, or 150 mM concentration, buffer with a range of pH (5.0 - 8.5), and precipitant with concentration ranging between 2–15% (typically PEG2000-MME was used). A fraction of solutions also contained divalent ions at millimolar concentrations. Due to stochastic nature of crystal growth, I have set up at least three sparse matrix screens, each with 96 conditions (with some overlap between screening conditions), for each LSD1/CoREST/nucleosome complex to maximize the chances of obtaining crystals.

It was crucial to check for proteolysis of LSD1/CoREST/nucleosome samples in crystal trays over time as the degradation of the proteins is detrimental to crystallization. Therefore, I have saved the purified and concentrated LSD1/CoREST/nucleosome samples and incubated them at the same temperature as the crystal screen trays. Proteolysis was observed after a week for some samples stored at 21°C, indicating that LSD1/CoREST/nucleosome complexes in crystal

screen drops were likely proteolyzed as well. The samples stored in 4°C refrigerator appeared intact even after a month and therefore, crystallization was pursued at 4°C in most cases.



**Figure 4-5**. Initial and optimized crystals of LSD1/CoREST complexed with different nucleosomes. (a), (b), (c) Initial crystals of LSD1/CoREST on 161 (8N8) and 189 bp (22N22) mononucleosomes, and 322 bp (8N16N8) dinucleosome from screens respectively. (d), (e), (f) Optimized crystals of LSD1/CoREST/ nucleosomes from (a), (b), and (c) respectively.

Crystals typically started to grow 3 to 14 days after they were set up. In the case of the human LSD1/CoREST/nucleosome complexes, 11 out of 21 different complexes formed crystals of different morphologies in several conditions (Table 4-2). From the selected drops, the crystals were washed and run on the SDS-PAGE gel to confirm that the crystals contained both the nucleosomes, LSD1, and CoREST proteins. For the human LSD1/CoREST $\Delta$ 3/ nucleosome complexes, 1 of 4 complexes grew crystals (Table 4-3). Crystals grew only when human LSD1/CoREST was used, and nucleosomes complexed with the zebrafish LSD1/CoREST did not

produce any crystals (Table 4-4). However, it should be noted that only a few different nucleosomes were scanned with zebrafish LSD1/CoREST, and more thorough crystallization studies with zebrafish ortholog should be done before this ortholog is abandoned completely. For each complex that produced crystals, conditions that produced single crystals with three dimensions, birefringence, and sharp edges were optimized by varying the pH, salt concentration, and precipitant amount around the initial condition. Optimization usually yielded larger crystals with fewer flaws (Figure 4-5).

Post-crystallization treatments were carried out to improve the diffraction quality of the protein crystals. In particular, dehydration, which can induce reorganization of the crystal lattice and lower solvent content, and cryoprotection, which can prevent crystallization of water to achieve the vitrified state during the cooling process, were employed (Rhodes, 2006). The optimized crystals were soaked in various cryoprotectants and dehydrating agents including methylpentanediaol (MPD), glycerol, PEG400, PEG550-MME, PEG600, and PEG750-MME either at room temperature or in the 4°C cold room, and the crystals were flashed cooled by plunging into liquid nitrogen. The diffraction data of the crystals were collected at Advanced Photon Source's NE-CAT beamline 24-ID-E or 24-ID-C. A majority of the crystals diffracted poorly and only a few of them diffracted to  $\sim 8$  Å or higher resolution. Out of the crystals that diffracted, two are noteworthy: LSD1/CoREST with 161 bp (8N8) nucleosome crystal that diffracted to 7.2 Å resolution (Figure 4-6a, b), and LSD1/CoREST with 189 bp (22N22) nucleosome containing the H3 K4M mutation that diffracted to 5.3 Å resolution (Figure 4-6c, d). These two diffraction data were processed and molecular replacement solution structures were obtained without further refinement (Supplementary Table 4-1). The fact that DNA extensions beyond the nucleosome core particle, which are not part of the input models in molecular replacement, are shown in electron density map suggests that the solutions are probably substantially correct (Supplementary Figure 4-2). LSD1/CoREST with the 161 bp nucleosomes

yielded a dinucleosome substrate model, and LSD1/CoREST with the 189 bp nucleosomes yielded a mononucleosome substrate model, both of which will be further described in the following sections.



**Figure 4-6**. Crystals of LSD1/CoREST/nucleosome and diffraction images. (a) Crystals of LSD1/CoREST complexed with 161 bp (8N8) nucleosome used to solve the structure of the dinucleosome substrate model. (b) Crystal from (a) diffracted to 7.2 Å resolution. (c) Crystals of LSD1/CoREST complexed with 189 bp (22N22) H3 K4M nucleosome used to solve the structure of the mononucleosome substrate model. (d) Crystal from (c) diffracted to 5.3 Å resolution. (e) Crystals of LSD1/CoREST $\Delta$ 3 complexed with 189 bp (22N22) H3 K4M nucleosome. (f) Crystal from (e) diffracted to 5.8 Å resolution.

#### 4.3.3 The Dinucleosomes substrate model of LSD1/CoREST/nucleosome

A dinucleosome substrate model was obtained from the crystals of LSD1/CoREST complexed with 161 bp (8N8) nucleosome. Even though the mononucleosome was used to grow the crystals, the nucleosomes in the molecular replacement solution structure resembled the structure of a dinucleosome, hence I refer to this as the dinucleosome substrate model. In addition, a very similar structure was obtained from molecular replacement solution of LSD1/CoREST/nucleosome with 163 bp (9N9) DNA crystals (data not shown). I would like to emphasize that the relative positions of loops and side chains shown in figures should be viewed with skepticism as they are modeled based on electron density at 7.2 Å resolution.

## 4.3.3.1 Overview of the model

In this structure, two nucleosomes with 161 bp (8N8) DNA emulates a dinucloesome structure with 16 bp linkers (Figure 4-7) as the DNA ends pack against each other head-to-head and mediate crystal contacts. Nucleosomes are oriented orthogonal to one another, as observed in the 9 Å crystal structure of tetranucleosome (Schalch *et al.*, 2005) and in the cryo-EM structure of ISW1a chromatin remodeling factor on the dinucleosome (Yamada *et al.*, 2011).



**Figure 4-7**. Dinucleosome model based on the molecular replacement solution of 161 bp (8N8) nucleosomes in the crystal structure of LSD1/CoREST/nucleosome complex. Sixteen bp B-form DNA (gray) was modeled onto the electron density formed by 8 bp DNA extension on both sides of two 161 bp nucleosomes, depicted in cartoon representation. This and the subsequent figures were prepared using PyMOL molecular graphics software (Schrodinger, 2011).





The asymmetric unit contains one copy of a 'half nucleosome' and one copy of the LSD1/CoREST complex, resulting in a stoichiometry of 4:1 LSD1/CoREST:dinucleosome complex. Each LSD1/CoREST molecule makes interaction with two nucleosome core particles in the dinucleosome simultaneously, with the AOD of LSD1 contacting the nucleosomal DNA of the nucleosome core particle (called nucleosome 1 because it is closer to the substrate binding site of the LSD1/CoREST complex for catalysis than the other nucleosome core particle in the dinucleosome) and the SANT2 domain of CoREST contacting the nucleosome disk surface of the nucleosome 2 (Figure 4-8). Therefore a dinucleosome appears to be the substrate for LSD1/CoREST complex in this model. If we imagine the LSD1/CoREST complex binding to the dinucleosome with much longer linker lengths than 16 bp, there are two possible orientations of the LSD1/CoREST molecule on the nucleosome. The LSD1/CoREST complex may bind the nucleosome core particle through the interaction between the LSD1 AOD and the nucleosomal DNA with the CoREST SANT2 domain making contacts with the linker DNA, similar to the orientation of the LSD1/CoREST molecule 1 on the nucleosome 1 in Figure 4-8. Alternatively, the LSD1/CoREST complex may bind to the nucleosome core particle through the interaction between the CoREST SANT2 domain and the nucleosome disk face with the LSD1 AOD making contacts with the linker DNA, similar to the orientation of the LSD1/CoREST molecule 1 on the nucleosome 2 in Figure 4-8.

In addition, LSD1/CoREST forms a dimer to bind to each side of the dinucleosome, with a total of four LSD1/CoREST molecules binding to a dinucleosome simultaneously in the crystal. The structure suggests five residues in each LSD1/CoREST molecule to be responsible for dimerization (Figure 4-9). The residue Lys432 in LSD1, and residues Glu341 and Gln348 in CoREST of one molecule contact the SWIRM domain of the other LSD1/CoREST molecule, while the LSD1 Asp399 and Lys550 contact the same two residues of the other molecule (Figure 4-9a). Although dimerization of LSD1/CoREST was observed in the dinucleosome substrate

model, no evidence of a LSD1/CoREST dimer in solution has been reported. Therefore, it is possible that LSD1/CoREST dimerization is an artifact of crystallization, and might not occur in solution.



**Figure 4-9.** Two LSD1/CoREST molecules homodimerize using three contact points. (a), (b) Orthogonal views of the LSD1/CoREST dimer. The complex is depicted in cartoon representation with LSD1 and CoREST of one complex colored in yellow and red, and LSD1 and CoREST of another colored in blue and green respectively. Five residues in each complex suggested to be responsible for dimerization are depicted in sphere representation in a lighter shade of respective protein color.

# 4.3.3.2 LSD1/CoREST – DNA interaction

The AOD of LSD1 makes contact with the end of nucleosomal DNA of the nucleosome 1 in the dinucleosome substrate model (Figure 4-10a). This orientation positions the catalytic site of LSD1/CoREST in close proximity to where H3 tail exits the DNA gyres. Six basic residues in LSD1, Lys355, Lys357, Lys359, Thr561, Ser563, and Arg568, are in close proximity to the nucleosomal DNA (Figure 4-10b). In addition, the negatively charged acidic patch near the substrate binding site may thread the positively charged H3 tail into the active site.



**Figure 4-10.** The AOD of LSD1 binds DNA near the end of nucleosomal DNA of the nucleosome 1. (a) LSD1 AOD – nucleosomal DNA interaction. (b) Zoomed view of showing six residues of LSD1 AOD suspected to interact with DNA, depicted in sphere representation. LSD1/CoREST and histones are depicted in cartoon representation and DNA is depicted in cartoon representation. For visualization purposes, the three other LSD1/CoREST complexes are not shown.



**Figure 4-11**. N-terminus of CoREST CoRESTΔ1 binds to linker DNA. (a) CoREST N-terminal extension linker DNA interaction. (b) Zoomed view of showing three residues of CoREST suspected to interact with linker DNA, depicted in stick representation and colored in pink. LSD1/CoREST and histones are depicted in cartoon representation and DNA is depicted in *stick* representation. For visualization purposes, the three other LSD1/CoREST complexes are not shown.

My dinucleosome substrate model also suggests that the N-terminus of CoREST $\Delta 1$  is responsible for interacting with the linker DNA of dinucleosome (Figure 4-11). This region is particularly rich in positively charged residues (Lys294, Lys295, Lys297, Lys303, Arg305, Lys307, Arg308, Lys309, and Lys312). Although the residues 286-307 in the N-terminus of CoREST $\Delta 1$  are missing in the structure, my HI-FI nucleosome binding experiments suggest the possibility that this region may be responsible for interaction with the linker DNA (Figure 4-2, Table 4-1).

Interestingly, these two regions of LSD1/CoREST complex that interact with DNA, the LSD1 AOD, and the CoREST $\Delta$ 1 N-terminus, form three 'basic stripes' observed in dimeric LSD1/CoREST (Figure 4-12a, b). Three 'basic stripes' are aligned with the end of the nucleosomal DNA of two NCPs and linker DNA by 10 bp apart (Figure 4-12c). Due to the specific distance between these basic stripes, it seems unlikely that LSD1/CoREST dimer would bind as well to dinucleosomes with significantly shorter or longer linker DNA than 16 bp.

The SWIRM domains of ADA2a and Swi3 are known to bind to DNA (Da *et al.*, 2006, Qian *et al.*, 2005), but DNA binding interaction by the LSD1 SWIRM domain was not observed in our model. This is consistent with the observation made by Yang *et al.* that residues implicated in DNA binding in the SWIRM domains of Ada2a and SWi3 are poorly conserved in LSD1 SWIRM (Yang *et al.*, 2006). In addition, AOD region partially blocks the putative DNA binding surface of LSD1 SWIRM, making it unlikely for LSD1 SWIRM to bind to DNA in the same manner as SWIRM domain of Ada2a and Swi3 (Yang *et al.*, 2006). Instead, the LSD1 SWIRM domain seems to be involved in protein-protein interaction for homodimer formation LSD1/CoREST complexes in my dinucleosome substrate model.



**Figure 4-12.** The 'basic stripes' of LSD1/CoREST homodimer are aligned with end of the nucleosomal DNA and linker DNA. (a) LSD1/CoREST homodimer depicted in cartoon representation. (b) Electric potential surface representation of LSD1/CoREST homodimer. Negatively charged, positively charged, and nonpolar surfaces are is represented in red, blue, and white respectively. The regions of LSD1/CoREST that make up three 'basic stripes' are indicated. (c) Dinucleosome with 16 bp linker between the two nucleosome core particles is overlaid on the electric potential surface representation of LSD1/CoREST homodimer. Histones and DNA are depicted in cartoon representation. For visualization, two other LSD1/CoREST complexes and the linker DNA on the other side of the dinucleosome were hidden from the structure. Surface electrostatic potential of the LSD1/CoREST dimer is calculated with ABPS (Baker *et al.*, 2001)

### 4.3.3.3 CoREST – nucleosome disk face interaction

The SANT2 domain of CoREST is in close contact with histone disk face of the nucleosome 2 (Figure 4-13). In particular, the SANT2 domain is situated above the histone H4 globular domains and also nucleosomal DNA approximately 10 bp away from the nucleosome dyad. The CoREST residues Arg425, Arg426, and Arg427 apper to interact with histone H4 whereas the residues CoREST Lys378 and Arg382 could be interacting with nucleosomal DNA. The CoREST SANT2 is critical for facilitating LSD1-mediated demethylation of nucleosomal substrates (Shi *et al.*, 2005), and it is conceivable as that LSD1/CoREST complex would be unable to recognize nucleosome core particle without the CoREST SANT2 domain as it is the only domain that interacts with a feature unique to the nucleosome core particle.



**Figure 4-13**: CoREST SANT2 is in close proximity to histone H4 surface and nucleosomal DNA. (a) CoREST SANT2 interaction – nucleosome disk face interaction. (b) Zoomed view of (a) showing three residues of CoREST suspected to interact with histone H4 and two residues in close proximity to nucleosomal DNA, depicted in stick representation and colored in pink. LSD1/CoREST and nucleosomes are depicted in cartoon representation. For visualization, three other LSD1/CoREST molecules were hidden from the structure.

#### 4.3.3.4 Distance between LSD1 catalytic site and H3 substrate

In my dinucleosome substrate model, the catalytic AOD of LSD1 is situated near where H3 tail exits the nucleosome through the DNA gyres (Figure 4-14a). The crystal structure of LSD1/CoREST complexed with H3-peptide shows density for H3 from Ala1 to Pro16 (PDB ID: 2V1D) (Forneris *et al.*, 2007) and the crystal structure of nucleosome with Widom 601 DNA sequence shows the histone H3 tail exits the DNA gyres (PDB ID: 3LZ0) (Vasudevan *et al.*, 2010). Aligning these structures in my molecular replacement solution, and calculating the distance between the C-alpha positions of H3 Pro16 in LSD1/CoREST from H3 Thr45 in the nucleosome 1 gives ~35 angstroms (Figure 4-14a). Even though there are three other H3 tail substrates in the dinucleosomes, it is unlikely that LSD1/CoREST will select these other H3 tails since they are further away with this binding mode. The distances between the H3 Pro16 in LSD1/CoREST and the Thr45 of the other histone H3 in the nucleosome 1, the residues H3 Thr45 in the histones H3 of the nucleosome 2 are ~72, ~63, and ~88 Å respectively (Figure 4-14b). The biochemical studies to evaluate the effect of varying the distance between the catalytic LSD1 site and the H3 substrate are discussed in Chapter 5.



**Figure 4-14**: Distance between the LSD1 substrate binding site and its targets in dinucleosome substrate model. (a) The distance measured between Pro16 of H3 peptide bound in LSD1 and Thr45 of the closest histone H3 in nucleosome 1 is about 35Å (green; calculated using PyMOL). (b) The distances measured between Pro16 of H3 peptide bound in LSD1 and Thr45 of the other three histone H3 in the nucleosomes (magenta). Histones, DNA, and LSD/CoREST are depicted in cartoon representation. For visualization, three other LSD1/CoREST complexes were hidden from the structure.

#### 4.3.4 The mononucleosome substrate model of LSD1/CoREST/nucleosome

After successive experiments performed to validate the dinucleosome substrate model that will be discussed in the following chapter, I concluded that the dinucleosome is not a better substrate than a mononucleosome with extranucleosomal DNA. My dinucleosome substrate model suggests two possible orientations of the LSD1/CoREST on the nucleosome core particle. The LSD1/CoREST molecule 1 may engage the nucleosome 1 by the interaction between the LSD1 AOD and the nucleosomal DNA, or the same molecule can bind to the nucleosome 2 by the interaction between the CoREST SANT2 and the nucleosome disk face (Figure 4-17 a-c). The orientation of SANT2 domain on the nucleosome 2 suggests that the LSD1/CoREST actually targets the H3 in the nucleosome 2 in the dinucleosome substrate model. With this assumption, the LSD1 AOD domain may bind to DNA extension longer than 16 bp that is absent in my my dinucleosome substrate model (Figure 4-17c). Instead, the LSD1 AOD bound the end of the nucleosomal DNA of the adjacent nucleosome in the crystal packing. To test this hypothesis that LSD1/CoREST may require ~ 20–24 bp on both sides of the nucleosomes for optimal binding, I attempted to crystallize LSD1/CoREST/nucleosome using the longer DNA (187 to 193 bp).

I successfully crystallized LSD1/CoREST on 189 bp (22N22) nucleosomes, and the resulting molecular replacement suggests a mononucleosome substrate model. The LSD1/CoREST complex on the nucleosome in model resembles the LSD1/CoREST molecule 1 oriented on the nucleosome 2 in the dinucleosome substrate model (Figure 4-17c, d). In addition, a very similar structure was obtained from molecular replacement solution of LSD1/CoREST/ nucleosome using human LSD1/CoREST $\Delta$ 3 and 189 bp (22N22) nucleosome crystals (Figure 4-6e, f). I would like to emphasize that the relative positions of the side chains shown in figures should be viewed with skepticism as they are modeled based on the electron density at 5.3 Å resolution.

## 4.3.4.1 Overview of the model

The structure shows one LSD1/CoREST molecule bound per one nucleosome, resulting in 1:1 stoichiometry of LSD1/CoREST:nucleosome complex, consistent with our HI-FI stoichiometry assay (Figure 3-5b). The CoREST SANT2 recognizes the disk face of the nucleosome near histone H3 and H4, and nucleosomal DNA around ~10 bp from the nucleosome dyad. The LSD1 AOD domain is in close proximity to interact with ~13-22 bp DNA extension, in contrast to the observed interaction with the nucleosomal DNA in the dinucleosome substrate model. There was an apparent electron density for extranucleosomal DNA on one side of the nucleosome that was bound by LSD1/CoREST, while the electron density for DNA extension on the other side was weak. Therefore, I modeled in 22 bp of B-form extranucleosomal DNA on one side of the nucleosome (Figure 4-15). The structure also suggests the possibility that LSD1/CoREST undergoes conformational change through loop regions between LSD1 AOD and Tower domain when binding to its nucleosomal substrate (Figure 4-18).



**Figure 4-15**. Overview of the LSD1/CoREST/nucleosome structure in the mononucleosome substrate model from the molecular replacement solution. Nucleosome and 22 bp DNA extension (gray) are depicted in surface representation, and LSD1/CoREST is depicted in cartoon representation with its domains colored as shown.

Even though my mononucleosome substrate model shows one LSD1/CoREST bound per one nucleosome, modeling in the second LSD1/CoREST complex molecule into the structure does not reveal any steric clashes (Figure 4-16). It is plausible that I have crystallized the 1:1 LSD1/CoREST:nucleosome complex, and that two LSD1/CoREST molecules may bind and act on the nucleosome simultaneously in solution.



**Figure 4-16**. Model for 2:1 LSD1/CoREST:nucleosome complex. The LSD1/CoREST molecule 2 and 22 bp DNA extension are modeled into the mononucleosome substrate model. The nucleosome and the DNA extensions (gray) are depicted in surface representation, and LSD1/CoREST molecules are depicted in cartoon representation in respective colors.

The dimerization of the LSD1/CoREST complex apparent in the dinucleosome substrate model was absent in the mononucleosome substrate model. However, many of the LSD1/CoREST interactions with the nucleosome highlighted in the dinucleosome substrate model were observed in the mononucleosome substrate model as well: CoREST $\Delta$ 1 N-terminusextranucleosomal DNA, and CoREST SANT2-nucleosome disk face interactions (Figure 4-15). The mononucleosome substrate model was strikingly similar to LSD1/CoREST in the dinucleosome substrate model oriented on the nucleosome 2. This resemblance could be visualized by removing the nucleosome 1 and three other LSD1/CoREST molecules in the dinucleosome substrate model (Figure 4-17c, d). The fact that same interaction surfaces and the orientation of LSD1/CoREST on nucleosomes are observed in different crystal forms suggest that these interactions are not due to crystal artifacts.



**Figure 4-17**. Comparison of the di- and mononucleosome substrate models. (a) The dinucleosome substrate model showing just one LSD1/CoREST molecule oriented on the N16N dinucleosome. The LSD1 and CoREST proteins are shown in yellow and red, respectively. (b) LSD1/CoREST oriented on the nucleosome 1 in the dinucleosome substrate model. (c) LSD1/CoREST oriented on the nucleosome 2 in the dinucleosome substrate model. (b) The mononucleosome substrate model. The LSD1 and CoREST proteins are shown in blue and green respectively.

#### 4.3.4.2 LSD1/CoREST undergoes a conformational change when binding to the nucleosome

When the LSD1/CoREST complex (PDB ID: 2V1D; LSD1/CoREST/H3-peptide structure from Forneris *et al.*, 2007) was used as a whole as a search model in the molecular replacement solution from the crystals of LSD1/CoREST complexed with 189 bp (22N22) nucleosome, the CoREST SANT2 domain clashes with the LSD1 AOD domain of the crystal symmetry mate. Therefore the SANT2 domain could not be positioned in the electron density correctly and was situated above and away from the nucleosome disk face (Figure 4-18b). Therefore, I divided up LSD1/CoREST into two segments (AOD domain and Tower domain) for the molecular replacement solution calculations: segment 1 = LSD1(171-405, 517-836) and segment 2 = LSD1(406-516) + CoREST(308-440) (Figure 4-16c, d). Molecular replacement solution with LSD1/CoREST segments allowed both the AOD and SANT2 domains to be positioned correctly into the electron density without steric clashes (data not shown).



**Figure 4-18.** LSD1/CoREST undergoes a conformational change when binding to the nucleosome. (a), (b) Alignment of LSD1/CoREST complex from LSD1/CoREST/nucleosome molecular replacement solution (darker colors) with the previously determined structure of LSD1/CoREST bound to H3 peptide (lighter colors, PDB ID: 2V1D). (c), (d) Segmentation of LSD1/CoREST complex to solve the molecular replacement solution. Segment 1 (LSD1 171-405, 517-836) is shown in green and segment 2 (LSD1 406-516 + CoREST 308-440) is shown in blue. Two hinge loops (hinge loop 1 = LSD1 398-405; hinge loop 2 LSD1 513-520) are indicated in red. Nucleosomes (gray) and LSD1/CoREST are represented in cartoon representation.

Alignment of the LSD1/CoREST bound to the nucleosome and the LSD1/CoREST bound to H3-peptide (Figure 4-18a, b) suggests that the enzyme undergoes a conformational change, possibly though two hinge loops LSD1 398-405 and 513-520 (Figure 4-18c, d). Flexing of the hinge loops also results in shifting down of the LSD1 Tower and CoREST linker that connects the SANT2 domain towards the nucleosome disk face (Figure 14-18b). This observation is in line with the molecular dynamics simulation studies which indicate that the distance between the LSD1 SWIRM and the CoREST SANT2 oscillate significantly in solution due to molecular flexibility. The different relative position of the LSD1 AOD domain with respect to the nucleosome due to the conformational change also brings the H3 tail of the nucleosome closer to the LSD1 catalytic pocket (~63 Å in the dinucleosomes substrate model versus ~49 Å in the mononucleosome substrate model) . However, I cannot rule out the possibility that the observed conformational change is an artifact of crystallization.

# 4.3.4.3 CoREST – nucleosome disk face interaction

The interaction between CoREST SANT2 and the nucleosome disk face observed in the dinucleosome substrate model is also conserved in the mononucleosome substrate model. Likewise, the SANT2 domain is situated above the histone H4 globular domains and also nucleosomal DNA approximately 10 bp away from the nucleosome dyad (Figure 4-19a). The CoREST residues Lys378 and Lys382 are positioned close to DNA backbone of nucleosomal DNA (Figure 4-19b). The residue Lys378 is conserved as either lysine or arginine in *Drosophila, Xenopus,* zebrafish, and human, and Arg382 is perfectly conserved among four species (Figure 4-20). Conservation of these positively-charged residues across species indicates an importance of these residues for CoREST's function. Residue Lys418 in CoREST SANT2 is conserved as either lysine or arginine in four species (Figure 4-20) and the studies using the NMR spectroscopy has

indicated that this residue can bind to free DNA (Yang *et al.*, 2006). My structure suggests that this residue is in close proximity to interact with either histone H4 or nucleosomal DNA ~ 6 bp away from the nucleosome dyad (Figure 4-19c). The CoREST residues Arg425, Arg426, and Arg427 are directly above the disk face of histone H3 and H4. These triple arginine residues are perfectly conserved in D*rosophila, Xenopus,* zebrafish and human (Figure 4-20). In particular, Arg426 is very close to H3 Asp77, and Arg427 is in close proximity to H4 Asp24 (Figure 4-19d). The synergistic interaction made by these five residues may be responsible for recognizing the nucleosome disk face and for orienting the LSD1/CoREST complex on the nucleosome for catalysis.



**Figure 4-19.** CoREST SANT domain is in close proximity to the histone face and nucleosomal DNA. (a) CoREST SANT2 – nucleosome disk face interaction with relevant side chains highlighted. (b), (c), (d) Zoomed views showing residues predicted to be involved in interaction. LSD1/CoREST and histones are depicted in cartoon representation and DNA is depicted in stick representation.

		SANT2 domain				
		α1	α2	α3	α4	
	* *			*	***	
Human	356 NSALKEKLDGGIEPYRLPEVIQKCNARWT	T <mark>EE</mark> QLLAVQAIRKYGF	RDFQAISDVIGNK	SVVQV <mark>K</mark> NFFV	NYRRRFNIDEVLQEW	AE 440
Zebrafish	329 NSALREKLDTGVDEFRPSESNQKFNTRWT	T <mark>EE</mark> QLLAVQAIRKYGF	RDFQAISDVIGNK	SVVQVKNFFV	NY <mark>RRR</mark> FNL <mark>DE</mark> VLQEW	AE 413
Drosophila	292 AAANT <mark>E</mark> SAQPRISARWL	P <mark>DE</mark> IQVALLAI <mark>RE</mark> YG	(NFPTIAKVVATK	T <mark>E</mark> AHVRTFYL	NNRRRYNLDQIVKEY	A- 364
Xenopus	305 NSAFKEKLQGGIEDYRIQEVSQKFNARWT	T <mark>EE</mark> QLLAVQAIRMYGF	RDFQAISDVIGNK	SVVQV <mark>K</mark> NFFV	NY <mark>RRR</mark> FNI <mark>D</mark> QVLQ <mark>E</mark> WI	AE 389
positively charged residue						

\* residue discussed in text

**Figure 4-20**. Sequence alignment of CoREST SANT domains from four orthologs: human, zebrafish, *Drosophila*, and *Xenopus* (UniProt accession codes: Q9UKL0, A1L228, Q59E36, and Q90WN5 respectively). Positively and negatively charged residues are colored in blue and red respectively. Key residues discussed in text are indicated with an asterisk.

# 4.3.4.4 LSD/CoREST – extranucleosomal DNA interaction

The N-terminus of CoRESTA1 that is predicted to bind to the linker DNA in the dinucleosome substrate model is suggested to interact with the extranucleosomal DNA in the mononucleosome substrate model (Figure 4-21). The CoREST residues Arg308, Lys309 and Lys312 are in close contact with DNA extension 5-10 bp from the end of the end of nucleosomal DNA (Figure 4-21). However, due to its flexible nature, the N-terminus of CoRESTA1, including the region 286-307 not shown in the structure, is rich in basic residues and could make extensive interactions with extranucleosomal DNA. Through the interactions of CoREST SANT2 with the nucleosome disk face and the N-terminus with extranucleosomal DNA, CoREST could function not only to position LSD1 in the right orientation for catalysis, but also to recognize a nucleosome in open chromatin and to drive specificity of LSD1's demethylation of H3 K4 for transcriptional repression.

The residues in LSD1 AOD (Lys355, Lys357, and Lys359) suggested to interact with the end of nucleosomal DNA in the dinucleosome substrate model are shown to interact with extranucleosomal DNA in the mononucleosome substrate model (Figure 4-21a). In addition to Lys355, Lys357, and Lys359, the LSD1 residues Lys372, Arg568 and Lys699 in AOD are also suspected to interact with DNA extension ~13-22 bp from the end of the nucleosomal DNA

(Figure 4-21c, d). Specifically, the structure suggests that Lys372 located on the end of helix bind to DNA extension ~13 bp, Lys355, Lys357, and Lys359 bind to DNA extension ~ 17 bp, and finally Arg568 and Arg699 interact with DNA extension approximately 20 bp away from the end of the nucleosomal DNA.



**Figure 4-21.** LSD1/CoREST complex makes extensive interactions with extranucleosomal DNA. (a) LSD1/CoREST-extranucleosomal DNA interaction with relevant side chains highlighted. DNA bp extension predicted to make contacts with CoREST and LSD1-AOD are indicated in light green and cyan respectively. (b) Zoomed view showing N-terminus extension of CoREST making interactions with DNA extension ~5-10 bp from the end of nucleosomal DNA. (c), (d) Zoomed views of showing residues of AOD predicted to be involved in interaction with DNA extension ~13-22 bp from the end of nucleosomal DNA. Histones, DNA, and relevant side chains are depicted in sphere representation, and LSD1/CoREST and histones are depicted in cartoon representation.

The LSD1 AOD is structurally highly similar to maize polyamine oxidase (mPAO) (Binda *et al.*, 2001) but helix 372-395 is one of the least conserved between the two proteins (Yang *et al.*, 2006). In particular, this helix is oriented differently from mPAO and is involved in making association with CoREST through the CoREST N-terminus extension and its short helix (Yang *et al.*, 2006) in addition to making contacts with H3 K14-P16 (Forneris *et al.*, 2007). Therefore, it seems plausible that the helix 372-395, unique to LSD1 AOD, plays an important role in binding to CoREST, histone H3 substrate, and also extranucleosomal nucleosomal DNA. It is also noteworthy that LSD1 Lys355, Lys357, Lys359, Arg568, and Lys699 are located on the loop region and the flexibility of the loop regions may allow them to be in closer proximity to extranucleosomal DNA. It is also conceivable that DNA bending may occur beyond ~ 18 bp DNA extension to make contacts with LSD1 AOD, which results in the expense of the binding energy, consistent with the increased dissociation constant of LSD1/CoREST on nucleosomes containing DNA longer than 181 bp (18N18) (Figure 3-4).

## 4.3.4.4 Distance between LSD1 catalytic site and H3 substrate

In contrast to the dinucleosome substrate model in which the catalytic AOD domain of LSD1 is located close to the nucleosome core particle, the mononucleosome substrate model shows that the LSD1 AOD is located rather further away and interacts with the extranucleosomal DNA. Using the same approach as in the dinucleosome substrate model, the distance between the LSD1 catalytic site and the H3 substrate was measured. The H3 residue Pro16 from the LSD1/CoREST/H3-peptide structure is now approximately 49 Å from H3 Thr45 on the nucleosome core particle (Figure 4-22). In addition, H3 tail exits nucleosome DNA gyres "behind" the extranucleosomal DNA, on the same side that LSD1 substrate binding site located

so that the H3 tail does not need to wrap around the extranucleosomal DNA to get into the active site (Figure 4-22b).



**Figure 4-22.** Distance between LSD1 substrate binding site and its target in the mononucleosome substrate model. (a),(b) The distance measured between Pro16 of H3 peptide bound in LSD1 and Trh45 of histone H3 in the nucleosome is about 49Å (calculated using PyMOL) and indicated in green. Histones, DNA, and LSD/CoREST are depicted in cartoon representation.

#### **4.4 Discussion**

My structures show how LSD1/CoREST binds to the nucleosome, albeit at low resolution. Based on my biochemical experiments, which are discussed in detail in the following chapter, I favor the mononucleosome substrate model over the dinucleosome substrate model. My structure suggests that the CoREST SANT2 domain binds to the nucleosome disk face to help orient the LSD1 AOD domain on the extranucleosomal DNA. In addition, the structure suggests that the N-terminus extension of CoREST interact with DNA extension close to the end of the nucleosomal DNA, and LSD1 AOD to interact with DNA extension that is further away.

Even though the dinucleosome and the mononucleosome substrate models propose different nucleosomal substrates for the enzyme, many of the observed interactions between the LSD1/CoREST and the nucleosome are conserved between the two models despite the different crystal forms, suggesting that these interactions are not due to crystal artifacts. The mononucleosome substrate model is comparable to LSD1/CoREST oriented on the nucleosome 2 in the dinucleosome substrate model (Figure 4-17c, d). In addition, the interactions between the CoREST SANT2 and the nucleosome disk face, and between the CoREST N-terminus extension and the DNA extension are strikingly similar. The LSD1 AOD's interaction with the end of the nucleosome substrate model distinguish one model from another. It seems conceivable that the AOD of LSD1/CoREST complexed with 161 bp (8N8) nucleosome required the longer DNA extension for interaction, but due to its absence, it bound to the end of the nucleosomal DNA of the adjacent nucleosome instead. And this was made possible because 161 bp nucleosomes happened to pack via end-to-end contact between extranucleosomal DNA.

If the LSD1 AOD binds extranucleosomal DNA beyond ~18 bp away from the end of the nucleosomal DNA, as suggested in the mononucleosome substrate model, why did lengthening

nucleosomes symmetrically to 185 (20N20) or 207 bp (31N31) decrease the binding affinity (shown in Chapter 3) compared to 181 bp (18N18) when nucleosome labeled on H3 K27C were used? The increase in dissociation constant with nucleosomes containing DNA length beyond 181 bp is possibly due to the expense of binding energy from the conformational change in LSD1/CoREST complex to engage the extranucleosomal DNA, as suggested by the mononucleosome substrate model (Figure 4-18). In addition, the LSD1 AOD residues predicted to make contacts with ~ 20 bp extranucleosomal DNA are located in the loop region (Figure 4-21), which could also undergo conformational changes that expend binding energy. My mononucleosome substrate model shows 22 bp DNA extension as a linear B-form DNA, but it is also possible that extranucleosomal DNA beyond 181 bp may be distorted to make contacts with LSD1/CoREST.

The SANT domain is identified as an approximately 50-amino-acid motif that is present in nuclear receptor co-repressors, and it was subsequently found in the subunits of many chromatin-remodelling complexes (Aasland *et al.*, 1996). The SANT domains in ADA2, SMRT, and c-Myb bind to histone tails whereas the SANT domains of Myb-related proteins interact with DNA (Yu *et al.*, 2003, Boyer *et al.*, 2002, Mo *et al.*, 2005). The binding studies of the CoREST SANT2 domain to histone tails suggest that the SANT domain in CoREST does not bind to isolated, unmodified histone tails (Yang *et al.*, 2006). On the other hand, studies of CoREST SANT2 domain binding to DNA fragments (Yang *et al.*, 2006, Pilotto *et al.*, 2015), and its high similarity with SANT domain of v-Myb suggest that the CoREST SANT2 domain could act as a DNA-binding domain. Although the CoREST SANT2 domain consists of 4  $\alpha$ -helices instead of 3  $\alpha$ -helices in the canonical SANT domain,  $\alpha$ -helices 1-3 in CoREST SANT2 adopt a protein fold highly similar to SANT domain of v-Myb and the DNA-binding residues in v-Myb are conserved in CoREST SANT2 as well (Yang *et al.*, 2006) (Figure 4-23a, b). The SANT domain of v-Myb binds to DNA by inserting its  $\alpha$ -helix 3 into the major groove of the DNA (Figure 4-23b). However, in my models, the corresponding  $\alpha$ -helix 3 is above the histone disk face, with Lys418, Asn419 and Arg427 possibly interacting with the nucleosomal DNA backbone and Arg426 pointing towards the  $\alpha$ -helix 1 of histone H3, suggesting a different DNA binding mode from v-Myb (Figure 4-23d, e).

The two crystal structures of chromatin remodeling factors ISW1a and Chd1 (chromodomain helices DNA-binding protein 1) in complex with DNA reveal how the SANT domains in each chromatin remodelers bind DNA (Yamada *et al.*, 2011, Sharma *et al.*, 2011). In both proteins, SANT and SLIDE domains are required for efficient nucleosome sliding and predicted to bind to extranucleosomal DNA. ISW1a residues K889, L890, R893 and T897 and Chd1 residues R1016 and K1020 in the N-terminal helix of the SANT domain contact DNA phosphate backbone (Figure 4-24c, f). To examine if the CoREST SANT2 in my model binds to DNA in a similar manner as the SANT domains in ISW1a and Chd1, I have aligned the SANT2 of LSD1/CoREST to the ISW1a and Chd1 SANT domains (Figure 4-24b, e). The surface of the N-terminal helix of the SANT domain that is responsible for interacting with DNA in ISW1a and Chd1 instead makes contacts with the LSD1 Tower domain in LSD1/CoREST complex, precluding the region from making interactions with DNA (Figure 4-24c, f). Therefore, the CoREST SANT2 domain probably does not bind DNA in the same manner as the SANT domains in ISW1a and Chd1 chromatin remodelers.



Binding mode of the CoREST SANT2 in my mononucleosome substrate model

**Figure 4-23**. Alignment of v-Myb and CoREST SANT domains. (a) Sequence alignment of the SANT2 domains of CoREST and v-Myb with their secondary structural elements indicated above and below the sequences, respectively. The conserved DNA-binding residues are labeled red, whereas other conserved residues are labeled yellow. (b) Structure of CoREST SANT2 (red) overlaid on the structure of v-Myb SANT2 (blue) bound to DNA Ribbon drawing of the structure of CoREST SANT2 (red) overlaid on the structure of v-Myb SANT2 (blue) bound to DNA Ribbon drawing of the structure of CoREST SANT2 (red) overlaid on the structure of v-Myb SANT2 (blue) bound to DNA. The conserved DNA-binding residues are shown as purple (CoREST) and yellow (v-Myb) sticks. (c) Molecular surface of CoREST SANT2 in the CoREST SANT2-DNA model in the same orientation as in (b) with the positive and negative electrostatic potentials colored blue and red, respectively. (d), (e) The actual binding mode of CoREST SANT2 in the mononucleosome substrate model. (Panels a-c adapted from Yang, M., J. C. Culhane, L. M. Szewczuk, C. B. Gocke, C. A. Brautigam, D. R. Tomchick, M. Machius, P. A. Cole and H. Yu. "Structural Basis of Histone Demethylation by Lsd1 Revealed by Suicide Inactivation." *Nat Struct Mol Biol* 14, no. 6 (2007): 535-9.)



**Figure 4-24.** Alignment of SANT2 domain of LSD1/CoREST complex with SANT domains in ISW1a and Chd1. (a) Structure of ISW1a bound to DNA depicted in cartoon representation (PDB ID: 2Y9Z). The ISW1a SANT domain is shown in green and the rest of the ISW1a is shown in light green. (b) Structure of the LSD1/CoREST (PDB ID: 2VID) with the CoREST SANT2 domain aligned to ISW1a SANT domain bound to DNA. The LSD1 protein is shown in light orange, CoREST is shown in pink, and SANT2 domain is shown in red. For visualization, the ISWIa structure except the SANT domain is not shown. (c) Zoomed view of the aligned SANT domains in (b). (d) Structure of Chd1 bound to DNA depicted in cartoon representation (PDB ID: 3TED). The Chd1 SANT domain is shown in yellow and the rest of the Chd1 is shown in light yellow. (e) Structure of the LSD1/CoREST (PDB ID: 2VID) with the CoREST SANT2 domain aligned to Chd1 SANT domain bound to DNA. The LSD1 protein is shown in light orange, CoREST is shown in pink, and SANT2 domain is shown in red. For visualization, the STED). The Chd1 SANT domain is shown in yellow and the rest of the Chd1 is shown in light yellow. (e) Structure of the LSD1/CoREST (PDB ID: 2VID) with the CoREST SANT2 domain aligned to Chd1 SANT domain bound to DNA. The LSD1 protein is shown in light orange, CoREST is shown in pink, and SANT2 domain is shown in red. For visualization, the Chd1 structure except the SANT domain is not shown. (f) Zoomed view of the aligned SANT domains in (e).

My mononucleosome substrate model is also consistent with my previous fluorescence quenching experiment. Very little or no change in fluorescence was observed for the probes incorporated into the nucleosome at H2A E91, H2A S113, and H4 Q27 upon LSD1/CoREST binding. In hindsight, it is likely that labeling at H4 Q27, which is right on the interaction surface with CoREST SANT2, inhibits binding of LSD1/CoREST to the nucleosome (Figure 4-25). Although the specific distance between the probe and the binding protein to produce fluorescence quenching is not known, both H2A E91 and H2A S113 are far away from LSD1/CoREST in the mononucleosome substrate model. However, a modest change in fluorescence was detected at the H2A D72 and H2B S120 10C label even though they seem far away from the LSD1/CoREST's binding site. It is likely that the fluorescence changes at these probes occurred due to nonspecific binding of LSD1/CoREST on the nucleosomal DNA at 50 mM NaCl in addition to probes located on H2A and H2B tails (Figure 3-2). In addition, my model shows that the LSD1/CoREST does not employ an arginine anchor motif to bind to the nucleosome acidic patch, as demonstrated by the H1-F1 binding assay performed on nucleosome containing the acidic patch mutation (Table 3-2).

Blaine Bartholomew group's photocrosslinking studies suggest that CoREST binds to DNA close to nucleosomal dyad, and also to DNA approximately one helical turn from the end of nucleosome particle end, and LSD1 binds to DNA two helical turns from the nucleosome particle end (Figure 3-7). These crosslinking results are mostly consistent with my model, with CoREST SANT2 making contact with DNA ~10 bp away from the dyad (Figure 4-26). The N-terminus of CoREST∆1is also in close proximity to DNA 9-12 bp from the end of the nucleosomal DNA, and it is plausible that the CoREST 286-307 not shown in structure may also make contacts with this region. The structure suggests that LSD1 AOD domain is responsible for making contacts with DNA 18-22 bp from the nucleosomal DNA (Figure 4-26). Crosslinking to nucleosomal DNA at

-33 and +42 bp from the nucleosome dyad likely reflects nonspecific binding of LSD1/CoREST to nucleosomal DNA because these interactions are weakened by increased salt concentrations even in the presence of extranucleosomal DNA as explained in chapter 3 (Figure 3-3). I currently cannot explain why positions +6, 0, and 2 crosslinks to CoREST while my mononucleosome substrate model shows that CoREST is positioned far away from these positions. However, the photocrosslinking studies substantiates and favors the mononucleosome over the dinucleosome substrate model.



**Figure 4-25.** Fluorescence positions on the nucleosome affected by LSD1/CoREST binding mapped onto the mononucleosome substrate model. The fluorescence quenching results from Chapter 3 are mapped onto the mononucleosome substrate model on the nucleosome core particle with Widom 601 DNA sequence (PDB ID: 3LZ0). The fluorescence quenching of probes on histone tails are not shown. The nucleosome is depicted in sphere representation and LSD1/CoREST is depicted in cartoon representation. Probe positions that produced 10-15% fluorescence quenching are indicated in pink, and probe positions that produce less than 10% fluorescence quenching are indicated in blue. The four H2A residues in the nucleosome acidic patch, E61, E64, D90 and E92, are shown in orange.

My dinucleosome substrate and mononucleosome models provide low resolution structures of LSD1/CoREST bound to nucleosomes. The two models suggest different nucleosomal substrates for LSD1/CoREST and provide ways that LSD1/CoREST can engage each substrate. Nevertheless, based on both models, it seems likely that CoREST SANT2 interacts with the nucleosome disk face above H4, and the N-terminus of CoREST makes contacts with extranucleosomal DNA to position LSD1's AOD domain to bind to H3 N-terminal tail for subsequent catalysis. Biochemical studies to validate the two models and to determine the true substrate for LSD1/CoREST will be discussed in the next chapter. My low resolution models also highlight the need for a higher resolution structure of LSD1/CoREST/nucleosome to visualize and understand precisely how LSD1/CoREST binds to its nucleosome substrate.



**Figure 4-26.** Photocrosslinking results mapped onto the mononucleosome substrate model. Positions along the phosphate backbone crosslinked to LSD1, CoREST and both LSD1/CoREST are shown in green, blue, and red. Nucleosomal DNA positions (light blue) are labeled with the base pair position from the nucleosome dyad whereas extranucleosomal DNA positions (gray) are labeled with the base pair position from the end of the nucleosome core particle.

# **4.5 Supplementary Figures**



**Supplementary Figure 4-1**. Binding curves for 147 (yellow), 161 (blue), and 181 bp (pink) nucleosomes interacting with LSD1/CoREST complex fluorescently labeled on LSD1 V415C at 75 mM NaCl. The nucleosomes of different lengths were titrated against the labeld LSD1/CoREST (kept at 4 nM).


**Supplementary Figure 4-2**. Electron density maps for (a) LSD1/CoREST/161M nucleosome and (b) LSD1/CoREST/189M nucleosome data at 1.0 $\sigma$ . The molecular replacement solutions of dinucleosome substrate model (yellow) and mononucleosome substrate model (purple) are superposed into each electron density map (gray). DNA extensions beyond the nucleosome core particle, which are not part of the input models (nucleosome (PDB ID: 3LZ0), LSD1-CoREST (PDB ID: 2V1D)), are shown in electron density map, suggesting that the solutions are probably substantially correct.

# 4.6 Supplementary Table

	LSD1/CoREST/161 bp nucleosome	LSD1/CoREST/189 bp nucleosome
	Dinucleosome substrate model	Mononucleosome substrate model
Data collection		
Space group	P 62 2 2	P 21 21 21
Cell dimensions		
a, b, c (Å)	212.1, 212.1, 488.4	103.0, 172.8, 331.4
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	7.24	5.29
<i>R</i> <sub>pim</sub>	0.023 (0.462)	0.033 (0.614)
<i>Ι/</i> σ ( <i>Ι</i> )	16.7 (1.6)	13.9 (1.3)
Completeness (%)	99.8 (99.4)	96.8 (84.8)
Redundancy	18.3 (18.0)	6.3 (5.3)
Number of reflections	181618	136420
Number of unique reflections	9919	21739
Molecular Replacement		
Resolution (Å)	7.25	5.31
Number of reflections	9298	20402
R <sub>work</sub> / R <sub>free</sub>	0.38/0.44	0.33/0.39
LLG	1153	2108

**Supplementary Table 4-1.** Data collection statistics determined by XDS, and molecular replacement statistics calculated by Phaser and REFMAC. Values in parentheses are for highest-resolution shell. Molecular replacement yielded single solution for each model

# 4.7 Acknowledgements

Mike Doyle, Kevin Thyne, and Bryan Tornabene prepared the recombinant mono- and dinucleosomes used in crystal trials, and also helped me express and purify the human and zebrafish LSD1/CoREST complexes.

#### **4.8 References**

- Aasland, R., A. F. Stewart and T. Gibson. "The Sant Domain: A Putative DNA-Binding Domain in the SWI-SNF and ADA Complexes, the Transcriptional Co-Repressor N-CoR and TFIIB." *Trends Biochem Sci* 21, no. 3 (1996): 87-8.
- Andrykovitch, M., K. M. Routzahn, M. Li, Y. Gu, D. S. Waugh and X. Ji. "Characterization of Four Orthologs of Stringent Starvation Protein A." *Acta Crystallogr D Biol Crystallogr* 59, no. Pt 5 (2003): 881-6.
- Baker, N. A., D. Sept, S. Joseph, M. J. Holst and J. A. McCammon. "Electrostatics of Nanosystems: Application to Microtubules and the Ribosome." *Proc Natl Acad Sci U S A* 98, no. 18 (2001): 10037-41.
- Binda, C., R. Angelini, R. Federico, P. Ascenzi and A. Mattevi. "Structural Bases for Inhibitor Binding and Catalysis in Polyamine Oxidase." *Biochemistry* 40, no. 9 (2001): 2766-76.
- Boyer, L. A., M. R. Langer, K. A. Crowley, S. Tan, J. M. Denu and C. L. Peterson. "Essential Role for the SANT Domain in the Functioning of Multiple Chromatin Remodeling Enzymes." *Mol Cell* 10, no. 4 (2002): 935-42.
- Da, G., J. Lenkart, K. Zhao, R. Shiekhattar, B. R. Cairns and R. Marmorstein. "Structure and Function of the SWIRM Domain, a Conserved Protein Module Found in Chromatin Regulatory Complexes." *Proc Natl Acad Sci U S A* 103, no. 7 (2006): 2057-62.
- Forneris, F., C. Binda, A. Adamo, E. Battaglioli and A. Mattevi. "Structural Basis of LSD1-CoREST Selectivity in Histone H3 Recognition." J Biol Chem 282, no. 28 (2007): 20070-4.
- Forneris, F., C. Binda, M. A. Vanoni, E. Battaglioli and A. Mattevi. "Human Histone Demethylase LSD1 Reads the Histone Code." J Biol Chem 280, no. 50 (2005): 41360-5.

- Kim, S. A., N. Chatterjee, M. J. Jennings, B. Bartholomew and S. Tan. "Extranucleosomal DNA Enhances the Activity of the LSD1/CoREST Histone Demethylase Complex." *Nucleic Acids Res* 43, no. 10 (2015): 4868-80.
- Makde, R. D., J. R. England, H. P. Yennawar and S. Tan. "Structure of RCC1 Chromatin Factor Bound to the Nucleosome Core Particle." *Nature* 467, no. 7315 (2010): 562-6.
- Makde, R. D. and S. Tan. "Strategies for Crystallizing a Chromatin Protein in Complex with the Nucleosome Core Particle." *Anal Biochem* 442, no. 2 (2013): 138-45.
- Mo, X., E. Kowenz-Leutz, Y. Laumonnier, H. Xu and A. Leutz. "Histone H3 Tail Positioning and Acetylation by the C-Myb but Not the V-Myb DNA-Binding SANT Domain." *Genes Dev* 19, no. 20 (2005): 2447-57.
- Pilotto, S., V. Speranzini, M. Tortorici, D. Durand, A. Fish, S. Valente, F. Forneris, A. Mai, T. K. Sixma, P. Vachette and A. Mattevi. "Interplay among Nucleosomal DNA, Histone Tails, and Corepressor Corest Underlies LSD1-Mediated H3 Demethylation." *Proc Natl Acad Sci U S A* 112, no. 9 (2015): 2752-7.
- Qian, C., Q. Zhang, S. Li, L. Zeng, M. J. Walsh and M. M. Zhou. "Structure and Chromosomal DNA Binding of the SWIRM Domain." *Nat Struct Mol Biol* 12, no. 12 (2005): 1078-85.
- Rhodes, G. *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*. 3rd ed. San Diego, CA: Elsevier/Academic Press, 2006.
- Schalch, T., S. Duda, D. F. Sargent and T. J. Richmond. "X-Ray Structure of a Tetranucleosome and Its Implications for the Chromatin Fibre." *Nature* 436, no. 7047 (2005): 138-41.
- Sharma, A., K. R. Jenkins, A. Heroux and G. D. Bowman. "Crystal Structure of the Chromodomain Helicase DNA-Binding Protein 1 (Chd1) DNA-Binding Domain in Complex with DNA." J Biol Chem 286, no. 49 (2011): 42099-104.
- Shi, Y. J., C. Matson, F. Lan, S. Iwase, T. Baba and Y. Shi. "Regulation of LSD1 Histone Demethylase Activity by Its Associated Factors." *Mol Cell* 19, no. 6 (2005): 857-64.

- Vasudevan, D., E. Y. Chua and C. A. Davey. "Crystal Structures of Nucleosome Core Particles Containing the '601' Strong Positioning Sequence." *J Mol Biol* 403, no. 1 (2010): 1-10.
- Yamada, K., T. D. Frouws, B. Angst, D. J. Fitzgerald, C. DeLuca, K. Schimmele, D. F. Sargent and T. J. Richmond. "Structure and Mechanism of the Chromatin Remodelling Factor ISW1a." *Nature* 472, no. 7344 (2011): 448-53.
- Yang, M., J. C. Culhane, L. M. Szewczuk, C. B. Gocke, C. A. Brautigam, D. R. Tomchick, M. Machius, P. A. Cole and H. Yu. "Structural Basis of Histone Demethylation by LSD1 Revealed by Suicide Inactivation." *Nat Struct Mol Biol* 14, no. 6 (2007): 535-9.
- Yang, M., C. B. Gocke, X. Luo, D. Borek, D. R. Tomchick, M. Machius, Z. Otwinowski and H. Yu. "Structural Basis for CoREST-Dependent Demethylation of Nucleosomes by the Human LSD1 Histone Demethylase." *Mol Cell* 23, no. 3 (2006): 377-87.
- Yu, J., Y. Li, T. Ishizuka, M. G. Guenther and M. A. Lazar. "A SANT Motif in the SMRT Corepressor Interprets the Histone Code and Promotes Histone Deacetylation." *EMBO J* 22, no. 13 (2003): 3403-10.

# Chapter 5

# **Biochemical Characterization of LSD1/CoREST on nucleosomes**

# **5.1 Abstract**

My two models suggest two possible nucleosomal substrates for LSD1/CoREST: a dinucleosome with 16 bp linker, and a mononucleosome with extranucleosomal DNA. Here I describe the biochemical studies to determine the true substrate for LSD1/CoREST and to validate my models. My studies of LSD1/CoREST's enzyme activity on dinucleosomes and mononucleosomes show that LSD1/CoREST is unable to demethylate H3K4 in dinucleosomes. I have also performed mutational studies combined with the HI-FI binding assay to validate my mononucleosome substrate model. The HI-FI nucleosome binding experiments suggest that the LSD1 AOD binds extranucleosomal DNA ~1.5-2 turns from the end of the nucleosomal DNA while the N-terminus of CoRESTA1 makes contacts with the extranucleosomal DNA ~1 turn from the end of the nucleosomal DNA. In addition, I have identified the residues in the CoREST SANT2 domain crucial for LSD1/CoREST to bind to nucleosome independent of extranucleosomal DNA. These results are consistent with my mononucleosome substrate model of LSD1/CoREST/nucleosome structure.

#### **5.2 Introduction**

Currently known structural studies of LSD1/CoREST have been performed on the H3peptide or on the nucleosome core particle (Forneris *et al.*, 2007, Pilotto *et al.*, 2015). However, the better nucleosomal substrates of LSD1/CoREST contain extranucleoeomal DNA, which enhances both the activity and affinity of LSD1/CoREST to its targets (Kim *et al.*, 2015). In the previous chapter, I have described two models from the crystal structures of LSD1/CoREST on the nucleosomes with extranucleosomal DNA. To my knowledge, these are the first crystal structures of a histone modifying enzyme on a nucleosome containing one or two turns of extranucleosomal DNA albeit at low resolution.

My two models of LSD1/CoREST/nucleosome structure suggest three possibilities: 1. the mononucleosome is the true substrate, 2. the dinucleosome is the true substrate, 3. both monoand dinucleosomes are substrates for LSD1/CoREST. To distinguish among the three possibilities, I have performed the demethylase activity assay on dinucleosomes with different linker lengths and on mononucleosomes using Western Blot. I have mutated specific LSD1/CoREST residues suspected to be responsible for the interaction with the nucleosomes and measured the dissociation constants of the mutant LSD1/CoREST complexes on nucleosomes of various DNA lengths. In addition, the mononucleosome substrate model suggests the possibility that LSD1/CoREST complex undergoes a conformational change to bind to the nucleosome core particle and the extranucleosomal DNA. As a consequence, the position of the LSD1 AOD is closer to the nucleosome core particle relative to the position of the LSD1 AOD to the nucleosome 2 in the dinucleosome substrate model. To examine the minimum length of the H3 tail, and examined their effect in the context of the nucleosome. Overall, my biochemical studies support my mononucleosome substrate model and indicate that the mononucleosome with extranucleosomal DNA, but not the dinucleosome, is the true substrate for LSD1/CoREST.

### **5.3 Results**

## 5.3.1 Demethylase activity of LSD1/CoREST complex on mono- and dinucleosomes

# 5.3.1.1 LSD1/CoREST is not able to demethylate H3K4me2 in dinucleosome with 16 bp linker DNA

My dinucleosome substrate model suggests the possibility that LSD1/CoREST uses a dinucleosome contatining 16 bp linker DNA as a substrate. I therefore tested the demethylase activity of LSD1/CoREST on dinucleosome substrate by Western Blot to examine if this is a better substrate than mononucleosomes with 147 bp and/or extranucleosomal DNA. Recombinant nucleosome containing a methyl-lysine analog of dimethylated H3 Lys4 (H3 $K_{C}$ 4) was reconstituted with 306 bp nucleosomal DNA that contains the two identical Widom 601 nucleosome positioning sequences. Since this 306 bp DNA contains the two 601 positioning sequences at each end, it constitutes a N16N positioning sequence where N represents the central 145 bp 601 nucleosome core particle sequence, with 16 bp linker in between. Nucleosomes containing the histone H3K4me2 methyl-lysine analog were prepared with 147 (1N1), 161 (8N8), and 177 bp (16N16) and LSD1/CoREST's catalytic activity on these substrates were tested. If the dinucleosome was a substrate for LSD1/CoREST, the demethylase activity on dinucleosome was expected to be significantly higher than on the 147 bp nucleosome. In addition, if the dinucleosome was a better substrate than mononucleosomes with extranucleosomal DNA, LSD1/CoREST's activity was expected to be higher on N16N dinucleosome than on 161 (8N8) or 177 bp (16N16) mononucleosomes. My results showed that LSD1/CoREST was not able to demethylate H3K4 in N16N dinucleosome during 15 minutes (Figure 5-1a). This represented a similar demethylase activity for LSD1/CoREST on 147 bp (1N1) nucleosome. In contrast, a dramatically increased activity was detected on 161 and 177 bp mononucleosomes, consistent

with the previous results that lengthening the extranucleosomal DNA on nucleosomes enhances

LSD1/CoREST's activity (Figure 5-1a).



**Figure 5-1**. Demethylase activity of LSD1/CoREST on mono- and dinucleosome substrates using Western Blot with an antibody against H3K4me2. (a) Demethylase activity of LSD1/CoREST on 147 (1N1), 161 (8N8), and 177 bp (16N16) nucleosomes and 306 bp (N16N) dinucleosome with after 0, 10, and 15 minutes. (b) Demethylase activity of LSD1/CoREST on 301 (N11N), 306 (N16N), 311 (N21N), and 316 bp (N26N) dinucleosomes after 0, 15, 30 minutes, and 0, 1, and 7 hours.

# 5.3.1.2 Dinucleosomes with 11, 16, 21, and 26 bp linker DNA are not substrates for LSD1/CoREST complex

Considering the unexpectedly undetectable demethylase activity of LSD1/CoREST on

the N16N dinucleosome, I asked if shortening or lengthening the linker DNA in the dinucleosome

would enhance the activity of LSD1/CoREST. Since my mononucleosome substrate model

suggests that the LSD1 AOD engages the DNA extension ~20 bp from the end of the

nucleosomal DNA, I anticipated that lengthening the linker length beyond 20 bp would enable

LSD1/CoREST to perform its activity. Nucleosomal DNA of 301 (N11N), 311 (N21N), and 316

bp (N26N) were prepared and reconstituted into dinucleosomes containing the same preparation of histone H3K4me2 methyl-lysine analog. LSD1/CoREST was unable to demethylate H3K4 in any of the dinucleosomes even when the assay duration was extended for seven hours (Figure 5-1b).

My results indicated that either the dinucleosomes are not true substrates for LSD1/CoREST or that the prepared recombinant dinucleosomes with H3K4me2 methyl-lysine analog are defective. To address the latter problem, I examined the integrity of these dinucleosomes by testing if they can be ubiquitylated by the PRC1 ubiquitylation module. The surfaces formed by the nucleosome are required for the E3 ligase activity of PRC1, and H2A– H2B dimers and H2A alone are not substrates for the PRC1 ubiquitylation module (Bentley *et al.*, 2011, Olsen and Lima, 2013, McGinty *et al.*, 2014). This enzyme, which consists of E3 ubiquitin ligase subunits, Ring1B and Bmi1, was shown to ubiquitylate histone H2A Lys119 in 147 bp nucleosomes with an E2 ubiquitin-conjugating enzyme, Ubc H5c in our laboratory (McGinty *et al.*, 2014). Both the wild-type and dimethylated dinucleosomes with N16N DNA are ubiquitylated by the Ring1B/Bmi1 complex, demonstrating that these dinucleosomes are not defective (Figure 5-2). Interestingly, the dinucleosomes were mostly mono-ubiquitylated whereas the 147 bp nucleosome was poly-ubiquitylated. I do not have an explanation for this observation.



**Figure 5-2**. Ubiquitylation assay on mono- and dinucleosomes. Coomassie blue stained gel of ubiquitylation assay of Bmi1/Ring1B complex on 147 bp (1N1) wild-type nucleosomes, 306 bp (N16N) wild-type dinucleosome, and 306 bp (N16N) H3K4me2 dinucleosome.

#### 5.3.2 Binding studies of LSD1/CoREST complex on mono- and dinucleosome

# 5.3.2.1 LSD1/CoREST complex has higher binding affinity on the mononucleosome with 181 bp DNA than on the dinucleosomes

My activity assay indicated that LSD1/CoREST is not able to demethylate H3K4me2 in N11N, N16N, N21N and N26N dinucleosomes. I next compared the LSD1/CoREST's ability to bind to mononucleosomes and dinucleosomes with different linker lengths using the HI-FI nucleosome binding assay. I selected the same position (H3 K27C) in the nucleosome to use in my LSD1/CoREST nucleosome binding assay. In addition to the 306 bp (N16N) nucleosomal DNA, I prepared 322 bp (8N16N8) DNA containing the two Widom 601 positioning sequence with 16 bp linker and 8 bp DNA extension on the other end of each nucleosome core particle sequences.

I determined that LSD1/CoREST binds to 147 bp nucleosomes with a dissociation constant of  $3.15 \,\mu\text{M}$  at 75 mM NaCl. LSD1/CoREST bound to 157 (6N6) and 181 bp (18N18) nucleosomes with dissociation constants of 862 ± 27 and 392 ± 31 nM respectively (Table 5-1), with the similar trend of increasing binding affinity with an increase in DNA length observed previously at 50 mM NaCl (Table 3-2). The 322 bp (8N16N8) dinucleosome resulted in a dissociation constant of 502 ± 51 nM, weaker affinity than the 181 bp nucleosome, but higher affinity than the 147 bp and 157 bp nucleosomes. (Figure 5-3a). Removing the 8 bp DNA extension on each end of the nucleosome core particle from the 322 bp (8N16N8) dinucleosome to 306 bp (N16N) dinucleosome decreased the dissociation constant only slightly to 531 ± 21 nM (Table 5-1). The fact that LSD1/CoREST is able to bind to N16N dinucleosome but is unable to remove methyl groups from H3K4 in the same dinucleosome suggests that the enzyme binds to this dinucleosome in a non-productive mode.



**Figure 5-3.** Binding affinity of LSD1/CoREST towards dinucleosomes. (**a**) Binding curves for LSD1/CoREST interacting with 147 (yellow), 157 (blue), 181 bp (pink) mono- and 322 bp (green) dinucleosome fluorescently labeled on H3 K27C at 75mM NaCl. (**b**) Equivalent binding curves for LSD1/CoREST interacting with 157 bp (blue) mono-, 301 bp (red) and 322 bp (green) dinucleosomes.

DNA	Nucleosome	DNA	K <sub>d</sub> (nM)	relative to	Hill
length		format		147 bp	coefficient
147	mono	1N1	3152 ± 261	1.0	0.88 ± 0.08
157	mono	6N6	862 ± 27	3.7	0.96 ± 0.05
181	mono	18N18	392 ± 31	8.0	0.77 ± 0.02
185	mono	20N20	710 ± 83	4.4	0.67 ± 0.02
189	mono	22N22	410 ± 61	7.7	0.60 ± 0.03
197	mono	26N26	777 ± 161	4.1	0.73 ± 0.03
301	di	N11N	1576 ± 350	2.0	0.44 ± 0.01
306	di	N16N	531 ± 21	5.9	0.67 ± 0.04
311	di	N21N	1295 ± 72	2.4	0.49 ± 0.05
316	di	N26N	1082 ± 115	2.9	0.55 ± 0.03
322	di	8N16N8	502 ± 51	6.3	$0.82 \pm 0.06$

**Table 5-1.** Dissociation constants and Hill coefficients for LSD1/CoREST complex binding to mono- and dinucleosomes with different length linker DNA

# 5.3.2.2 Negative cooperative binding of LSD1/CoREST to dinucleosomes

Next, I questioned if changing the linker lengths in dinucleosomes would affect the

binding affinity of LSD1/CoREST, as lengthening the extranucleosomal DNA enhanced the

binding affinity of LSD1/CoREST towards the mononucleosomes. Dinucleosomes labeled on H3 K27C containing the 301 (N11N), 311 (N21N), and 316 bp (N26N) DNA were tested in the HI-FI nucleosome binding assay. Both shortening the linker length to 11 bp and lengthening it to 21 or 26 bp decreased the binding affinity of LSD1/CoREST towards dinucleosomes (Table 5-1). But more interestingly, the binding curves for dinucleosomes with 11, 21, and 26 bp linkers fit Hill coefficient  $(n_H)$  values significantly lower than 1, indicating negative cooperativity of LSD1/CoREST binding to these dinucleosomes. The Hill coefficients of 0.44, 0.49, and 0.55 for N11N, N21N and N26N dinucleosomes, respectively, are significantly lower than the Hill coefficients calculated for any other mononucleosome or dinucleosome binding curves (Table 5-1). Previous HI-FI binding assay indicated Hill coefficients of 0.88, 0.96 and 0.77 for 147, 157, and 181 bp nucleosomes, and 0.67 and 0.82 for 306bp (N16N) and 322bp (8N16N8) dinucleosomes, respectively, at 75 mM NaCl and therefore a Hill coefficient of 1.0 was used for subsequent analysis. However, a significant deviation of Hill coefficients for dinucleosomes with 11, 21, and 26 bp linker from 1.0 value prevented the binding curves to be fit with  $n_H$  of 1.0. Figure 5-3b shows that LSD1/CoREST binds to 301 bp (N11N) dinucleosome with negative cooperativity indicated by reduced slope steepness and with lower affinity compared to 157 bp mono- and 322 bp dinucleosomes. LSD1/CoREST binds the 311 (N21N) and 316 bp (N26N) dinucleosomes with similar Hill coefficient and binding affinity as 301 bp (N11N) dinucleosome (Table 5-1). These results suggest that LSD1/CoREST prefers 181 bp (18N18) mononucleosomes over dinucleosomes for binding and that dinucleosomes with 11, 21, and 26 bp linkers induce negative cooperativity binding of LSD1/CoREST.

Previously, I have also observed the decreased affinity for 185 (20N20) and 207 bp (31N31) nucleosomes compared to 181 bp (18N18) nucleosomes, which bound LSD1/CoREST with the highest affinity. However, my mononucleosome substrate model suggests that LSD1/CoREST binds DNA extension beyond 18 bp from the end of the nucleosome particle,

with the LSD1 AOD possibly interacting with DNA extension ~ 22 bp, which corresponds to 189 bp (22N22) of symmetrically extended nucleosome. To examine the binding affinities of LSD1/CoREST towards nucleosomes with symmetrically extended DNA between 181 bp and 207 bp, I have performed the HI-FI binding assay with nucleosomes labeled at H3 K27C at 75 mM NaCl. I determined that LSD1/CoREST binds 185 (20N20) and 197 bp (26N26) nucleosomes with dissociation constants of 710 nM and 777 nM respectively (Table 5-1), or ~2x weaker affinity compared to the 181 bp (18N18) nucleosomes. However, the 189 bp (22N22) nucleosomes bound LSD1/CoREST with similar affinity as 181 bp nucleosome, with a dissociation constant of 410 nM. It was surprising that adding an additional 2 bp on either side to 185 bp nucleosome produced an almost ~2-fold greater affinity, suggesting that these 2 additional bp engages LSD1/CoREST productively. The decreased binding affinity with 185 bp nucleosomes compared to 181 bp nucleosomes, and the similar trend with 197 bp nucleosomes compared to the 189 bp nucleosomes despite the longer extranucleosomal DNA suggests for either conformational change in LSD1/CoREST or DNA distortion in DNA extension beyond 18 bp, or combination of both when LSD1/CoREST binds to nucleosomes with substantially longer extranucleosomal DNA. The expense of binding energy due to conformational change or DNA distortion is compensated by tighter binding with 189 bp nucleosomes compared to 185 bp nucleosomes, but additional contacts are not made beyond 22 bp DNA extension for 197 bp nucleosomes.

### 5.3.3 Binding studies of LSD1/CoREST mutant complexes

## 5.3.3.1 Dimerization mutations

My dinucleosome substrate model suggests the possibility that the LSD1/CoREST binds to dinucleosome substrates as a dimer. To examine if the dimerization is obligatory for LSD1/CoREST's interaction with dinucleosomes, I have created the LSD1/CoREST complexes containing mutations at the dimerization interface (Figure 4-9). In particular, the residues suspected to mediate dimerization were mutated to tyrosine, a bulky residue that can disrupt dimer formation. My HI-FI binding results show that LSD1/CoREST complexs containing either the LSD1(K432Y) or the CoREST(E341Y/Q348Y) mutations bind to both 181 bp (18N18) mononucleosomes and 306 bp (N16N) dinucleosomes with similar affinity as the wild-type LSD1/CoREST at 100 mM NaCl (Figure 5-4, Table 5-2). These results suggest that dimerization is not required for LSD1/CoREST to bind to both mono- and dinucleosomes, and that dimerization may be an artifact of crystallization in the dinucleosome substrate model.



**Figure 5-4**. Effect of the dimerization mutations on the binding affinity of LSD1/CoREST towards mono- and dinucleosomes. The dissociation constants determined by the HI-FI binding assay at 100 mM NaCl are shown as bar graphs.

Mutations	DNA	DNA format	K <sub>d</sub> (nM)	relative to WT
	length			LSD1/CoREST
WT	181	18 + 145 + 18	1070	1.0
LSD1 K432Y	181	18 + 145 + 18	1090	1.0
CoREST E341Y/Q348Y	181	18 + 145 + 18	1260	1.2
WT	306	145 + 16 + 145	1600	1.0
LSD1 K432Y	306	145 + 16 + 145	1620	1.0
CoREST E341Y/Q348Y	306	145 + 16 + 145	1170	0.72

**Table 5-2.** Dissociation constants for dimerization mutant LSD1/CoREST complexes binding to nucleosomes at 100 mM NaCl.

#### 5.3.3.2 Mutations in CoREST SANT2



Both my mononucleosome and dinucleosome substrate models suggested that the CoREST SANT2 binds to the nucleosome disk face and nucleosomal DNA ~10 bp away from the nucleosome dyad. In particular, the CoREST residues Lys378 and Arg382 are positioned close to DNA backbone of nucleosomal DNA and the residues Arg425, Arg426, and Arg427 are above the disk face of histone H3 and H4 (Figure 4-13, Figure 4-19). To examine the effect of these residues, I have created LSD1/CoREST complexes containing either the K378E/R382E double mutations or the R425A/R426A/R427A triple mutations and tested their ability to bind to 147 (1N1), 157 (6N6), 181 bp (18N18) mononucleosomes and 322 bp (8N16N8) dinucleosomes with the HI-FI nucleosome binding assay. LSD1/CoREST containing the K378E/R382E mutations bound to 147 (1N1), 157 (6N6), and 181bp (18N18) nucleosomes with ~15x weaker affinity compared to the wild-type LSD1/CoREST complex (Figure 5-5a, Table 5-3), indicating an important role of these two basic residues in binding to nucleosomes. The dissociation constants of this mutant towards all 147, 157 and 181 bp nucleosomes could not be determined accurately as no plateau of fluorescence quenching was achieved even with the highest concentration of LSD1/CoREST complex used in the assay (Figure 5-5a), characteristic of weak interactions. Similarly, LSD1/CoREST complex containing the R425A/R426A/R427A mutations bound to 147 (1N1), 157 (6N6), 181 (18N18), and 322 bp (8N16N8) nucleosomes with ~5x weaker affinity compared to the wild-type LSD1/CoREST complex (Figure 5-5b, Table 5-3). Dynamic light scattering was used to check for aggregation of these and subsequent mutant LSD1/CoREST complexes, and the polydispersity was below 15%, indicating the samples were monodispersed, and the reduced binding is not due to protein aggregation.

Both sets of mutations decreased the ability of LSD1/CoREST to bind to 147 bp nucleosome significantly, and the DNA extension enhanced the affinity of the mutant complexes to nucleosomes, suggesting that CoREST residues K378/R382 and R425/R426AR427 likely bind to the nucleosome core particle and not to extranucleosomal DNA.



**Figure 5-5.** Effect of mutations in the CoREST SANT2 on binding affinity of LSD1/CoREST towards nucleosomes. (a) Effect of K378E/R382E mutations in CoREST SANT2. Binding curves for wild-type LSD1/CoREST interacting with 147 (yellow), 157 (blue), and 181 bp (pink) nucleosomes at 75mM NaCl are indicated by solid lines and binding curves for mutant LSD1/CoREST interacting with 147 (green), 157 (purple) and 181 bp (red) nucleosomes are indicated by dashed lines. (b) Effect of R425A/R426A/R427A mutations in CoREST SANT2. Binding curves for wild-type LSD1/CoREST interacting with 181 bp mononucleosomes (pink) and 322 bp dinucleosomes (green) are indicated by solid lines and binding curves for mutant LSD1/CoREST interacting with 181 bp mononucleosomes (yellow) and 322 bp dinucleosomes (green) are indicated by solid lines and binding curves for mutant LSD1/CoREST interacting with 181 bp mononucleosomes (yellow) and 322 bp dinucleosomes (green) are indicated by solid lines and binding curves for mutant LSD1/CoREST interacting with 181 bp mononucleosomes (yellow) and 322 bp dinucleosomes (green) are indicated by solid lines and binding curves for mutant LSD1/CoREST interacting with 181 bp mononucleosomes (yellow) and 322 bp dinucleosomes (blue) are indicated by dashed lines.

CoREST mutations	DNA	DNA DNA format		relative to WT
	length			LSD1/CoREST
K378E/R382E	147	1 + 145 + 1	> 23000	>7
K378E/R382E	157	6 + 145 + 6	11903	14
K378E/R382E	181	18 + 145 + 18	6732	17
R425A/R426A/R427A	147	1 + 145 + 1	> 17000	>5
R425A/R426A/R427A	157	6 + 145 + 6	3879 ± 827	4.5
R425A/R426A/R427A	181	18 + 145 + 18	2216 ± 447	5.7
R425A/R426A/R427A	322	8 + 145 + 16 + 145 + 8	3009	6.0

**Table 5-3.** Dissociation constants for SANT2 mutant LSD1/CoREST complexes binding to nucleosomes with different length linker DNA.

### 5.3.3.3 Truncations in CoREST N-terminus



Both my mono- and dinucleosome substrate models suggest a role for the N-terminus of CoREST $\Delta 1$  in binding to the DNA extension ~5-10 bp away from the end of the nucleosomal DNA. HI-FI binding assays comparing LSD1/CoREST $\Delta 2$  [CoREST $\Delta 2$  = CoREST(308-482)] with the equivalent LSD1/CoREST complex without the CoREST N-terminal deletion showed overall ~2-fold decrease in binding affinity towards nucleosomes with 147 (1N1), 157 (6N6) and 181 bp (18N18) nucleosomes (Table 4-1). However, similar to wild-type LSD1/CoREST, adding on additional 5 bp and 17 bp on either side to 157 (6N6) and 181 bp (18N18) nucleosome decreased the dissociation constant ~3.5 x and ~7x for LSD1/CoREST $\Delta 2$ , suggesting that CoREST residues 286-307 are not responsible for binding to 5-10 bp DNA extension.

To examine if the CoREST residues beyond 308 in the N-terminus of CoREST $\Delta 1$  are responsible for interaction with extranucleosomal DNA, I have created a LSD1/CoREST complex containing the truncated CoREST [CoREST $\Delta 4$  = CoREST(316-482)] (Figure 5-6a, b) and tested its binding affinity to the nucleosomes with various lengths of extranucleosomal DNA extended symmetrically. A gradual increase in binding affinity with increasing DNA lengths from 147 (1N1), 151 (3N3), 155 (5N5), 157 (6N6), 159 (7N7), 161 (8N8), to 181 bp (18N18) was observed with the wild-type LSD1/CoREST (Figure 5-6c). However, this concomitant increase in binding affinity was abolished by removing the residues 286-315 in CoREST. Extending the 147 bp DNA fragment by 2 bp on either side to 151 bp (3N3) resulted in a  $K_d$  of 6488 nM, ~ 1.5x greater affinity compared to  $K_d$  of 8917 nM for 147 bp nucleosome (Figure 5-6c, Table 5-4), the same trend observed with the wild-type LSD1/CoREST. However, adding additional 4-7 bp on either side to 147 bp nucleosome did not affect the dissociation constants for LSD1/ CoREST $\Delta$ 4 in contrast to the wild-type LSD1/CoREST. In particular, the LSD1/CoRESTA2 truncation mutant bound the 155 (5N5) and 157 bp (6N6) nucleosomes with weaker affinity than the 151 bp (3N3) nucleosome, suggesting that the CoREST residues 286-315 interact with DNA extension 5-6 bp away from the end of the nucleosomal DNA (Table 5-4). In addition, further lengthening the extranucleosomal DNA to 159 (7N7) and 161 bp (8N8) did not enhance the binding affinity of LSD1/CoRESTA4 significantly, with ~1.5-fold increase in affinity (Table 5-4). On the other hand, the binding affinity was increased by 2.4-fold with 181 bp (18N18) nucleosome, suggesting that the region other than the N-terminus of CoREST $\Delta$ 1 interacts with the DNA extension ~18 bp from the end of the nucleosomal DNA. In addition, the dissociation constant of 322 bp (8N16N8) dinucleosome (4498 nM) was similar to the  $K_d$  of 161 bp (8N8) nucleosome (5520 nM) for this truncation mutant complex.



**Figure 5-6**. Effect of N-terminus truncation in CoREST on the binding affinity of LSD1/CoREST towards nucleosomes with different length linker DNA. (a) Amino acid sequence of N-terminus of CoREST $\Delta$ 1. The basic residues are colored in red and the residues removed in the CoREST $\Delta$ 4 are underlined. (b) Schemes representing the constructs for CoREST $\Delta$ 1 and CoREST $\Delta$ 4 (c) Summary of fold increase in binding affinity of the wild-type (orange) and LSD1/CoREST $\Delta$ 4 complexes for 8 nucleosomes with different extranucleosomal lengths.

DNA length	DNA format	K <sub>d</sub> (nM)	relative to 147 bp
147	1 + 145 + 1	8917 ± 355	1.0
151	3 + 145 + 3	6488 ±1004	1.4
155	5 + 145 + 5	7217 ±1889	1.2
157	6 + 145 + 6	7359 ± 373	1.2
159	7 + 145 + 7	5795	1.5
161	8 + 145 + 8	5520 ± 2941	1.6
181	18 + 145 + 18	3697 ± 327	2.4
322	8 + 145 + 16 + 145 +8	4498	2.0

**Table 5-4**. Dissociation constants for LSD1/CoREST $\Delta$ 4 complex binding to nucleosomes with different length linker DNA.

#### 5.3.3.4 Mutations in LSD1 AOD



The mononucleosome substrate model suggests the possibility that the LSD1 AOD makes interactions with the extranucleosomal DNA beyond one helical turn from the end of the nucleosome core particle. To examine if the LSD1 AOD is responsible for binding to extranucleosomal DNA, I have made mutations on the basic residues in this region and tested their effect on binding affinities of LSD1/CoREST complexes to nucleosomes with different DNA lengths.

The mutations of the LSD1 AOD residues (K355E/K357E/K359E), a part of the basic stripe, significantly decrease the binding affinity of LSD1/CoREST towards all nucleosomes with symmetrically extended DNA of various length (Figure 5-7a, b, Table 5-5). This mutated LSD1/CoREST complex binds the 147 bp nucleosome with ~6-fold weaker affinity compared to the wild-type LSD1/CoREST, while it binds the 177 (16N16) and 181 bp (18N18) nucleosomes with >40-fold weaker affinity compared to the wild-type complex. In addition, the LSD1/CoREST with K355E/K357E/K359E mutations binds the 177 and 181 bp nucleosomes with weaker affinity than the 147, 151, or 161 bp nucleosomes, and adding on extra DNA to 185 bp (20N20) increase the affinity of LSD1/CoREST (Figure 5-7a, b, Table 5-5), suggesting that these lysine residues are critical in making contact with ~16-18 bp DNA from the end of the nucleosomal DNA.



**Figure 5-7**. Effect of mutations in the LSD1 AOD on the binding affinity of LSD1/CoREST towards nucleosomes with different length linker DNA. (a) Binding curves for LSD1/CoREST containing LSD1 K353E/K357E/K359E mutations interacting with 161 (blue), 177 (green) and 181 bp (pink) nucleosomes fluorescently labeled on H3 K27C at 75mM NaCl. (b) Summary of fold increases in binding affinity of the mutated LSD1/CoREST (K353E/K357E/K359E) towards nucleosomes with various linker DNA relative to 147 bp nucleosome. (c) Binding curves for LSD1/CoREST containing LSD1 K372A mutation interacting with 147 (orange), 155 (blue), 177 (green) and 181 bp (pink) nucleosomes fluorescently labeled on H3 K27C at 75mM NaCl. (b) Summary of fold increases in binding affinity of the mutated LSD1/CoREST (K372A) towards nucleosomes with various linker DNA relative to 147 bp nucleosome.

DNA	DNA format	K <sub>d</sub> (nM)	relative to 147	relative to WT
length			bp	LSD1/CoREST
147	1 + 145 + 1	19367±655	1.0	6.1
151	3 + 145 + 3	15612	1.2	8.0
157	6 + 145 + 6	12099 ± 4136	1.6	14
161	8 + 145 + 8	14651 ± 1228	1.3	22
177	16 + 145 + 16	22222 ± 1964	0.9	42
181	18 + 145 + 18	17143 ± 3475	1.1	44
185	20 + 145 + 20	13303 ± 2164	1.5	19

**Table 5-5**. Dissociation constants for LSD1/CoREST complex containing LSD1 K355E/K357E/K359E mutations binding to nucleosomes with different length linker DNA. The increase in the binding affinity relative to 147 bp nucleosomes for longer DNA length and the decrease in the binding affinity relative to the wild-type LSD1/CoREST complex for each nucleosome are indicated.

DNA	DNA format	K <sub>d</sub> (nM)	relative to 147	relative to WT
length			bp	LSD1/CoREST
147	1 + 145 + 1	5932 ± 566	1.0	1.9
151	3 + 145 + 3	4009 ± 805	1.5	2.1
155	5 + 145 + 5	2501 ± 294	2.4	2.4
157	6 + 145 + 6	2121 ± 329	2.8	2.5
161	8 + 145 + 8	1581 ± 217	3.8	2.3
177	16 + 145 + 16	1391 ± 171	4.3	2.6
181	18 + 145 + 18	992 ± 79	6.0	2.5

**Table 5-6**. Dissociation constants for LSD1/CoREST complex containing LSD1 K372A mutation binding to nucleosomes with different length linker DNA. The increase in the binding affinity relative to 147 bp nucleosomes for longer DNA length and the decrease in the binding affinity relative to the wild-type LSD1/CoREST complex for each nucleosome are indicated.

The mononucleosome substrate model suggests the possibility that the LSD1 residue Lys372 located on the end of the helix (Figure 4-21c) makes contact with DNA extension ~13 bp away from the end of the nucleosomal DNA. The LSD1/CoREST complex containing the LSD1 K372A mutation decreases the binding affinity of the complex towards the nucleosomes of various DNA lengths by ~2-fold (Figure 5-7c, d, Table 5-6). However, similar to the wild-type LSD1/CoREST's trend, adding on additional DNA extension on either side to 147 bp nucleosome gradually decreases the dissociation constants, suggesting that the LSD1 K372A mutation does not impair LSD1/CoREST's interaction with ~13 bp DNA extension as expected from the mononucleosome substrate model.

#### 5.3.3.5 Length of the H3 tail required for LSD1/CoREST demethylase activity

To examine the distance required between the LSD1 catalytic site and the nucleosome core particle for LSD1/CoREST's demethylase activity, I have made series of deletions in the histone H3 tail. The distance between the LSD1 catalytic site and the H3 substrate in my dinucleosome substrate model was ~ 35 Å if the LSD1/CoREST targets the H3 substrate in the nucleosome 1, or ~63 Å if the enzyme targets H3 substrate in the nucleosome 2 (Figure 4-14). The respective distance in my mononucleosome substrate models was ~49 Å (Figure 4-22). The distance was calculated between the C-alpha positions of H3 Pro16 bound by the LSD1/CoREST and H3 Thr45 in the nucleosome, which gives 29 amino acids between the two positions. Assuming that the peptide chain length in an extended peptide is 3.6 Å, then 5, 10, 15, 18 and 20 amino acid deletions in the H3 tail would result in distances of approximately 86, 68, 50, 40, and 32 Å, respectively, between the LSD1 catalytic site and the H3 in the nucleosome (Figure 5-8). Therefore, it seems unlikely that LSD1/CoREST would be able to demethylate H3K4 in nucleosomes containing the 18 and 20 amino acid deletions in the H3 tail with the binding mode observed in the mononucleosome substrate model. The 15 amino acid deletions in the H3 tail results in a 50 Å extended chain distance between the Pro16 and Thr45. The measured distance between the LSD1 catalytic site and the H3 substrate in the mononucleosome substrate model was 49 Å, and it is conceivable that some unraveling of the N-terminal H3 peptide in the LSD1 pocket might allow the 15 amino acid H3 tail deletion to enter the catalytic pocket. Recombinant nucleosomes containing the desired truncated histone H3 with H3K4me2 methyl-lysine analog were prepared with 197 bp (26N26) DNA and LSD1/CoREST's catalytic activity on these substrates were tested by Western Blot. At the time, our lab had generated histone H3 with only 5, 10 and 15 amino acid deletion, (histone H3 with 18 and 20 amino acid deletions are currently being prepared). My results show that LSD1/CoREST is able to demethylate H3K4 in 197 bp

nucleosomes with 5, 10 and 15 amino acid deletions, consistent with our hypothesis (Figure 5-9). This also suggests that LSD1/CoREST undergoes a conformation change when binding to the nucleosome with extranucleosomal DNA (Distance of ~63 Å is required for the demethylase activity without conformational change in contrast to ~49 Å). Given that LSD1/CoREST is not able to demethylate H3K4 in the dinucleosome substrates, I did not perform the equivalent experiments in the context of the dinucleosomes. Enzymatic assays on the nucleosomes containing the histone H3 with 18 and 20 amino acid deletions will be critical to evaluate the effect of varying the distance between the catalytic LSD1 site and the H3 substrate.

Distance												
(P16 to T45)	H3	1	6	11	16	21	26	31	36	41	46	
104 Å	WT	ARTK	QTARK	STGGK	APRKQ	LATKA	ARKSA	APATGG	Vĸĸph	RYRPG	TVAL	RE
86 Å	∆ 5aa	ARTK	QTARK	STGGK	APRKQ	!	ARKSA	APATGG	VKKPH	RYRPG	TVAL	RE
68 Å	Δ 10 aa	ARTK	QTARK	STGGK	(APRKQ			-PATGG	VKKPH	IRYRPO	TVAL	RE
50 Å	Δ15 aa	ARTK	QTARK	STGGK	(APRKQ				VKKPF	IRYRPO	<b>STVAL</b>	.RE
40 Å	Δ18 aa	ARTK	QTARK	STGGK	APRKQ				PH	RYRPO	<b>STVAL</b>	.RE
32 Å	∆ 20 aa	ARTK	QTARK	STGGK	APRKQ					-RYRPO	<b><i>STVAL</i></b>	.RE

**Figure 5-8**. Sequences of histone H3 representing the deletions in the H3 tail. The distance represents the approximate distance between H3 Pro16 in LSD1/CoREST from H3 Thr45 in the nucleosome with assuming 3.6 Å per peptide chain (calculated from the PyMOL software). The methyl-lysine analog at position 4 is colored in red, and Pro16 and Thr38 are colored blue.

DNA length	197M	197M	197M	197M	
xH3 tail		5aa del	10aa del	15aa del	
Time (min)	0 15 30	0 15 30	0 15 30	0 15 30	

# Blot: α-H3K4me2

**Figure 5-9**. Demethylase activity of LSD1/CoREST on 197 bp nucleosomes containing H3 tail truncations. The nucleosomes containing the wild-type H3, or 5, 10, and 15 amino acid deletions in the H3 tail were tested using Western Blot with an antibody against H3K4me2. The acivity was measured after 0, 15, and 30 minutes.

#### **5.4 Discussion**

My results suggest that the mononucleosome with extranucleosomal DNA, and not dinucleosomes, is the preferred substrate for LSD1/CoREST complex. The catalytic LSD1 AOD and the N-terminus of CoREST $\Delta$ 1 appear to make contact with extranucleosomal DNA, while the CoREST SANT domain recognizes the nucleosome core particle, possibly through the nucleosome disk face. These findings are consistent with my mononucleosome substrate model of LSD1/CoREST/nucleosome structure and elucidate how LSD1/CoREST interacts with extranucleosomal DNA beyond the nucleosome core particle to positively influence the histone demethylase activity of the complex.

LSD1/CoREST is shown to be functional on mononucleosomes with symmetrically extended DNA of various lengths (Chapter 3 and 5) and also maintains the catalytic activity on the asymmetric 177 bp (31N1) nucleosome (data not shown). However, my enzymatic studies suggest that LSD1/CoREST is not able to demethylate H3K4 in dinucleosomes with 11, 16, 21 and 26 bp linker DNA. My nucleosome binding studies showed optimal binding affinity of LSD1/CoREST to nucleosome with symmetrically extended 181 bp (18N18) extranucleosomal DNA (for nucleosomes labeled on H3 K27C). Although LSD1/CoREST is enzymatically inactive on the dinucleosomes, the complex bound the 306 (N16N) and 322 bp (8N16N8) dinucleosomes with only slightly reduced affinity compared to the 181 bp nucleosome, and with higher affinity than 147 (1N1) or 157 bp (6N6) nucleosomes. If LSD1/CoREST can bind to these dinucleosomes with reasonably high affinity, why is LSD1/CoREST binds to dinucleosomes in a nonproductive way that prevents the enzyme to specifically remove methyl groups from H3K4. The LSD1 protein is able to demethylate not only the Lys4 on histone H3, but also Lys9 on H3 and is also implicated in targeting non-histone substrates such as p53 and DNMT (DNA methyltransferase 1) (Huang *et* 

al., 2007, Wang et al., 2009). The specificity of LSD1 protein to H3K4 in nucleosome substrates is driven by the association with CoREST, which may consequently induce the breathing of the LSD1 pocket and molecular flexibility of LSD1/CoREST complex indicated by molecular dynamics simulation studies (Baron and Vellore, 2012a, Baron and Vellore, 2012b). In addition, the decrease in binding affinity was observed with 185 (20N20), 197 (26N26), and 207 bp (31N31) nucleosomes compared to 181 bp (18N18) nucleosomes, which suggests for either conformational change and/or DNA distortion in DNA beyond 18 bp extension. It is plausible that dinucleosome prevents the conformational change in LSD1/CoREST or inhibits the flexibility of the linker DNA to make appropriate contacts with the LSD1/CoREST complex, consequently precluding the enzyme to specifically remove methyl groups from the H3K4 position (Figure 5-10). On the other hand, in the mononucleosome substrates, the DNA extensions have free ends that are not bound by another nucleosome core particle which allows for DNA breathing and flexibility. However, it is unlikely that such DNA breathing can occur easily in the context of the dinucleosomes. When the LSD1/CoREST complex from the mononucleosome substrate model was modeled into the dinucleosome substrate model, steric clashes were observed between the LSD1 AOD domain and the linker DNA, suggesting that the LSD1/CoREST cannot bind to N16N dinucleosomes using the mononucleosome binding mode (Figure 5-10).

Additionally, both shortening and lengthening the linker length to 301 (N11N), 311 (N21N), and 316 bp (N26N) dinucleosomes decreased the binding affinity for LSD1/CoREST and led to negative cooperative binding, indicated by significantly low Hill coefficients. The dinucleosome substrate model shows that one LSD1/CoREST molecule can bind the nucleosome 2 on the nucleosome disk face via CoREST SANT2 and the nucleosome 1 on the end of the nucleosomal DNA via the LSD1 AOD domain. The 16 bp linker could have been the coincidental

linker length that allowed the LSD1/CoREST to bind to the dinucleosome, and either shortening or lengthening the linker length disables the N-terminus of the CoREST or the LSD1 AOD domain to make appropriate contacts with DNA. In all, both my enzymatic and binding experiments indicate that the mononucleosome with extranucleosomal DNA is the preferred substrate over the dinucleosomes.



**Figure 5-10.** The LSD1/CoREST complex from the mononucleosome substrate model superimposed onto the dinucleosome substrate model. (a) The LSD1/CoREST complex from the mononucleosome substrate model (LSD1, yellow; CoREST, red) was modeled onto the dinucleosome substrate model. The LSD1 and CoREST proteins in the dinucleosome substrate model is shown in blue and green respectively. (b) (c) Steric clashes of the LSD1 AOD domain with the linker and nucleosomal DNA of the nucleosome 1 are observed.

My nucleosome binding studies with the mutant LSD1/CoREST complexes also substantiate my mononucleosome substrate model. The basic stripe in the LSD1 AOD, suggested to make contacts with the DNA extension two turns away from the end of the nucleosome core particle, is involved in interaction with extranucleosomal DNA 16-18 bp away from the end of the nucleosomal DNA. The triple lysines (Lys355, Lys357, and Lys359) are located on the loop region of the LSD1 and it is possible that there is a conformational change in this loop region to make contacts with the extranucleosomal DNA. On the other hand, the LSD1 residue Lys372, located on the end of the helix that is positioned close to the N-terminus of  $CoREST\Delta 1$ , does not affect the affinity of LSD1/CoREST dependent on extranucleosomal DNA. Rather, the LSD1 K372A mutation decreases the binding affinity of LSD1/CoREST by 2-fold across all nucleosomes of different DNA extension length DNA (Table 5-6). The decrease in binding affinity of LSD1/CoREST may be due to the fact that the LSD1 helix 372-395 is involved in making contact with the H3 residues Lys14-Pro16 and stabilizing the association with the CoREST, as the N-terminus loop of CoREST∆1 is close to this LSD1 helix. However, there are other residues in the LSD1 AOD that are suspected to bind to extranucleosomal DNA that still need to be tested.

The mononucleosome substrate model suggest the N-terminus of CoREST $\Delta$ 1, which is rich in basic residues, contacts extranucleosomal DNA one helical turn from the end of the nucleosomal DNA. I show that removing the CoREST residues 286-315 affects the binding affinity of LSD1/CoREST $\Delta$ 4 complex in an extranucleosomal DNA-dependent manner. The Nterminus 286-315 seems to be responsible for making interactions with DNA extension 5-6 bp away from the end of the nucleosomal DNA.

The CoREST SANT2 binds the nucleosome disk face in both di- and mononucleosome substrate models. The triple arginines (Arg425, Arg426, and Arg427) in the CoREST SANT2

play a critical role in binding to the nucleosomes. Mutating these arginine residues to alanines decreased the affinity of LSD1/CoREST towards 147 (1N1), 157 (7N7), and 181 bp (18N18) nucleosomes 5-fold compared to the wild-type LSD1/CoREST. The CoREST residues Lys378 and Arg382 located on the loop of SANT2 domain are suggested to make contacts with nucleosomal DNA in my model. Mutations of these residues to glutamic acid decrease the binding affinity of LSD1/CoREST complex towards the 147 bp (1N1) nucleosome significantly, and reduce the affinity towards the 157 (6N6) and 181 bp (18N18) nucleosomes ~15-fold, suggesting the critical roles for Lys378 and Arg382. The combination of the five lysine and arginine residues in the CoREST SANT2 domain seems to be responsible for recognizing the nucleosome disk face. This is consistent with the previous finding that SANT2 domain is indispensable in stimulating LSD1's demethylase activity towards the nucleosome (Yang *et al.*, 2006).

My biochemical analyses validate my mononucleosome substrate model. In particular, the LSD1 AOD and the N-terminus of CoRESTA1 appears to interact with the extranucleosomal DNA while the CoREST SANT2 domain binds the nucleosome core particle. Further investigations should be performed, including mutating H3 and H4 of the nucleosome disk face contacted by the CoREST SANT2 domain. In summary, the biochemical studies are consistent with my crystallographic mononucleosome substrate model as well as the photocrosslinking results, and provide a roadmap for further crystallization studies of LSD1/CoREST on the nucleosomes.

# **5.5 Acknowledgements**

Song Tan has designed and cloned the plasmid constructs for the dinucleosomal DNA with 11, 16, 21, and 26 bp linkers. Recombinant mono- and dinucleosomes used in the demethylase activity and the HI-FI binding assays were prepared by Mike Doyle, Kevin Thyne, and Bryan Tornabene. The LSD1/CoREST mutants were cloned by Turner Pecen and Victoria Spadafora, undergraduate students in our laboratory. The truncations in the histone H3 tail were cloned and prepared by James Johnstone, an undergraduate student in our laboratory.

### **5.6 References**

- Baron, R. and N. A. Vellore. "LSD1/CoREST Is an Allosteric Nanoscale Clamp Regulated by H3-Histone-Tail Molecular Recognition." *Proc Natl Acad Sci U S A* 109, no. 31 (2012): 12509-14.
- Baron, R. and N. A. Vellore. "LSD1/CoREST Reversible Opening-Closing Dynamics: Discovery of a Nanoscale Clamp for Chromatin and Protein Binding." *Biochemistry* 51, no. 15 (2012): 3151-3.
- Bentley, M. L., J. E. Corn, K. C. Dong, Q. Phung, T. K. Cheung and A. G. Cochran. "Recognition of UbcH5c and the Nucleosome by the Bmi1/Ring1b Ubiquitin Ligase Complex." *EMBO* J 30, no. 16 (2011): 3285-97.
- Forneris, F., C. Binda, A. Adamo, E. Battaglioli and A. Mattevi. "Structural Basis of LSD1-CoREST Selectivity in Histone H3 Recognition." *J Biol Chem* 282, no. 28 (2007): 20070-4.
- Huang, J., R. Sengupta, A. B. Espejo, M. G. Lee, J. A. Dorsey, M. Richter, S. Opravil, R. Shiekhattar, M. T. Bedford, T. Jenuwein and S. L. Berger. "P53 Is Regulated by the Lysine Demethylase LSD1." *Nature* 449, no. 7158 (2007): 105-8.
- Kim, S. A., N. Chatterjee, M. J. Jennings, B. Bartholomew and S. Tan. "Extranucleosomal DNA Enhances the Activity of the LSD1/CoREST Histone Demethylase Complex." *Nucleic Acids Res* 43, no. 10 (2015): 4868-80.
- Olsen, S. K. and C. D. Lima. "Structure of a Ubiquitin E1-E2 Complex: Insights to E1-E2 Thioester Transfer." *Mol Cell* 49, no. 5 (2013): 884-96.
- Pilotto, S., V. Speranzini, M. Tortorici, D. Durand, A. Fish, S. Valente, F. Forneris, A. Mai, T. K. Sixma, P. Vachette and A. Mattevi. "Interplay among Nucleosomal DNA, Histone Tails,

and Corepressor CoREST Underlies LSD1-Mediated H3 Demethylation." *Proc Natl Acad Sci U S A* 112, no. 9 (2015): 2752-7.

- Wang, J., S. Hevi, J. K. Kurash, H. Lei, F. Gay, J. Bajko, H. Su, W. Sun, H. Chang, G. Xu, F. Gaudet, E. Li and T. Chen. "The Lysine Demethylase LSD1 (KDM1) Is Required for Maintenance of Global DNA Methylation." *Nat Genet* 41, no. 1 (2009): 125-9.
- Yang, M., C. B. Gocke, X. Luo, D. Borek, D. R. Tomchick, M. Machius, Z. Otwinowski and H. Yu. "Structural Basis for CoREST-Dependent Demethylation of Nucleosomes by the Human LSD1 Histone Demethylase." *Mol Cell* 23, no. 3 (2006): 377-87.

# Chapter 6

# **Summary and Future Experiments**

## 6.1 Summary

### 6.1.1 Importance of extranucleosomal DNA for LSD1/CoREST

In this study, I established that extranucleosomal DNA enhances the activity and affinity of LSD1/CoREST on the nucleosomes. Although the interaction of LSD1/CoREST with extranucleosomal DNA was first suggested by Nilanjana Chatterjee's photocrosslinking assay, and the enhancement of LSD1/CoREST's activity by extranucleosomal DNA was first demonstrated by Matt Jenning's fluorescence demethylase assay, I was able to determine the dissociation and kinetic constants of LSD1/CoREST on nucleosomes of various DNA lengths. My nucleosome binding studies demonstrated a substantial increase in binding affinity when the nucleosome core particle is symmetrically extended with extranucleosomal DNA up to 181 bp (18N18), consistent with an enhanced catalytic activity on 185 bp nucleosome. In addition, LSD1/CoREST prefers the nucleosomes with extranucleosomal DNA on both sides over the nucleosomes with extranucleosomal DNA on one side for maximal binding.

The arginine anchor is a commonly used nucleosome recognition motif employed by all chromatin proteins (RCC1, Sir3 silencing protein, PRC1 ubiquitylation complex, and SAGA deubiquitylation module) crystallized on the nucleosomes. These chromatin proteins insert an arginine side chain into an acidic patch on the histone dimer surface of the nucleosome. However, my HI-FI binding assay on nucleosome with acidic patch mutation indicated that LSD1/CoREST does not make a critical interaction with the nucleosome acidic patch. Rather, LSD1/CoREST
seems to bind to nucleosomes primarily through interactions with nucleosomal and extranucleosomal DNA. These biochemical analyses provide only a low resolution information for how LSD1/CoREST binds to its nucleosome substrate. Nevertheless, my results set the stage for structural studies with more appropriate nucleosome substrates for LSD1/CoREST.

#### 6.1.2 Low resolution structure of LSD1/CoREST on nucleosomes

I was able to crystallize LSD1/CoREST on nucleosomes with extranucleosomal DNA of various lengths and obtain the low resolution diffraction data from which I solved the molecular replacement solutions. The crystals of LSD1/CoREST on the 161 bp (8N8) nucleosome resulted in the dinucleosome substrate model, suggesting that LSD1/CoREST uses a dinucleosome with 16 bp linker as a substrate. On the other hand, the crystals of LSD1/CoREST on the 189 bp (22N22) nucleosome showed LSD1/CoREST binding to the mononucleosome with extranucleosome JDNA. Even though the two models propose different nucleosome disk face and the N-terminus of CoRESTΔ1 making contacts with extranucleosomal DNA. The dinucleosome substrate model evokes two possible orientations of LSD1/CoREST on the nucleosome 2 (Figure 4-17). In the mononucleosome substrate model, the LSD1 AOD is in close proximity to the extranucleosomal DNA two turns from the end of the nucleosomal DNA, corroborating the latter orientation of the dinucleosome substrate model.

My biochemical studies suggest that the dinucleosome is not the preferred substrate for LSD1/CoREST, as the enzyme is unable to demethylate H3K4 in dinucleosome substrates with 11, 16, 21, and 26 bp linker. Furthermore, negative cooperativity is observed for the binding of

LSD1/CoREST on dinucleosomes with 11, 21, and 26 bp linkers, suggesting that 16 bp was coincidental linker length that allowed LSD1/CoREST molecules to bind without inhibiting each other.

The interactions of the LSD1 AOD and the N-terminus of CoRESTA1 with DNA extensions, and the interaction of the CoREST SANT2 with the nucleosome disk face are observed in both the mono- and dinucleosome substrate models despite the different crystal forms, suggesting that they are not due to crystal artifacts. In addition, these interactions are validated by my biochemical studies. The examination of mutations in LSD1/CoREST using HI-FI binding assay indicates that K378E/R382E and R425A/R426A/R427A mutations significantly decreases LSD1/CoREST's affinity to nucleosomes as suspected in both the dinucleosome and the mononucleosome substrate models. This decrease in binding affinity was paralleled in all 147 (1N1), 157 (6N6), and 181 bp (18N18) nucleosomes. In addition, removing the CoREST residues 286-315 in the N-terminus prevented the 5-10 bp DNA extension on both sides of the nucleosome to enhance the affinity of LSD1/CoREST for nucleosomes. However, the binding affinity of this LSD1/CoREST truncation mutant was still enhanced by the 18 bp DNA extension on both sides of the nucleosome. In addition, the LSD1 AOD, in particular, the residues Lys355, Lys357, and Lys359, have shown to be responsible for making contacts with DNA extension ~16 bp from the end of the nucleosomal DNA.

#### **6.2 Future Experiments**

#### 6.2.1 Crystallization of LSD1/CoREST/nucleosomes for higher a resolution structure

At the time when I started to pursue crystallization of LSD1/CoREST on nucleosomes with DNA length between 185–193 bp, our lab had 185 (20N20), 189 (22N22), and 193 bp (24N24) nucleosomal DNA plasmid constructs available but lacked the 187 (21N21) and 191 bp (23N23) DNA constructs. Successful crystallization of LSD1/CoREST on 189 (22N22) and 193 bp (24N24) nucleosomes and diffraction limit of 5.3 Å resolution from the crystals of LSD1/CoREST on 189 bp (22N22) suggests that 187 and 191 bp nucleosomes are good candidates for crystallization studies. A single base pair on either end of the nucleosome can change diffraction properties as seen with the crystals of nucleosome core particles containing Widom 601 sequence. The 147 bp nucleosomes yielded poor crystals that diffracted to ~7 Å whereas the crystals of the 145 bp nucleosomes diffracted to ~2.5 Å (Vasudevan *et al.*, 2010), indicating how an extra bp prevented the DNA end to DNA end crystal packing in the 147 bp nucleosome crystals.

Shortly after successful crystal growth of LSD1/CoREST on 189 and 193 bp nucleosomes, the cloning of the 187 and 191 bp 601 DNA plasmid constructs was pursued and completed by Turner Pecen, an undergraduate student in our laboratory. Recombinant nucleosomes containing either the wild-type histones or the H3 K4M mutant histone were reconstituted with 187 and 191 bp DNA and I set up crystal trials using these nucleosomes with the human LSD1/CoREST or LSD1/CoREST $\Delta$ 3. I also wondered if removing the histone tail could improve crystallization and diffraction properties. Using the HI-FI nucleosome binding assay, I determined that removing both histone H2A (1-13, 119-130) and H2B (1-23) tails does not affect the binding affinity of LSD1 towards nucleosomes (Supplementary Figure 6-1,

167

Supplementary Table 6-1). Recombinant nucleosomes containing the tailless dimers (H2A 14-118 and H2B 24-122) and 189 bp (22N22) 601 DNA were reconstituted and the crystal trials were set up with LSD1/CoREST on the 189 bp nucleosomes containing tailless dimer. Recent crystallization attempts yielded new crystals with different morphologies in various conditions (Figure 6-1, Table 6-1). It was interesting that the crystals of LSD1/CoREST on the 189 bp nucleosome with tailless dimer grew with different morphologies than the crystals containing LSD1/CoREST on the 189 bp nucleosomes with the wild-type dimer (Figure 4-6c and Figure 6-1b), suggesting for different crystal packing and possibly diffraction properties. I plan on optimizing the crystals that grew in three dimensions, birefringence, and sharp edges, soaking them in various conditions, and collecting diffraction data, hoping that they will produce higher resolution data.



**Figure 6-1.** Recent crystals of LSD1/CoREST/nucleosomes. (a), (b), (c), (d) Initial crystals of LSD1/CoREST on various nucleosome substrates as indicated. (e), (f) Initial crystals of LSD1/CoREST $\Delta$ 3 on various nucleosome substrates as indicated.

Enzyme construct	DNA length	histones	DNA format	Crystal
LSD1/CoREST	187	WT	21+ 145+21	х
LSD1/CoREST	191	WT	23+ 145 + 23	0
LSD1/CoREST	187	H3 K4M	21+ 145+21	0
LSD1/CoREST	189	tailless dimer	22+ 145+22	0
LSD1/CoREST	191	H3 K4M	23+ 145+23	0
LSD1/CoREST∆3	187	WT	21+ 145+21	х
LSD1/CoREST∆3	191	WT	23+ 145 + 23	х
LSD1/CoRESTA3	187	H3 K4M	21+ 145+21	0
LSD1/CoRESTA3	191	H3 K4M	23+ 145+23	0

**Table 6-1**. Components of LSD1/CoREST and nucleosome substrates used in recent crystallization attempts and their crystallization results.

So far, my crystallization studies of LSD1/CoREST/nucleosome were done on nucleosomes with symmetrically extended DNA with an assumption that extranucleosomal DNA on both sides are required for optimal binding of LSD1/CoREST. However, my diffraction data at 5.3 Å resolution from the crystals of LSD1/CoREST on the 189 bp (22N22) nucleosome showed weak electron density for extranucleosomal DNA on one side of the nucleosome. The crystal packing shows that while the 22 bp DNA extension bound by the LSD1/CoREST molecule mediates crystal contacts with the nucleosome in the adjacent symmetric mate, the DNA extension on the other side does not. To improve diffraction properties, nucleosomes with asymmetrically extended DNA can be used grow LSD1/CoREST/nucleosomes. By removing the DNA extension that occupies space in crystals but does not actually mediate helful crystal contacts, the resulting optimized LSD1/CoREST/nucleosome complex might pack more tightly, resulting in higher resolution diffraction.

#### 6.2.2. Further validation of current models of LSD1/CoREST/nucleosome structures

While the most effort should be focused on producing crystals for higher resolution structures, further biochemical studies can be performed to test the mononucleosome substrate model. Mutations on the histone disk face contacted by the SANT2 domain may confirm the interaction between the CoREST SANT2 and the nucleosome disk face. However, we have experienced problems with expressing the histone H4 containing the D24R mutation at this interface. There are also other mutations on LSD1/CoREST that could be tested, including the LSD1 residue Lys699 which is suspected to bind on the DNA extension ~20 bp from the end of the nucleosome core particle. In addition, I would like to test the effect of deleting 18 and 20 amino acids in the H3 tail on LSD1/CoREST's demethylase activity on the nucleosomes. The truncated histone H3 with 18 and 20 amino acid deletions (H3(1-19, 38-135) and H3(1-19, 40-135) respectively) were cloned by Turner Pecen, and the recombinant nucleosomes containing the dimethylated and truncated H3 histones are currently being prepared by Bryan Tornabene and Lauren McCarl in our laboratory.

# 6.2.3 Structural studies of LSD1/CoREST/nucleosome complex using electron cryomicroscopy

Electron cryomicroscopy (Cryo-EM) is increasingly becoming popular in studying the structure of protein complexes. The cryo-EM requires less sample and eliminates the crystallization step, the rate-limiting step for x-ray crystallography (Bai *et al.*, 2015). In addition, the same methods for preparing and purifying the LSD1/CoREST/nucleosome samples can be applied for cryo-EM. Song Tan, my supervisor, was able to collect images on LSD1/CoREST on 181 bp (18N18) nucleosome in Laboratory of Molecular Biology in Cambridge, England, and the data is being processed now. Although we do not expect to solve a high resolution EM structure this time, we are hoping for some structural insights into how LSD1/CoREST binds the mononucleosome, in particular, if the EM structure is consistent with the molecular replacement solution structure

# 6.3 Supplementary Data



**Supplementary Figure 6-1**. Binding curves for LSD1/CoREST interacting with 147 (yellow), 157 (blue), and 181 bp (red) nucleosomes containing tailless dimers at 75 mM NaCl.

DNA length	DNA format	K <sub>d</sub> (nM)	relative to WT	relative to 147
			nucleosome	bp
147	1N1	2029 ± 163	0.64	1.0
157	6N6	846 ± 37	0.98	2.4
181	18N18	429 ± 23	1.09	4.7

**Supplementary Table 6-1**. Dissociation constants for LSD1/CoREST complex binding to 147, 157, and 181 bp nucleosomes containing tailless dimers.

# 6.4 Acknowledgements

Turner Pecen has cloned the plasmid constructs for 187 (21N21), 191 bp (23N23), 165 (1N19), and 166 bp (1N20) nucleosomal DNAs. Turner has also prepared the plasmid constructs of the histone H3 containing the 18 and 20 amino acid deletions. Recombinant nucleosomes and LSD1/CoREST complexes used in crystallization studies were prepared by Bryan Tornabene and Lauren McCarl.

### **6.5 References**

- BAI, X. C., MCMULLAN, G. & SCHERES, S. H. 2015. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci*, 40, 49-57.
- VASUDEVAN, D., CHUA, E. Y. & DAVEY, C. A. 2010. Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J Mol Biol*, 403, 1-10.

# VITA

## Sang-Ah Kim

## Education

Ph.D candidate, Biochemistry, Microbiology, and Molecular Biology
The Pennsylvania State University, University Park, PA
B.S., Biochemistry and Molecular Biology
Clark University, Worcester, MA
B.A., Management – Innovation and Entrepreneurhsip
Clark University, Worcester, MA

### **Publications**

**Kim, S.A.**, Chatterjee, N., Jennings, M.J., Bartholomew, B. and Tan, S. (2015) Extranucleosomal DNA enhances the activity of the LSD1/CoREST histone demethylase complex. *Nucleic acids research*, **43**, 4868-4880.

### **Meeting Abstracts**

**Sang-Ah Kim,** Nilanjana Chatterjee, Matthew J. Jennings, Blaine Bartholomew and Song Tan, "Extranucleosomal DNA enhances the activity of the LSD1/CoREST histone demethylase complex" 34<sup>th</sup> Summer Symposium in Molecular Biology – Chromatin and Epigenetic Regulation of Transcription. Penn State University, 2015 (selected for oral presentation)

### **Honors and Awards**

2011	Homer F. Braddock and Nellie H. and Oscar L. Roberts Fellowships
	(The Pennsylvania State University)
2011	Graduation Honor: summa cum laude
	(ClarkUniversity)
2010	Lise Anne and Leo E. Beavers Fellowship
	(Clark University)