Efforts to Elucidate the Binding Interaction between Lysine-Specific Demethylase 1 (LSD1/KDM1) and CoREST and Their Roles in Breast Cancer

by

Sunhee Hwang

Department of Chemistry
Duke University

Date: _____________________
Approved: _____________________

___________________________
Dewey G. McCafferty, Supervisor

___________________________
Jiyong Hong

___________________________
Pei Zhou

___________________________
Katherine J. Franz

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry in the Graduate School of Duke University

2012
ABSTRACT

Efforts to Elucidate the Binding Interaction between Lysine-Specific Demethylase 1 (LSD1/KDM1) and CoREST and Their Roles in Breast Cancer

by

Sunhee Hwang

Department of Chemistry
Duke University

Date: ______________________

Approved:

___________________________________________
Dewey G. McCafferty, Supervisor

___________________________________________
Jiyong Hong

___________________________________________
Pei Zhou

___________________________________________
Katherine J. Franz

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry in the Graduate School of Duke University

2012
Abstract

Histone modifications play important roles in regulating gene expression in various cellular processes by altering the underlying chromatin structure and thus influencing related pathological conditions. Histone methylation is one such modification that was thought to be static and enzymatically irreversible until the recent discovery of histone demethylases. Lysine specific demethylase 1 (LSD1) is a unique histone demethylase that removes methyl groups on lysine residues and mediates expression of many genes important in cancer progression. LSD1 activity and its substrate specificity are mainly regulated by association with a number of co-regulatory proteins. CoREST is one such important binding protein that endows LSD1 with the ability to associate with and demethylate nucleosomal substrates.

Given the significance of CoREST in directing LSD1 activity, herein we report our efforts to regulate LSD1 enzymatic activity by modulating its binding interaction with CoREST in order to provide a new means to inhibit the demethylation activity of LSD1 and thereby suggest a novel therapeutic intervention for breast cancer where LSD1 is implicated. Towards this end, initial steps have been taken to elucidate the molecular basis of the binding interaction between LSD1 and the functional region of CoREST, denoted as CoREST286-482, by conducting isothermal titration calorimetry (ITC) studies. We found that the proteins tightly interact in a 1:1 stoichiometric ratio with a dissociation binding constant ($K_d$) of $15.9 \pm 2.07$ nM, and that their binding interaction is
characterized by a favorable enthalpic contribution near room temperature with a smaller entropic penalty at pH 7.4. Furthermore, we defined that the linker region of CoREST\textsuperscript{286-482} is a key determinant for the tight binding interaction toward LSD1. Accordingly, the peptide corresponding to the linker region of CoREST\textsuperscript{286-482} (designated as the Linker peptide) was used as a potent competitive modulator of the binding interaction between LSD1 and CoREST\textsuperscript{286-482} with our expectation that the Linker peptide will compete with CoREST\textsuperscript{286-482} for binding to LSD1 and thus sequester LSD1 from nucleosomal contexts to inhibit its demethylation activity.

Indeed, the Linker peptide has shown successful inhibitory activity against the binding of LSD1 to CoREST\textsuperscript{286-482} at nanomolar concentration. We evaluated the potential of the Linker peptide to inhibit LSD1 demethylation activity in cellular models of estrogen receptor α (ERα)-positive breast cancer cell line (MCF7) where LSD1 is known to co-localize with ERα and affect ERα-transcription activities. We were able to observe that the disruption of the binding interaction between LSD1 and CoREST by the Linker peptide affects not only the demethylation activity of LSD1 on both H3K4 and H3K9 but also the ERα-recruitment to promoters of target genes and processes required for proliferation of breast cancer cells. Thus it is believed that LSD1 demethylation activity through its physical and functional interaction with CoREST is essential in the development of ERα-positive breast cancer. Furthermore, we have observed that not LSD1 but CoREST physically interacts with ERα, which suggests a potential role of
CoREST as a bridge connecting LSD1 to ERα. At the moment although the precise mechanisms underlying the implication of LSD1 through CoREST interaction in ERα-positive breast cancer still remain elusive, our study has verified that modulating LSD1 enzymatic activity by interrupting a critical protein-protein interaction offers a distinctively new avenue of inhibition of LSD1 activity, which has anti-cancer potential.
To all my family who supported me every step of the way.
## Contents

Abstract ........................................................................................................ iv

List of Tables .................................................................................................. xiii

List of Figures ................................................................................................ xv

List of Schemes ............................................................................................. xxvi

List of Abbreviations ..................................................................................... xxviii

Acknowledgements ........................................................................................ xxxiv

1. Introduction ................................................................................................ 1

   1.1 Epigenetics ............................................................................................. 1

       1.1.1 Chromatin, Nucleosomes, and Histones ........................................ 1

       1.1.2 Post-translational Modifications of Histones .............................. 7

1.2 Histone Methylation and Demethylation .............................................. 15

1.3 LSD1 ........................................................................................................ 24

       1.3.1 LSD1 Mechanism and Structure Overview .................................. 24

       1.3.2 The Role of LSD1 in Gene Regulation ......................................... 30

       1.3.3 Associated Factors of LSD1 ....................................................... 32

           1.3.3.1 Overview of CoREST ......................................................... 36

           1.3.3.2 The Importance of CoREST in LSD1 Activity ..................... 39

       1.3.4 Strategies to Inhibit the Demethylation Activity of LSD1 .......... 41

1.4 Thesis Objectives ..................................................................................... 45

2. Expression and Purification of LSD1 and CoREST ................................. 48

   2.1 Background ......................................................................................... 48
2.2 Expression and Purification of LSD1 ................................................................. 48
2.3 Expression and Purification of CoREST$^{286-482}$ ............................................. 52
  2.3.1 Gene Cloning for CoREST$^{286-482}$ ............................................................ 52
  2.3.2 Expression and Purification of CoREST$^{286-482}$ ........................................ 54
2.4 Enzymatic Assay of LSD1 in the Presence of CoREST$^{286-482}$ .................... 59
2.5 Discussion and Future Directions ................................................................. 63
2.6 Experimental Section .................................................................................. 64

3. Characterization of the Binding Interaction between LSD1 and CoREST$^{286-482}$ .... 67
  3.1 Background ................................................................................................. 67
    3.1.1 Protein-Protein Interactions ................................................................. 68
    3.1.2 Overview of Isothermal Titration Calorimetry and Thermodynamics ......... 69
  3.2 Calorimetric Titration of CoREST$^{286-482}$ against LSD1 ............................. 77
    3.2.1 Analysis of the Thermodynamic Parameters for the Interaction between LSD1 and CoREST$^{286-482}$ ................................................................. 78
    3.2.2 Surface Plasmon Resonance (SPR) Measurement .................................. 83
  3.3 Calorimetric Titration of CoREST$^{286-482}$ against LSD1 in Different Buffers .... 85
  3.4 Calorimetric Titration of CoREST$^{286-482}$ against LSD1 in Different Temperatures .. 88
    3.4.1 Calculation of the Solvent Accessible Surface Area (SASA) ..................... 93
  3.5 Calorimetric Titration of CoREST$^{286-482}$ against LSD1 as a Function of pH. .... 95
  3.6 Effect of the Disulfide Bond Formation upon the Binding Interaction ............. 97
  3.7 Discussion and Future Directions ................................................................ 98
  3.8 Experimental Section ................................................................................ 105
4. Determination of the Key Residues of the Binding Interface between LSD1 and CoREST<sup>286-482</sup> .......................................................... 109

4.1 Background ............................................................................................................. 109

4.1.1 Targeting Protein-Protein Interactions .......................................................... 109

4.2 Calorimetric Titration of CoREST Linker<sup>293-380</sup> and CoREST SANT2<sup>381-450</sup> against LSD1 .......................................................... 112

4.3 Site-Directed Mutagenesis Study ............................................................................. 116

4.3.1 Calorimetric Titration of CoREST<sup>286-482</sup> Mutants against LSD1 .................. 116

4.4 Peptide Fragment Analysis ....................................................................................... 121

4.5 Spot Synthesis .......................................................................................................... 126

4.6 Discussion and Future Directions ............................................................................ 133

4.7 Experimental Section ............................................................................................... 136

5. Inhibition of LSD1 Activity by Disrupting the Binding Interaction with CoREST in ERα-Positive Breast Cancer................................................................. 146

5.1 Background ............................................................................................................. 146

5.1.1 Breast Cancer ........................................................................................................ 147

5.1.2 Estrogen and Estrogen Receptor (ER) .............................................................. 153

5.1.3 Current Efforts to Modulate ERα Activity ......................................................... 162

5.1.4 Epigenetics in Breast Cancer ............................................................................. 168

5.2 In vitro Inhibition of the Binding Interaction between LSD1 and CoREST<sup>286-482</sup> with the Linker peptide ........................................................................ 172

5.3 In vivo Inhibition of the Binding Interaction between LSD1 and CoREST in ERα-Positive Breast Cancer Cell ........................................................................ 174

5.3.1 Overexpression of CoREST<sup>293-380</sup> in ERα-Positive Breast Cancer Cells .......... 174
5.3.2 Inhibition of the Binding Interaction with the Fusion Linker protein with HIV-TAT domain ................................................................. 174

5.3.2.1 Preparation of the Linker peptide with HIV-TAT domain .................. 174

5.3.2.2 In vitro Inhibition of the Binding Interaction between LSD1 and CoREST\(^{286-482}\) with the Linker –TAT peptide ...................................................... 179

5.3.3 Examination of the Global Histone Methylation Levels ....................... 181

5.3.4 Examination of the Stability of LSD1 ............................................ 184

5.3.5 Examination of the Effects of the Disruption of the Binding Interaction in ER\(\alpha\)-Positive Breast Cancer .................................................... 187

5.3.5.1 Decreased Recruitment of LSD1/CoREST/ER to ER\(\alpha\)-Dependent Target Gene Promoter ................................................................. 187

5.3.5.2 Decreased Transcription Levels on ER\(\alpha\)-Dependent Genes .............. 192

5.3.5.3 Decreased Proliferation of ER\(\alpha\)-Positive Breast Cancer cells .......... 195

5.4 In vitro Pull-Down Assay to Investigate Interaction between LSD1 and ER\(\alpha\) .... 199

5.5 Discussion and Future Directions ..................................................... 202

5.6 Experimental Section ...................................................................... 212


6.1 Background ...................................................................................... 221

6.2 Generation of Stable Human Doxycycline-Inducible Breast Cancer MCF7 cells Overexpressing the Linker Gene ...................................................... 227

6.3 Conclusion and Future Directions ..................................................... 230

6.4 Experimental Section ...................................................................... 233

Appendix A. Expression and Purification of CoREST\(^{53-482}\) ................................ 238

A.1 Background ..................................................................................... 238
**List of Tables**

Table 1: Different classes of modifications identified on histones (Adapted from [26]). ..13

Table 2: Enzymes associated with histone modifications (Adapted from [26, 27, 28]). ....14

Table 3: Identified histone lysine demethylases and their roles in human diseases (Adapted from [47]). ............................................................23

Table 4: Kinetic parameters of the catalytic activity of LSD1 with the H3K4diMe peptide substrate (21 amino acids). ..........................62

Table 5: Thermodynamic parameters for binding of LSD1 to CoREST$^{286-482}$ at 25 °C. b ......81

Table 6: Thermodynamic parameters for binding of LSD1 to CoREST$^{286-482}$ in various buffers at 25 °C. c .................................................................87

Table 7: Thermodynamic parameters for binding of LSD1 to CoREST$^{286-482}$ at different temperatures. d .................................................................90

Table 8: Thermodynamic parameters for binding of LSD1 to CoREST$^{286-482}$ at different temperatures with the addition of 10 mM NaCl in buffer. e .........................................................91

Table 9: Thermodynamic parameters for binding of LSD1 to CoREST$^{286-482}$ at different temperatures with the addition of 100 mM NaCl in buffer. f .................................................................91

Table 10: Change in solvent accessible surface area ($\Delta$ASA) upon binding between LSD1 and CoREST$^{286-482}$. g .................................................................94

Table 11: Thermodynamic parameters for binding of LSD1 to CoREST$^{286-482}$ at different pH (25 °C). h .............................................................................96

Table 12: Thermodynamic parameters for binding of LSD1 to CoREST$^{286-482}$ without a reducing reagent (25 °C). i ............................................................................97

Table 13: Thermodynamic parameters for binding of LSD1 to CoREST Linker$^{293-382}$ at 25 °C. j ............................................................................115

Table 14: Thermodynamic parameters for binding of LSD1 to CoREST$^{286-482}$ mutants at 25 °C. k ............................................................................120
Table 15: 20-mer peptides corresponding to CoREST Linker<sup>293-380</sup> sequences. ..................122

Table 16: Sequences of peptides synthesized on the membrane. Reactivity was determined by visual observation: +++ strong, ++ average, + weak reactivity. ..........128

Table 17: Steroid hormones and their respective receptors. The biological effects of estrogen are mediated by binding to estrogen receptors. ..................................................154

Table 18: Genes probed by RT-PCR. All information was obtained from NCBI Gene database. ...........................................................................................................194

Table 19: Expression levels of CoREST<sup>53-482</sup> under various culture conditions. The coding sequence of CoREST<sup>53-482</sup> was cloned into pET151/D-TOPO. .................................243

Table 20: Expression levels of CoREST<sup>53-482</sup> under various culture conditions. The coding sequence of CoREST<sup>53-482</sup> was cloned into pDEST15. .........................................................244
List of Figures

Figure 1: Nucleosome structure and histone octamer organization. Nucleosome consists of 147 bp (base pairs) of DNA (yellow) wrapped around the histone octamer being made up of two copies of core histone proteins - H2A (purple), H2B (cyan), H3 (blue), and H4 (pink). Figure was generated using PyMoL (PDB file: 1EQZ). .................................................................3

Figure 2: The major structures in DNA compaction. Chromatin that makes up chromosomes is composed of nucleosomes appearing as beads on a string. Each nucleosome consists of DNA wrapped around histone proteins, and the tails of these histones protrude from nucleosomes (adapted from [6]). ........................................................................4

Figure 3: Schematic representation of structural differences between (a) heterochromatin and (b) euchromatin. ..................................................................................................................6

Figure 4: Post-translational modifications of the core histones. Histone tails are methylated at lysines and arginines (red), acetylated at lysines, phosphorylated at serines or threonines (blue), and ubiquitylated at lysines (green). ..............................................................................8

Figure 5: Histone code. The modifications probably act as specific receptors for protein modules that mediate downstream events leading to gene transcription. Abbreviations: HAT: histone acetyltransferase; HDAC: histone deacetylase; HMT: histone methyltransferase; Ac: acetyl group; Me: methyl group; P: phosphate group (Adapted from [11]). ..................................................................................................10

Figure 6: Methylation states of (a) arginine and (b) lysine. Methyl groups are indicated in red. ..........................................................................................................................18

Figure 7: (a) Domain and (b) crystal structure of LSD1 with an N-terminal deletion of first 150 residues (NΔ150). LSD1 tower domain (residues 418-522) is colored in blue, amino oxidase domain (N-terminal residues 271-417 and C-terminal residues 523-833) in green, and SWIRM domain (residues 172-270) in purple, respectively. (c) Surface structure of AOD (green) and SWIRM domain (purple). The 16-mer peptide substrate (red) is located at the center of the AOD. The C-terminal of the histone H3 tail is expected to locate at the binding interface between the AOD and SWIRM domain indicated by a pink arrow. All structural figures were generated using PyMoL (PDB file: 2IW5). ........................................................................................................29

Figure 8: Key histone demethylase complexes. LSD1 plays an important role in transcriptional regulation by associating with a variety of proteins. (a) LSD1-CoREST
complex. (b) LSD1-NuRD complex. (c) LSD1-nuclear hormone receptor (AR/ER) complex. Abbreviations: AR/ER: androgen receptor/estrogen receptor; BRAF35: BRCA2-associated factor 35; BRCA2: breast cancer 2 susceptibility protein; BHC80: BRAF and histone deacetylase complex 80; HDAC1/2: histone deacetylase 1/2; MBD: methyl-CpG-binding domain protein 2; Mi2: myositis autoantigen 2; MTA: metastasis associated protein; RbAp: retinoblastoma-associated protein (Adapted from [54]).

Figure 9: (a) Domain structure of CoREST. The essential domains of CoREST, the linker region and SANT2 domain, for LSD1 demethylation activity are highlighted in red and pink, respectively. (b) Crystal structure of C-terminal segment of CoREST including the linker region (red) and SANT2 domain (pink). (c) LSD1 in complex with the C-terminal segment of CoREST. All structural figures were generated using PyMoL (PDB file: 2IW5).

Figure 10: Mechanism-based inactivators of LSD1.

Figure 11: Mechanism-based inhibitors of LSD1 in complex with FAD. The structures show the covalent adducts involving (a) FAD and tranylcypromine [74] and (b) FAD and propargyl-Lys derivatized peptide inhibitor [78].

Figure 12: Representative AKTA FPLC profiles and pictures of 10 % SDS-PAGE gels showing LSD1 purification on (a) chelating sepharose Ni (II) affinity chromatography and (b) Hiprep 26/60 Sephacryl S200 gel filtration chromatography. A blue line represents UV absorbance at 280 nm as a function of retention volume (mL), and a green line represents a buffer gradient. Red arrows indicate the peaks where LSD1 eluted from the column. Details are given in Experimental Sections.

Figure 13: Representative AKTA FPLC profiles and pictures of 15 % SDS-PAGE gel showing LSD1 purification on (a) chelating sepharose Ni (II) affinity chromatography, (b) Hiprep 26/60 Sephacryl S200 gel filtration chromatography, and (c) Q-sepharose anion exchange chromatography. A blue line represents UV absorbance at 280 nm as a function of retention volume (mL), and a green line represents a buffer gradient. Red arrows indicate the peaks where LSD1 was eluted off from the column. Further purification of LSD1 with an additional ion exchange chromatography improved the purity of the protein. Details are given in Experimental Sections.

Figure 14: (a) Picture of 1 % agarose gel of PCR product of the extraction of the gene encoding CoREST^{286-482}. PCR was performed at different annealing temperatures. (b) A vector map of pET28b indicates the portion of CoREST^{286-482} gene.
Figure 15: Expression of CoREST\textsuperscript{286-482} in (a) BL21 (DE3) \textit{E. coli} strain, (b) BL21 Star (DE3) \textit{E. coli} strain, and (c) C43 (DE3) \textit{E. coli} strain. Pre: before IPTG induction; 1h, 2h, 4h: time points after IPTG induction; S: soluble fraction; IS: insoluble fraction. Red arrows indicate the location of CoREST\textsuperscript{286-482} on each gel. ........................................................................................................56

Figure 16: Representative AKTA FPLC profiles and pictures of 12 % SDS-PAGE gel showing CoREST\textsuperscript{286-482} purification on (a) chelating sepharose Ni (II) affinity chromatography and (b) CM-sepharose cation exchange chromatography. A blue line represents UV absorbance at 280 nm as a function of retention volume (mL), and a green line represents a buffer gradient. Red arrows indicate the peaks where CoREST\textsuperscript{286-482} was eluted off from the column........................................................................................................57

Figure 17: A graphic representation of amino acid sequence of CoREST\textsuperscript{286-482} identified by tandem mass spectrometry. The sequences in the blue box represent the ones corresponding to CoREST\textsuperscript{286-482}. The yellow regions represent the sequences identified and matched with original sequence of CoREST, and the green color represents the oxidized methionines ........................................................................................................58

Figure 18: Sequence of the first N-terminal 30 amino acids of histone 3 (H3). The numbers above residues indicate the length of peptides........................................................................................................60

Figure 19: Initial velocity curves of the catalytic activity of LSD1 using the first 21 amino acid residues of H3 with dimethylated K4 as a substrate. The assay was performed in 50 mM Tris buffer, pH 7.85 at 25 °C, and the final concentration of both proteins was 0.5 \textmu M. Data were fitted to the Michaelis-Menten equation. The black curve represents the activity of LSD1, whereas the green curve represents the activity of LSD1 in the presence of CoREST\textsuperscript{286-482}. ........................................................................................................62

Figure 20: Schematic representation of isothermal titration calorimetry (ITC) instrument. The sample cell is filled with the protein (blue) and the injection syringe, which stirs the solution to assure proper mixing, is filled with the partner protein (pink). At specific time intervals, as the protein in the syringe is injected into the cell, heat is generated or absorbed. The feedback system will either increase or decrease power to the sample cell to maintain equal temperature with the reference cell. The heat per unit time (\textmu cal/sec) supplied to the sample cell is monitored in the ITC experiment (Adapted from [90]). ....72

Figure 21: A representative calorimetric titration of 3 \textmu M of LSD1 with 30 \textmu M of CoREST\textsuperscript{286-482} in 50 mM sodium phosphate and 1 mM DTT, pH 7.4. Thermogram (top panel) represents the heat released after each injection of CoREST\textsuperscript{286-482} into LSD1 solution. Binding isotherm (bottom panel) shows the integrated peak area plotted as a
function of molar ratio (CoREST\textsuperscript{286-482}/LSD1). The following parameters were observed: $n = 1.14 \pm 0.00$, $\Delta H = -21.3 \pm 0.19$ kcal/mol, $K_a = 6.29e7 \pm 8.18e6$ M\textsuperscript{-1}. The red line indicates the best fit of the ITC data to a one site binding model.

Figure 22: (a) Surface representation of the complex between LSD1 and CoREST. Two main domains responsible for the binding interaction are shown: LSD1 tower domain (residues 418-522, blue) and CoREST linker region (residues 293-380, red). The dark blue and dark red surfaces represent the residues involved in the binding interaction. (b) Possible residues involved in the hydrogen bond and salt bridges are represented by dashed lines. (c) Nonpolar residues are populated at the binding interface. Nonpolar residues of each protein are represented as red and blue surface diagrams, respectively. All structural figures were generated using PyMoL (PDB file: 2IW5).

Figure 23: Surface plasmon resonance (SPR) measurement curves obtained during and after injection of LSD1 on chip surfaces with immobilized CoREST\textsuperscript{286-482}. Sensograms show the association of LSD1 to and dissociation from CoREST\textsuperscript{286-482}. The association rate constant ($k_{on}$, $5.63 \times 10^4 \pm 1.93 \times 10^3$ M\textsuperscript{-1}s\textsuperscript{-1}) and dissociation rate constant ($k_{off}$, $7.78 \times 10^5 \pm 1.27 \times 10^5$ s\textsuperscript{-1}) were used to calculate a dissociation binding constant ($K_d$) of 1.4 \pm 6.61 nM.

Figure 24: Plot of the calorimetric enthalpy ($\Delta H_{cal}$) obtained from binding of LSD1 to CoREST\textsuperscript{286-482} as a function of the enthalpy of buffer ionization ($\Delta H_{ion}$) at pH 7.4. The slope of the plot indicates the number of protons transferred between the complex and the buffer upon binding, and the y-intercept indicates the intrinsic $\Delta H_{cal}$ without a protonation effect. At pH 7.4, 0.83 \pm 0.09 protons are absorbed to the complex from the buffer.

Figure 25: Temperature dependence of the binding enthalpy change ($\Delta H_{cal}$) for the interaction between LSD1 and CoREST\textsuperscript{286-482} with different NaCl concentrations in buffers at pH 7.4. Blue diamond marks (●) indicate the enthalpy change with no NaCl in buffer, pink square (■) with 10 mM NaCl, and green circle (●) with 100 mM NaCl. With addition of buffer additive, NaCl, larger binding heats ($\Delta H_{cal}$) were observed, but the $\Delta C_p$ values were not changed significantly. The $\Delta C_p$ values with the addition of 10 mM NaCl and 100 mM NaCl were determined to be -1.04 \pm 0.04 kcal/mol·K and -0.84 \pm 0.1 kcal/mol·K, respectively.
Figure 27: A heptad repeat model of a coiled-coil interaction. Typically the coiled-coil is defined by the heptad amino acid repeat sequence \((abcdefg)_n\). A 3.5 residue periodicity along the coil axis aligns \(a\) and \(d\) position hydrophobic residues, stabilizing the interaction interface, while residues at the \(e\) and \(g\) positions are typically oppositely charged and flank the hydrophobic core, forming salt bridges between helical chains.

Figure 28: (a) Trimeric coiled-coil interaction between LSD1 tower region (blue) and CoREST linker region (red) (side view and bottom view). (b) The side chains at \(a\) and \(d\) position of a heptad repeat model are represented as sticks. The binding interface between two proteins exhibits a unique trimeric coiled-coil structure, which has been less preceded. All structural figures were generated using PyMoL (PDB file: 2IW5).

Figure 29: Expression and purification of (a) CoREST Linker\(^{293-380}\) and (b) CoREST SANT2\(^{381-450}\). Both proteins were overexpressed in BL21 Star (DE3) \(E.\) coli stain. The first and fourth panels show the expression of each protein at 37 °C, the second and fifth show the expression of each protein at 23 °C. The third and sixth panels show purified proteins indicated by arrows.

Figure 30: (a) A calorimetric titration of LSD1 with CoREST Linker\(^{293-380}\) and (b) a titration of LSD1 with CoREST SANT2\(^{381-450}\) in 50 mM sodium phosphate and 1 mM DTT, pH 7.4. The binding interaction between LSD1 and CoREST Linker\(^{293-380}\) yielded the qualitatively similar thermodynamic parameters as compared to the parameters of the binding interaction between LSD1 and CoREST\(^{286-482}\), while no binding isotherm curve was observed between LSD1 and CoREST SANT2\(^{381-450}\).

Figure 31: (a) A side view and (b) bottom view of LSD1 tower domain (blue) in complex with CoREST\(^{286-482}\) mutants. The mutated residues of CoREST\(^{286-482}\) are represented as sticks in orange color. All structural figures were generated using PyMoL (PDB file: 2IW5).

Figure 32: Five 20-mer peptides corresponding to the sequence of CoREST Linker\(^{293-380}\) are synthesized, which are represented in different colors. Cyan represents CoR1, yellow CoR2, pink CoR3, green CoR4, and brown CoR5 as described in Table 15. Structural figure was generated using PyMoL (PDB file: 2IW5).

Figure 33: Western blot probing LSD1 after binding to GST-CoREST\(^{286-482}\) with potential peptide inhibitors listed in Table 15. Mostly those peptides showed weak bindings to LSD1 with poor inhibition activity at excessive concentration of 500 \(\mu\)M.
Figure 34: (a) Reactivity of 12-mer peptides corresponding to CoREST Linker293-380 with LSD1 and (b) crystal structure reflecting the reactivity of the peptides on the spot membrane. Sequences colored in red and light red represent relatively strong and weak reactivities toward LSD1 binding. The regions labeled as a, b, and c exhibited relatively strong reactivities in binding to LSD1. Structural figure was generated using PyMoL (PDB file: 2IW5). .......................................................... 131

Figure 35: A putative model of the potential crystal structure of LSD1 (blue) in complex with the linker region of CoREST (orange). The entire linker region of CoREST may be required for complete inhibition of the binding interaction between LSD1 and CoREST286-482. Structural figure was generated using PyMoL (PDB file: 2IW5). .......... 132

Figure 36: Estimated cancer cases and deaths by sex in 2011 (United States); (a) both sexes, (b) male, and (c) female. Blue bar indicates total estimated cancer cases, while red bar indicates resultant deaths. Breast cancer is the highest cause of deaths in women [135, 136].................................................................................. 149

Figure 37: Major steroid hormones. ................................................................................. 152

Figure 38: Functional domains of human estrogen receptors (ERs). Both ERα and ERβ share a highly conserved DNA- and C-terminal ligand binding domain (LBD) except the AF-1 domain. No apparent AF-1 domain is present in ERβ. Numbers in boxes indicate the numbers of residues in each domain [147]. ................................................................. 157

Figure 39: Crystal structure of ERα ligand binding domain (LBD). ERα LBD is arranged in an antiparallel α-helical sandwich fold. Particularly, the helix colored in pink functions as a docking module to recruit various co-regulators. Estrogen is represented as sticks in green. Structural figures were generated using PyMoL (PDB file: 1QKU)...158

Figure 40: A classical model of ERα action. Estrogen (E2) travels through blood stream and passes through the membrane of target cells. When estrogen molecule enters the cell, it binds to its receptors (ERα), which induces conformational changes within ERα and thus promotes dimerization. The estrogen-ERα complex then binds to specific DNA sites, called estrogen-responsive elements (EREs), near genes that are transcriptionally regulated by estrogens. This mechanism activates gene transcription and then synthesis of specific proteins that can influence cell behavior in different ways.................................160

Figure 41: Schematic representation of ERα-related pathways targeted by small molecules and peptides used to block ERα action and breast cancer cell growth. Currently small molecules mainly target estrogen synthesis and competition with
estrogen for binding in ERα. Alternative approaches include inhibitions of ERα binding to ERE DNA sites and coregulators.

Figure 42: Structures of ERα antagonists. Exemestane, letrozole, and anastrozole belong to aromatase inhibitors [152]; Tamoxifen and raloxifene compete with estrogen for binding to ERα [153, 154, 155]; TPBM (theophylline, 8-[(benzylthio)methyl]-(7CI, 8CI)) and DIBA (2,2,’-dithiobisbenzamide-1) inhibit ERα binding to ERE DNA sites [159, 160].

Figure 43: A heat map illustrating expression levels of FAD-dependent amine oxidases in breast cancer cell lines. Red and blue colors indicate high and low expression levels, respectively. Among nine FAD-dependent amine oxidases, LSD1 is the most-highly expressed, suggesting that LSD1 can be a useful target in the treatment of breast cancers [133].

Figure 44: Western blot probing for LSD1 binding to GST-CoREST286-482. The Linker peptide indeed disrupts the binding interaction between LSD1 and CoREST286-482 in a dose-dependent manner. The eluted LSD1 was immunoblotted with anti-LSD1.

Figure 45: Preparation of the Linker-TAT peptide (18 kDa). (a) The gene encoding the Linker was cloned into pTAT-HA expression vector. (b) Sequence of the Linker-TAT peptide. His6 tag is colored in green, TAT domain in purple, HA tag in blue, and Linker in orange. (c) The Linker-TAT peptide (indicated by a pink arrow) was purified via Ni (II) affinity chromatography. However the undesired proteins indicated by a blue arrow coexisted with the Linker-TAT peptide through several purifications steps such as size exclusion chromatography and ion exchange chromatography.

Figure 46: Selective precipitation of the Linker-TAT peptide (indicated by a pink arrow) with addition of ammonium sulfate ((NH4)2SO4). (a) In a preliminary test with the cell lysate, the ammonium sulfate concentration was increased stepwise, and each protein present in the lysate started to precipitate at different concentration of ammonium sulfate. Particularly, the Linker-TAT peptide precipitated at 30 % ammonium sulfate concentration and was separated from the undesired proteins (indicated by a blue arrow). We applied the ammonium sulfate precipitation method to polish the Linker-TAT peptide after the purification via Ni (II)-affinity chromatography. (b) The disulfide dimerization formed after the salt addition was reduced to the monomer in addition of 1 mM DTT, and finally the Linker-TAT peptide was purified to near homogeneity.

Figure 47: Western blot probing for LSD1 binding to GST-CoREST286-482. (a) The Linker peptide with an additional TAT domain still inhibits the binding interaction between
LSD1 and CoREST in a dose-dependent manner almost at the same concentration of LSD1 present. (b) GFP-TAT peptide was used as a control did not affect the binding interaction at all as expected.

Figure 48: Global dimethylation levels of H3K4 and H3K9 in ERα-positive MCF7 breast cancer cells after treatment with various concentrations of the Linker-TAT peptide (0-200 nM). GFP-TAT peptide was used as a control. (a) GFP and (b) Linker peptide were transduced into MCF7 cells in a dose-dependent manner by aids of TAT domain. (c) The methylation levels on H3K4 and H3K9 were not changed with treatment of MCF7 cells with GFP-TAT peptide, but (d) significantly increased methylation levels on both marks were observed with treatment with Linker-TAT peptides. This indicates that disruption of the binding interaction affects the global levels of H3K4 and H3K9 dimethylation and that the Linker-TAT peptide is most likely specific to modulate the binding interaction between LSD1 and CoREST. H3 was used as a loading control. Abbreviation: NT; non-treated, BF; PBS buffer-treated.

Figure 49: The binding of LSD1 with CoREST through helical interactions is required for LSD1 stability in vivo. (a) Immunostaining of siRNA CoREST-transfected cells with DAPI, which marks the transfected cells (left column), and LSD1 antibody (right column). The knockdown of CoREST led to a reduction of LSD1 expression, but when the Linker-TAT peptide was added, the LSD1 expression was restored. (b) Fluorescence intensity of LSD1 antibody analyzed for LSD1 levels after addition of the Linker-TAT peptide. 63 % of LSD1 was restored as compared to the LSD1 levels in siRNA-CoREST treated cells. (c) Western blotting analysis of the knockdown of CoREST. The knockdown of CoREST decreased LSD1 expression, but the addition of the Linker TAT peptide restored the LSD1 expression. Mock is untreated cell and siMED represents a scrambled siRNA. GAPDH was included as a loading control.

Figure 50: ChIP/qPCR recruitment analysis of ERα, CoREST, and LSD1 on the proximal and distal binding sites of the promoters of ERα target genes, (a) pS2 and (b) PR. CoREST and LSD1 associate with ERα-promoter targets in MCF7 cells. Data is presented as ± SEM for triplicate wells. IgG was used as a negative control.

Figure 51: Disruption of the binding interaction between LSD1 and CoREST with the Linker-TAT peptide leads to the decreased recruitments of ERα, CoREST, and LSD1 on the proximal binding site of the promoters of ERα target genes, (a) pS2 and (b) PR. Data is presented as ± SEM for triplicate wells. IgG was used as a negative control.

Figure 52: Disruption of the binding interaction between LSD1 and CoREST with GFP-TAT peptide (control peptide) does not affect the recruitments of ERα, CoREST, and
LSD1 on the proximal binding site of the promoters of ERα target genes, (a) pS2 and (b) PR. Data is presented as ± SEM for triplicate wells. IgG was used as a negative control.

Figure 53: RT-qPCR analysis after the treatment of ERα-positive MCF7 cells with the Linker-TAT peptides at various concentrations. Some genes were affected by the disruption of the binding interaction between LSD1 and CoREST. Data is presented as ± SEM for triplicate wells.

Figure 54: Proliferation of ERα-positive MCF7 cells after repeated treatment with indicated concentration of the Linker-TAT peptide over the course of 8 days. Disruption of the binding interaction leads to the reduced rate of cellular proliferation in both dose- and time-dependent manners.

Figure 55: Examination of the stability of the Linker-TAT peptide. The Linker-TAT peptide was stable for 8 days without any significant degradation in MCF7 cells. GAPDH was used as a loading control.

Figure 56: Viability of ERα-positive MCF7 cells after treatment for 24 h with indicated concentrations of the Linker-TAT peptide. The various concentrations of the Linker-TAT showed little effects on the cell viability, which indicates that cytostatic capability of the Linker-TAT peptide is not ascribed to its cytotoxicity.

Figure 57: Western blot analyses showing (a) immunoprecipitation of LSD1 and coimmunoprecipitation of CoREST from ERα-positive MCF7 cell extracts by using a LSD1 antibody and (b) coimmunoprecipitation of LSD1 by ERα antibody from MCF7 cell extracts [61]. This suggests the potential coexistence of LSD1/CoREST/ERα in ERα-positive breast cancer cells. (c) A 12 % SDS-PAGE gel showing normalized GST fusion proteins used for a pull-down assay (performed by Dr. Julie A. Pollock in our group). Yellow bands indicate the location of the indicated proteins. (d) Radiometric blot showing the binding interaction between CoREST and ERα.

Figure 58: Two potential ERα-binding motifs of CoREST<sub>286-482</sub> are represented in green color. Figure was generated using PyMol (PDB file: 2IW5).

Figure 59: (a) Global dimethylation levels of H3K4 and H3K9 in ERα-negative MDA-MB-231 breast cancer cells after treatment with Linker-TAT peptide in a dose-dependent manner. The methylation levels on H3K4 and H3K9 were significantly changed with the addition of Linker-TAT peptide to cells. (b) ERα-negative cell proliferation after repeated treatment with indicated concentration of the Linker-TAT peptide over the course of 8 days. Disruption of the binding interaction leads to the reduced rate of
cellular proliferation in both dose- and time-dependent manners. Abbreviation: NT; non-treated, BF; PBS buffer treated.

Figure 60: A tetracycline (Tet)-inducible system. In the Tet-Off system, tTA (tetracycline-controlled transactivator protein) binds to the tetracycline-responsive promoter elements (TRE) and activates target gene expression. In the presence of tetracycline (Tc) or doxycycline (Dox), tTA is unbound and gene expression is turned off. In the Tet-On system, reverse tetracycline-controlled transactivator protein (rtTA) is unbound in the absence of Tc or Dox, resulting in an inactive target gene. In the presence of Tc or Dox, rtTA binds to the TRE and activates the expression of the target gene.

Figure 61: Cloning steps for preparation of the lentiviral plasmid. A picture of 2 % agarose gel shows the PCR product of the genes encoding the Linker peptide extracted from the pET28b vector containing genes for CoREST at two different annealing temperatures, 58 °C and 60 °C. The PCR product was then cloned into a pM vector containing a GAL4 DNA-binding domain (DBD). Genes of the Linker and GAL4 DBD were extracted from the pM vector and cloned into the Gateway entry vector (pENTR2B), which were finally transferred to a pLenti CMV/TRE3G PuroDEST destination vector.

Figure 62: 15 % SDS-PAGE gels showing (a) the purified GAL4 protein (23 kDa) and (b) the purified Linker-GAL4 protein (33 kDa). Western blot analyses probing for LSD1 binding to GST-CoREST with addition of (c) the GAL4 alone and (d) Linker-GAL4. The additional GAL4 domain affects the capability of the Linker peptide to disrupt the binding interaction to some degree by inhibiting the binding interaction at a higher concentration than the Linker-TAT peptide (see Figure 47-(a) for comparison). Nonetheless, the Linker-GAL4 peptide specifically inhibits the binding interaction between LSD1 and CoREST. GAL4 was used as a control and did not affect the binding interaction as expected.

Figure 63: Domain structure of CoREST. CoREST entails its role as a bridge between two enzymes, LSD1 and HDAC1, and nucleosomes. The ELM2 and SANT1 domains are essential for deacetylation activity of HDAC1/2, whereas the Linker region and SANT2 domain are required for LSD1 demethylation activity toward nucleosomes.

Figure 64: Pictures of 10 % SDS-PAGE gels showing representative expression levels of CoREST in (a) BL21 Star (DE3) E.coli strain, (b) BL21 (DE3) E.coli strain, (c) BL21 (DE3)-codonplus RIL E.coli strain, (d) Rosetta (DE3) E.coli strain, (e) BL21 (DE3) E.coli strain with co-transformation of a plasmid expressing the chaperone GroESEL, and (f) BL21 (DE3) Star E.coli strain. The cells were grown at 37 °C for 4 h after 0.5 mM IPTG.
induction except (f). (f) Cells were cultured at 4 °C for 7 days with 0.5 mM IPTG induction. Pre: before IPTG induction, 1h-4h: time points after IPTG induction, d1-d7: day points after IPTG induction S: soluble, IS: insoluble. Expressed CoREST<sup>53-482</sup> is indicated by red arrows.

Figure 65: A picture of 10 % SDS-PAGE gel showing the progress for the preparation of the refolded CoREST<sup>53-482</sup>. The insoluble CoREST<sup>53-482</sup> was washed multiple times to remove unwanted proteins associated with or entrapped in the insoluble protein and unfolded by addition of 8 M urea. CoREST<sup>53-482</sup> was then refolded by gradually removing the denaturant through step dialysis. IS: insoluble CoREST<sup>53-482</sup>, w1-w3: washing the insoluble protein, UF: unfolded CoREST<sup>53-482</sup>, RF: refolding CoREST<sup>53-482</sup>, CRF: concentrated refolded CoREST<sup>53-482</sup>. A red arrow indicates CoREST<sup>53-482</sup>.

Figure 66: A representative AKTA FPLC profile and pictures of 10 % SDS-PAGE gels showing CoREST<sup>53-482</sup> purification on chelating sepharose Ni (II) affinity chromatography. A blue line represents UV absorbance at 280 nm as a function of retention volume (mL), and a green line represents a gradient of elution buffer. Arrows in the chromatography indicate the putative peaks where CoREST<sup>53-482</sup> might be eluted off from the column. Consistently analysis of 10 % SDS-PAGE gels shows that CoREST<sup>53-482</sup> was eluted off from the column throughout the entire run. The protein in the 2nd lane indicates the concentrated, refolded CoREST<sup>53-482</sup> loaded onto the column for purification.

Figure 67: A picture of 8% native PAGE gel showing the formation of the LSD1-CoREST<sup>53-482</sup> complex. lane 1: protein maker; lane 2: LSD1 with molecular weight of 80 kDa; lane 3: CoREST<sup>53-482</sup> with molecular weight of 53 kDa; lane 4: a LSD1-CoREST complex with molecular weight of approximately 130 kDa. The native PAGE could not clearly prove as to whether the refolded CoREST<sup>53-482</sup> is functionally active and forms a complex with LSD1.
List of Schemes

Scheme 1: Chemical mechanisms of demethylation by (a) PADI4 (peptidyl arginine deiminase 4), (b) LSD1 (lysine-specific demethylase 1), and (c) JHDM (JmjC domain-containing histone demethylase).

Scheme 2: Hypothetical mechanisms by which LSD1 achieves its oxidation reaction with H3K4 dimethylated substrates [49]. (a) In the hydride transfer mechanism, a hydride is directly transferred to the flavin. (b) In the nucleophilic mechanism, the flavin forms a covalent bond with the methylated substrate. (c) Single electron transfer mechanism includes a transfer of a proton and electron. The imine intermediate produced by either mechanisms described above then hydrolyzes to produce formaldehyde byproduct and the demethylated lysine side chain on histone H3.

Scheme 3: Enzymatic scheme for coupled assay to detect LSD1 activity. HRP (horseradish peroxidase) converts a byproduct of LSD1 enzymatic activity, H$_2$O$_2$, to O$_2$ while reducing amplex red into fluorescent resorufin.

Scheme 4: GST pull-down assay. CoREST$^{286-482}$ is GST- tagged, which is captured on glutathione agarose beads. Peptides (indicated by P in light-orange color) corresponding to the sequences of CoREST Linker$^{293-380}$ are added to GST-CoREST$^{286-482}$, which is followed by the addition of LSD1. The mixture is incubated for 1 h at room temperature. Peptides are expected to disrupt the binding interaction by competitively binding to LSD1. The LSD1-peptide complex is washed away and any bound LSD1 to GST-CoREST$^{286-482}$ is eluted, resolved by SDS-PAGE, and analyzed by Western blotting. As a control, no CoREST$^{286-482}$ is loaded onto the glutathione agarose beads; thus, LSD1 is washed away.

Scheme 5: Spot synthesis technique mainly involves three phases: preparation of the cellulose membrane, stepwise couplings of the amino acids, and the cleavage of the side-chain protection groups. The presence of bound LSD1 to peptides can be detected via Western blot analysis.

Scheme 6: Schematic outline of the xenograft experiment. The 293FT cell line is transfected with three plasmids (an expression plasmid, a packaging plasmid, and an envelope plasmid) for production of lentiviral particles. The lentiviruses are then used to transduce the Tet-controlled MCF7 cells where the viral gene is integrated into the DNA of MCF7 cells to be expressed. This is followed by selection of stably transduced MCF7 cells. The cells should be tested to validate whether they express the Linker peptide in a controlled manner and affect ERα-mediated functions before being xenotransplanted.
into nude mice. Once transplanted, the volume of tumor will be observed after 7-8 weeks of the treatment with the Linker peptide expressed. Abbreviation: Tet: tetracycline.

Scheme 7: A schematic flow chart for refolding process of insoluble proteins.
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACES</td>
<td>N-(2-acetamido)-2-aminoethanesulfonic acid</td>
</tr>
<tr>
<td>aDMA</td>
<td>asymmetric Dimethylarginines</td>
</tr>
<tr>
<td>AOF2</td>
<td>Amine oxidase domain 2</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen receptor</td>
</tr>
<tr>
<td>BHC</td>
<td>BRAF-HDAC complex</td>
</tr>
<tr>
<td>Bp</td>
<td>Base pairs</td>
</tr>
<tr>
<td>BRAF</td>
<td>BRCA2-associated factor</td>
</tr>
<tr>
<td>BRCA2</td>
<td>Breast cancer 2 susceptibility protein</td>
</tr>
<tr>
<td>CBP</td>
<td>p300/cyclic-AMP-response-element binding protein</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>CoREST</td>
<td>Co-repressor for RE1 Silencing Transcription factor</td>
</tr>
<tr>
<td>CPP</td>
<td>Cell penetrating peptide</td>
</tr>
<tr>
<td>DAPI</td>
<td>4’6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA binding domain</td>
</tr>
<tr>
<td>DCM</td>
<td>Dichloromethane</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
</tbody>
</table>
DNA  Deoxyribonucleic acid
Dnmt1  DNA methyltransferase 1
Dox  Doxycycline
DTT  Dithiothreitol
E2  Estradiol
EDTA  Ethylenediaminetetraacetic acid
ELM2  Egl-27 and MTA1 homology 2
ER  Estrogen Receptor
ERE  Estrogen responsive element
FAD  Flavin adenine dinucleotide
Fmoc  Fluorenylmethyloxycarbonyl chloride
FP  Fluorescence polarization
FPLC  Fast protein liquid chromatography
FRET  Fluorescence resonance energy transfer
GAPDH  Glyceraldehyde 3-phosphate dehydrogenase
GFP  Green fluorescent protein
GST  Glutathione S-transferase
HA  Hemagglutinin
HAT  Histone acetyltransferase
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>HMT</td>
<td>Histone methyltransferase</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>HSV</td>
<td>Herpes simplex viruses</td>
</tr>
<tr>
<td>IP</td>
<td>Immunoprecipitation</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>ITC</td>
<td>Isothermal Titration Calorimetry</td>
</tr>
<tr>
<td>JmjC</td>
<td>Jumomji C-domain containing protein</td>
</tr>
<tr>
<td>$K_a$</td>
<td>Association binding constant</td>
</tr>
<tr>
<td>$K_d$</td>
<td>Dissociation binding constant</td>
</tr>
<tr>
<td>$k_{on}$</td>
<td>Association rate constant</td>
</tr>
<tr>
<td>$k_{off}$</td>
<td>Dissociation rate constant</td>
</tr>
<tr>
<td>KDM</td>
<td>Lysine demethylase</td>
</tr>
<tr>
<td>LB</td>
<td>Lennox Broth</td>
</tr>
<tr>
<td>LBD</td>
<td>Ligand binding domain</td>
</tr>
<tr>
<td>LSD1</td>
<td>Lysine Specific Demethylase 1</td>
</tr>
<tr>
<td>LSD2</td>
<td>Lysine Specific Demethylase 2</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MAO</td>
<td>Monoamine oxidase</td>
</tr>
<tr>
<td>MCF7</td>
<td>Michigan Cancer Foundation-7</td>
</tr>
<tr>
<td>MES</td>
<td>2-(N-morpholino)ethanesulfonic acid</td>
</tr>
<tr>
<td>MMA</td>
<td>Monomethylarginines</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>n</td>
<td>Stoichiometry</td>
</tr>
<tr>
<td>NCoR</td>
<td>Nuclear receptor corepressor</td>
</tr>
<tr>
<td>NuRD</td>
<td>Nucleosome remodeling and deacetylase</td>
</tr>
<tr>
<td>PADI4</td>
<td>Peptidyl arginine deiminase 4</td>
</tr>
<tr>
<td>PAO</td>
<td>Polyamine oxidase</td>
</tr>
<tr>
<td>PIPES</td>
<td>Piperazine-N,N′-bis(2-ethanesulfonic acid)</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulphonyl fluoride</td>
</tr>
<tr>
<td>PR</td>
<td>Progesterone receptor</td>
</tr>
<tr>
<td>PRMT</td>
<td>Protein arginine N-methyltransferase</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative real time PCR</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>rtTA</td>
<td>Reverse tetracycline-controlled transactivator protein</td>
</tr>
<tr>
<td>SAHA</td>
<td>Suberoylanilide hydroxamic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosylmethionine</td>
</tr>
<tr>
<td>SANT</td>
<td>Swi3/Ada2/NCoR/Transcription factor</td>
</tr>
<tr>
<td>sDMA</td>
<td>symmetric dimethylarginines</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard error mean</td>
</tr>
<tr>
<td>SERM</td>
<td>Selective estrogen receptor modulator</td>
</tr>
<tr>
<td>SIN3</td>
<td>SWI-independent 3</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>SMRT</td>
<td>Silencing mediator for retinoid and thyroid hormone receptors</td>
</tr>
<tr>
<td>SNAG</td>
<td>Snail/Gfi domain</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>SPROX</td>
<td>Stability of proteins from rates of oxidation</td>
</tr>
<tr>
<td>SRC</td>
<td>Steroid receptor coactivator</td>
</tr>
<tr>
<td>SWIRM</td>
<td>Swi3p/Rsc8p/Moira</td>
</tr>
<tr>
<td>TAT</td>
<td>Trans-activator of transcription</td>
</tr>
<tr>
<td>TB</td>
<td>Terrific Broth</td>
</tr>
<tr>
<td>TCA</td>
<td>Trichloroacetic acid</td>
</tr>
<tr>
<td>TET</td>
<td>Tetracycline (Tc)</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>TIS</td>
<td>Triisopropylsilane</td>
</tr>
<tr>
<td>TRE</td>
<td>Tetracycline-responsive promoter element</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>TSA</td>
<td>Trichostatin A</td>
</tr>
<tr>
<td>tTA</td>
<td>Tetracycline-controlled transactivator protein</td>
</tr>
<tr>
<td>ΔASA</td>
<td>Change in solvent-accessible surface area</td>
</tr>
<tr>
<td>ΔC_p</td>
<td>Change in heat capacity</td>
</tr>
<tr>
<td>ΔG</td>
<td>Change in Gibbs free energy</td>
</tr>
<tr>
<td>ΔH</td>
<td>Change in enthalpy</td>
</tr>
<tr>
<td>ΔS</td>
<td>Change in entropy</td>
</tr>
</tbody>
</table>
Acknowledgements

I am marking the end of a long journey in obtaining my Ph.D. for which there are many people that I would like to thank for their love and support. First of all, a special thank you to Dr. Dewey McCafferty for guiding me on this great project and opening my eyes to the world of science over the years. I am grateful to you for moving me into the right direction whenever I was lost. Truthfully I have enjoyed working in your lab, and I believe my experiences here will be the cornerstone of my future career.

Thank you to my committee members, Dr. Jiyong Hong, Dr. Pei Zhou, and Dr. Katherine Franz. I really appreciate your advice and support throughout my graduate school years. Also I owe a tremendous debt of gratitude to Dr. Ching-yi Chang in the Department of Pharmacology and Cancer Biology. Thank you for your patience in answering all my questions.

I have met the most amazing group members while working in this lab. A special thanks to Dr. Helena Gaweska for helping me set up the experiments at the beginning of my research, Dr. Julie Pollock for giving me valuable advice and discussing my research problems all the time, and Kate Clancy for your support and companionship through the tough times over the years. I have loved every moment I have spent with all my lab members- Puffy vest day, Friday Fun day, step classes, chasing lab animals, the Christmas party, blaming Jon, and all other moments. My lab life with you all has
nourished me intellectually and culturally. I would have never finished my journey
toward the Ph.D. if not for these great moments!

Most especially, I thank all my family for their continuous support and love
throughout my time in graduate school. Talking to Mom over the phone almost every
day provided one of the greatest energies that helped me survive here. Thank you, Mom,
for listening to all my struggles in daily life and praying for me all the time. Thank you,
Dad, for making me strong and teaching me by example how to face life with a positive
attitude. I could have never learned to be brave and patient if not for your wonderful
advice. Also thank you, Jin, for sharing all my frustrations and great moments. I love
you all.
1. Introduction

1.1 Epigenetics

The word “epigenetic” literally means “in addition to changes in genetic sequence”, which was originally described by Conrad Waddington in the 1940s [1]. This term has evolved to include any process that alters gene expression without a change in DNA sequences. To date, DNA methylation and various covalent modifications of histones are the most studied epigenetic mechanisms. These mechanisms are involved in many normal cellular processes such as cell differentiation, proliferation, and development processes. On the other hand, they can also be responsible for some disease states when they switch off genes that we need to remain healthy such as those that suppress tumor progression. Accordingly, over the past decades, there has been an explosion in studies of epigenetics in human diseases, suggesting that there is a pressing need to understand how the epigenetic changes participate in human diseases and how they can be regulated. Particularly, modulation of epigenetic enzymes has become an attractive target for therapeutic intervention.

1.1.1 Chromatin, Nucleosomes, and Histones

Epigenetic changes are associated with changes in chromatin structure. Chromatin has the dynamic combination where DNA is highly folded and compacted with histone and nonhistone proteins. Its fundamental unit, the nucleosome, consists of
147 base pairs (bp) of DNA organized around an histone octamer that is made up of two copies of each core histone protein –H2A, H2B, H3, and H4 (Figure 1) [2]. These histone proteins are arranged as a central (H3-H4)2 tetramer that is capped on either side by an H2A-H2B dimer, forming a structured-globular domain [3]. Each histone core protein also contains an unstructured N-terminal tail domain that extends freely out of the nucleosome, which provides sites for an enormous number of post-translational modifications discussed later. The histone octamer represents positively charged basic elements that interact with the negatively charged phosphate groups of the DNA to form strong and stable contacts with each other and create the condensed structure of chromatin. Another histone protein called the linker histone H1 is believed to link the nucleosomes together, facilitating the formation of higher-order structures of chromatin [3, 4, 5]. Taken together, the configuration of nucleosomes can be described as “beads on a string” [5, 6]; the string is DNA, and each bead is a nucleosome core particle. The beads on a string represent the first level of chromosomal DNA packing (Figure 2).
Figure 1: Nucleosome structure and histone octamer organization. Nucleosome consists of 147 bp (base pairs) of DNA (yellow) wrapped around the histone octamer being made up of two copies of core histone proteins - H2A (purple), H2B (cyan), H3 (blue), and H4 (pink). Figure was generated using PyMoL (PDB file: 1EQZ).
Figure 2: The major structures in DNA compaction. Chromatin that makes up chromosomes is composed of nucleosomes appearing as beads on a string. Each nucleosome consists of DNA wrapped around histone proteins, and the tails of these histones protrude from nucleosomes (adapted from [6]).
The chromatin can fold upon itself to compact the nucleosomes, forming a highly condensed structure, referred to as heterochromatin (Figure 3-(a)). However, in order for genes to be transcribed into mRNA molecules, promoter regions of the genes need to interact with particular proteins such as transcription factors and other regulatory proteins that initiate the mRNA transcription. In the condensed heterochromatin structure these regions are mostly inaccessible to the transcription machinery; thus, the expression of the genes is inhibited [7, 8]. In order for the genes to be expressed, the heterochromatin structure needs to be relaxed in a controlled manner, so as to expose the promoter regions of the genes to the transcriptional apparatus. This more loose and relaxed structure is referred to as euchromatin (Figure 3-(b)). The change from heterochromatin to euchromatin can mainly be accomplished by chromatin-remodeling complexes and post-translational modifications of the histone tails by specific enzymes [7, 8, 9].
Figure 3: Schematic representation of structural differences between (a) heterochromatin and (b) euchromatin.
1.1.2 Post-translational Modifications of Histones

In recent years, much work has progressed toward correlating the structural and functional properties of chromatin with patterns of histone modifications and subsequently relating these modifications with various cellular processes.

Each histone protein, particularly H2A, H2B, H3, and H4, contains a long, flexible N-terminal tail that extends outward from the core nucleosome. These tails comprise ~25-30% of the mass of the histones. They do not contribute significantly to the structure of individual nucleosomes, but do play an important role in controlling the folding of nucleosomal array into higher-order structure by providing sites for a variety of post-translational modifications. The post-translational modifications include acetylation, phosphorylation, methylation, ubiquitination, sumoylation, and ADP-ribosylation, which markedly influence the ability of associated DNA to be transcribed by altering the normal interactions with DNA and thus modifying the recruitment of other regulatory proteins (Figure 4) [10, 11].
Figure 4: Post-translational modifications of the core histones. Histone tails are methylated at lysines and arginines (red), acetylated at lysines, phosphorylated at serines or threonines (blue), and ubiquitylated at lysines (green).
These modifications serve as important markers recognized by specific protein complexes. However, many of the modification sites are close each other, so that the modification by one enzyme can influence the efficiency of subsequent enzymes. This suggests the multifaceted collaboration of histone modifications to regulate unique biological outcomes, which is commonly referred as to the “histone code” hypothesis [11, 12, 13]. The histone code hypothesis postulates that combinatorial histone modifications, acting sequentially or in combination, regulate critical downstream functions (Figure 5) [12, 13, 14, 15]. However, how the histone code is established and maintained still remains elusive. Understanding the histone code and the specific roles for histone modifications will be key to understanding the mechanisms underlying epigenetic control of cellular activities.
Figure 5: Histone code. The modifications probably act as specific receptors for protein modules that mediate downstream events leading to gene transcription. Abbreviations: HAT: histone acetyltransferase; HDAC: histone deacetylase; HMT: histone methyltransferase; Ac: acetyl group; Me: methyl group; P: phosphate group (Adapted from [11]).
Post-translational modifications of histones occur by the various groups of proteins that are largely categorized by functions as writing, erasing, and reading [14]. Various post-translational modifications are introduced by ‘writers’, and the functions of such writers are antagonized by ‘erasers’ that remove the histone modifications. Some of the modifications are recognized by ‘readers’ that have conserved recognition of specific chemical groups in the histone sequence context [14, 15]. The highly diverse ensemble of these histone-modifying proteins plays important roles in regulating gene expression in various cellular processes and subsequently influencing pathological conditions.

Histone acetylation is the well-characterized histone modification because of its direct effect on chromatin structure. Histone acetylation was discovered over 30 years ago, but the first transcription-related histone acetyltransferase (HAT), Gca5, was identified in 1996 and characterized as a transcriptional co-activator protein [16]. Since then, most of the identified human HATs have long been functioned as transcriptional co-activators [16, 17]. Now a large number of co-activator proteins such as p300/cyclic-AMP-response-element binding protein (CBP), is recognized to have HAT activity [17]. Acetylation of lysine residues removes positive charges of the lysine side chain and reduces the affinity for the negatively charged DNA, thereby opening the condensed chromatin structure to allow the easier access of transcription factors to promoter regions. Histone acetylation is dynamically reversed by histone deacetylases (HDACs). HDACs induce transcriptional repression through chromatin condensation. So far 20
putative HDACs have been identified, which can be classified into three distinct families [18]. HDAC1 and HDAC2 are the best-characterized and main components of multiprotein transcriptional repression complexes such as SIN3-HDAC and NuRD complex [18, 19]. Due to the implications of HDACs in transcriptional repression of tumor suppressor gene expression, histone deacetylase inhibitors (HDACIs) are emerging as a novel class of anticancer agents. In fact, the treatment of tumor cells with inhibitors of class I and II HDACs results in growth arrest and apoptosis [20, 21, 22].

In addition, phosphorylation of histones has been the subject of the active areas of research. The histone H3 serine 10 (H3S10) has been an important phosphorylation site for transcription and is crucial for cell-cycle progression during mitosis and meiosis [23, 24]. Histone ubiquitination was one of the least understood histone modifications due to the lack of information regarding the responsible enzymes. Recently, however, many of the enzymes responsible for the addition and removal of ubiquitin from the histone H2A and H2B have been identified and characterized [25]. Histone ubiquitination plays critical roles in regulating transcription initiation and elongation, and DNA repair [25]. Each histone modification and its respective function are summarized in Table 1 and Table 2.
Table 1: Different classes of modifications identified on histones (Adapted from [26]).

<table>
<thead>
<tr>
<th>Chromatin Modifications</th>
<th>Residues Modified</th>
<th>Function Regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylation</td>
<td>K-ac</td>
<td>Transcription, Repair, Replication, Condensation</td>
</tr>
<tr>
<td>Methylation (Lysines)</td>
<td>K-me1 K-me2 K-me3</td>
<td>Transcription, Repair</td>
</tr>
<tr>
<td>Methylation (Arginines)</td>
<td>R-me1 R-me2a R-me2s</td>
<td>Transcription</td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>S-ph T-ph</td>
<td>Transcription, Condensation</td>
</tr>
<tr>
<td>Ubiquitination</td>
<td>K-ub</td>
<td>Transcription, Repair</td>
</tr>
<tr>
<td>Sumoylation</td>
<td>K-su</td>
<td>Transcription</td>
</tr>
<tr>
<td>ADP-ribosylation</td>
<td>E-ar (ADP-ribose)</td>
<td>Transcription</td>
</tr>
</tbody>
</table>
Table 2: Enzymes associated with histone modifications (Adapted from [26, 27, 28]).

<table>
<thead>
<tr>
<th>Histone Modifications</th>
<th>Residues Modified</th>
<th>Positions Modified</th>
<th>Enzymes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylation</td>
<td>K-ac</td>
<td>H3 K9</td>
<td>PCAF/GCN5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H3 K14</td>
<td>PCAF/GCN5, CBP/p300, SsSAS3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H3 K18</td>
<td>PCAF/GCN5, CBP/p300</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H3 K23</td>
<td>SsSAS3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H3 K56</td>
<td>ScRTT109</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H4 K5</td>
<td>CBP/p300, TIP60, HB01, HAT1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H4 K8</td>
<td>CBP/p300, TIP60, HB01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H4 K12</td>
<td>TIP60, HB01, HAT1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H4 K16</td>
<td>TIP60, SsSAS2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H2A K5</td>
<td>CBP/p300</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H2B K12, 15</td>
<td>CBP/p300</td>
</tr>
<tr>
<td>Deacetylation</td>
<td>K-ac</td>
<td>Non-histone</td>
<td>HDAC I (HDAC1, 2, 3, and 8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>proteins</td>
<td>HDAC II (HDAC4, 5, 6, 7, and 9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H4 K16</td>
<td>HDAC III (NAD+ dependent- SIRT family)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>: HDACs show no specificity for particular acetyl groups</td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>S-ph T-ph</td>
<td>H3 S10</td>
<td>Aurora B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H3 T3</td>
<td>Haspin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H3 S28</td>
<td>MSK1, MSK2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H4 S1</td>
<td>CKII</td>
</tr>
<tr>
<td>Ubiquitination</td>
<td>K-ub</td>
<td>H2A K119</td>
<td>hPRC1L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H2B K120</td>
<td>UbcH6, RNF20/40</td>
</tr>
<tr>
<td>Methylation (Lysines)</td>
<td>K-me1 K-me2 K-me3</td>
<td>H3 K4</td>
<td>MLL, ALL-1, Set9/7, ALR-1/2,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ALR, Set1, ASH1, Sc/Sp SET1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SUV39H1, SUV39H2, G9a,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ESET/SETDB1, EuHMTase/GLP,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CLL8, SpCir4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EZH2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HYPB, SMYD2, N5D1, Set2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DOT1, Sc/SpDOT1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SUV4 20h1, SUV4 20h2, SpSet9</td>
</tr>
<tr>
<td>Demethylation (Lysines)</td>
<td>K-me1 K-me2 K-me3</td>
<td>H3 K4</td>
<td>LSD1, Jarid1A-D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LSD1, JHDMA2A, JMJD2A-D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UTX, JMJD3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>JHDM1, JHDM2A-C, FBXL10</td>
</tr>
<tr>
<td>Methylation (Arginines)</td>
<td>R-me1, R-me2</td>
<td>H3 R2, R17, R26</td>
<td>CARM1 (PRMT4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H4 R3</td>
<td>PRMT1</td>
</tr>
</tbody>
</table>
1.2 Histone Methylation and Demethylation

Histone methylation has emerged as another important histone modification that regulates chromatin structure and function. Histone methylation, which occurs at lysine and arginine residues, has been linked to a number of cellular processes including DNA repair, replication, transcriptional activation, and repression [29, 30]. Unlike acetylation, methylation does not alter the charge of arginine and lysine residues, but instead, increasing methyl groups (mono, di or tri) elevates its basicity and hydrophobicity, which may alter inter- or intramolecular interactions or create new binding surfaces for other regulatory proteins, referred to as reader proteins, that bind preferentially to the methylated domain [30]. In other words, histone methylation serves as information storage marker as opposed to disruptors of DNA-histone interactions. As such, the mechanistic and functional consequences of methylation are typically restricted to the proteins and domains that recognize the modifications.

The methylation of arginine residues is catalyzed by the protein arginine N-methyltransferase (PRMT) family of enzymes [31, 32]. There are three main forms of methylated arginine identified in eukaryotes (Figure 6-(a)): monomethylarginines (MMA), asymmetric dimethylarginines (aDMA), and symmetric dimethylarginines (sDMA), each of which has potentially different functional outcomes. The methylated arginine residues are then recognized by methylarginine-binding domain. Currently, Todor domain is the only major group of protein modules that binds to methylated
arginines [33]. This domain mediates methylation-dependent protein-protein interactions and has crucial roles in orchestrating RNA metabolism and DNA damage repair [33]. Distinct enzymes to reverse arginine methylation have not been identified yet, but it was proposed that arginine could be converted to citrulline by deimination mechanism, which antagonizes arginine methylation (Scheme 1-(a)) [34]. Peptidyl arginine deiminase 4 (PADI4) has shown the potential methylarginine deiminase activity in the presence of Ca²⁺ ions [35].

The methylation of lysine has been more extensively studied because of its prominence and its array of important functions. Similar to arginine methylation, lysine methylation can occur as mono-, di-, or trimethylated forms on predominantly H3 and H4 by catalytic activities of various histone methyltransferases (HMTs), which use S-adenosylmethionine (SAM) as a cofactor (Figure 6-(b)). Histone methyltransferases (HMTs) identified are listed in the Table 2. Several lysine residues including lysines 4, 9, 27, 36, and 79 of histone H3 and lysine 20 of histone H4 have shown to be preferred sites of methylation [36]. Unlike acetylation, which generally correlates with transcriptional activation, histone lysine methylation can mediate either activation or repression, depending on the sites of methylation. Methylation of H3K9 and H3 K27 is generally associated with transcriptional repression, whereas methylation of H3K4, H3K36, and H3K79 has been implicated in the transcriptional activation processes [36]. Methylated lysine residues are recognized by one of three methyl-lysine binding domains: the
chromodomain, the tudor domains, or the WD40-repeat domain that are capable of specific interactions with methylated lysine residues [29, 37]. The chromodomain specifically interacts with methylation of H3K4, H3K9, and H3K27, the tudor domain with H3K79 and H4K20, and the WD40 repeat domain with H3K4. These domains not only interact with specific lysine residues, but also seem to play a crucial role in discriminating different methylated lysines; the chromodomain of HP1 protein interacts specifically with methylated H3K9, whereas the same domain of Polycomb (Pc) protein interacts with H3K27 [38]. Together, various proteins that contain these binding domains are recruited by specific methylated lysine residues, which may explain why different methylated lysines result in distinctive biological outcomes. However, additional levels of complexity still can exist because of different states of lysine methylation. The binding interaction of protein recruited to particular methylated lysines can be also affected by neighboring modifications such as histone acetylation. Acetylated H3 and H4 are known to be the preferential targets of histone methylation [30], suggesting that histone acetyltransferases and histone methyltransferases may act synergistically to regulate gene expression.
Figure 6: Methylation states of (a) arginine and (b) lysine. Methyl groups are indicated in red.
For many years, histone lysine methylation has been considered as a static and enzymatically irreversible modification with little turnover of the methyl groups due to the high thermodynamic stability of N-CH₃ bond [39]. Beyond this, the belief of methylation irreversibility was based on the experiments demonstrating that the half-life of histone methyl marks was approximately same as that of the histone itself [40]. However, in 1973, Paik and Kim, for the first time, reported an enzymatic demethylation of N-methylated calf thymus although the enzyme was never identified [41]. Since then, the clear descriptions regarding histone demethylation have been limited due to the lack of experimental evidence. In 2002, Shiekhattar, R. et al. discussed BCH110 (BRAF-Histone deacetylase Complex 110) as a component of histone deacetylase complexes and reported that it had similar activity to amine oxidases [42]. In late 2004, it was specifically named as lysine-specific demethylase 1 (LSD1). It is a flavin-dependent amine oxidase that is capable of selective demethylation of mono- or dimethylated lysine 4 of histone 3 (H3K4) (Scheme 1-(b)) [43]. The mechanism and structure of LSD1 will be discussed in details later in this chapter. The initial description of LSD1, hence, opened up a new chapter in histone methylation and demethylation field. Since then, a larger family of histone lysine demethylases, containing the JmjC domain as a signature motif, has been uncovered. The reaction mechanism of JmjC domain-containing histone demethylases (JHDM) relies on iron (II) and α-ketoglutarate as cofactors to catalyze direct hydroxylation of the lysine methylamine group, generating succinate and carbon
dioxide as reaction products (Scheme 1-(c)) [44]. The hydroxyl-methyl group is spontaneously lost as formaldehyde to release a methyl group from lysine residues. Based on the sequence homology in the JmjC domain and overall architecture of associated motifs, JHDMs have been classified into a different class of histone lysine demethylases from LSD1. Unlike LSD1, JHDMs can demethylate all three histone lysine methylation states (mono-, di-, and trimethylation) at H3K4, H3K9, H3K27, and H3K36; they are able to act on trimethylated lysines, which is mechanistically impossible for flavin-catalyzed oxidative demethylation reactions. As shown in the Table 3, several different subfamilies of the JmjC domain-containing histone demethylases have been identified, which have different histone selectivities toward methylation states. Despite the recent discovery of many JmjC-containing histone demethylases, still their functional characterization is at an early stage [44, 45].

Also recently, a second mammalian flavin-dependent histone demethylase, named LSD2, has been discovered. LSD2 represents a similar enzymatic activity to that of LSD1 by sharing 45 % overall identity in the amine oxidase domain with LSD1. Like LSD1, LSD2 displays distinctive substrate specificity for mono- and dimethylated H3K4, but little is known about the actual biological function of LSD2 [46].

The identification of histone demethylases has been considered a major breakthrough in the field of chromatin biology, as it showed that histone methylation is dynamically reversible like histone acetylation/deacetylation than anticipated. Despite
their recent discovery, they have already been demonstrated to perform important roles in diverse biological processes including cellular pathologies, but many important questions centered on their mechanisms and exact roles in transcriptional control have yet to be answered.
Scheme 1: Chemical mechanisms of demethylation by (a) PADI4 (peptidyl arginine deiminase 4), (b) LSD1 (lysine-specific demethylase 1), and (c) JHDM (JmjC domain-containing histone demethylase).
Table 3: Identified histone lysine demethylases and their roles in human diseases (Adapted from [47]).

<table>
<thead>
<tr>
<th>Demethylase</th>
<th>Alternate name</th>
<th>Site Specificity</th>
<th>Demethylase Activity</th>
<th>Transcription Effect</th>
<th>Association with human diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flavin-dependent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSD1</td>
<td>KDM1A; BHC110; AOF2</td>
<td>H3K4</td>
<td>mono, di</td>
<td>Repression</td>
<td>Overexpression in breast, lung, prostate, neuroblastoma, and bladder cancer (putative oncogene)</td>
</tr>
<tr>
<td>LSD2</td>
<td>KDM1B; AOF1</td>
<td>H3K4</td>
<td>mono, di</td>
<td>Repression</td>
<td>Not reported</td>
</tr>
<tr>
<td><strong>Fe (II), α-ketoglutarate dependent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JHDM1A</td>
<td>KDM2A; FBXL11</td>
<td>H3K36</td>
<td>mono, di</td>
<td>Repression</td>
<td>Suppression of NF-κB-dependent growth of colon cancer cells</td>
</tr>
<tr>
<td>JMDJ1A</td>
<td>KDM3A; TSGA</td>
<td>H3K9</td>
<td>mono, di</td>
<td>Activation</td>
<td>Co-activator of AR-mediated transcription</td>
</tr>
<tr>
<td>JMDJ2A</td>
<td>JMDJ3A</td>
<td>H3K9, H3K36</td>
<td>di, tri</td>
<td>Activation/Repression</td>
<td>Overexpression in prostate cancer</td>
</tr>
<tr>
<td>JMDJ2B</td>
<td>KDM4B; KIAA0876</td>
<td>H3K9</td>
<td>di, tri</td>
<td>Activation</td>
<td>Overexpression in prostate cancer</td>
</tr>
<tr>
<td>JMDJ2C</td>
<td>KDM4C; GASC1</td>
<td>H3K9, H3K36</td>
<td>di, tri</td>
<td>Activation/Repression</td>
<td>Overexpression in prostate cancer; Transcriptional activation of AR target genes</td>
</tr>
<tr>
<td>JARID1A</td>
<td>KDM5A; RBP2</td>
<td>H3K4</td>
<td>di, tri</td>
<td>Repression</td>
<td>Overexpression in gastric cancer; Elevated expression in drug-tolerant subpopulation of cancer cells</td>
</tr>
<tr>
<td>JARID1B</td>
<td>KDM5B; PLU-1</td>
<td>H3K4</td>
<td>di, tri</td>
<td>Repression</td>
<td>Overexpression in breast, testis, prostate, ovary cancer; Silencing of tumor suppressor genes (BRAC1)</td>
</tr>
<tr>
<td>JARID1C</td>
<td>KDM5C; SMCX</td>
<td>H3K4</td>
<td>di, tri</td>
<td>Repression</td>
<td>Regulation of human papilloma virus (HPV) oncogene expression</td>
</tr>
<tr>
<td>JARID1D</td>
<td>KDM5D; SMCY</td>
<td>H3K4</td>
<td>di, tri</td>
<td>Repression</td>
<td>Not reported</td>
</tr>
<tr>
<td>PHF8</td>
<td></td>
<td>H3K9</td>
<td>mono, di</td>
<td>Activation</td>
<td>Not reported</td>
</tr>
</tbody>
</table>
1.3 LSD1

1.3.1 LSD1 Mechanism and Structure Overview

As mentioned above, LSD1 (also known as KDM1, BHC110, AOF2) specifically acts on mono- and dimethylated lysine of histone H3 through an oxidative process that requires FAD (flavin adenine dinucleotide) as an essential redox cofactor. The oxidative reaction catalyzed by LSD1 generates an imine intermediate, which then hydrolyzes to produce formaldehyde byproduct and the demethylated lysine side chain on histone H3 (Scheme 1-(b)). Demethylation of trimethylated lysine, however, is prevented by the absence of a protonated nitrogen required for oxidation [43]. Incubation of recombinant LSD1 with various methyl-lysine peptides derived from histones results in robust demethylase activity with high level of specificity towards methylated lysine 4 of histone H3 (H3K4) [48].

A number of plausible mechanisms of LSD1 demethylation activity have been proposed, which include a direct hydride transfer from the substrate to the flavin, nucleophilic attack of the flavin by the substrate, forming a covalent intermediate, and a single electron transfer to form radical intermediate (Scheme 2) [49]. The isoalloxazine ring of FAD is capable of accepting either one or two electrons, which may increase the number of possible catalytic mechanisms of LSD1.

Analysis of the LSD1 structure has led to significant insight into its function. LSD1 is highly conserved in organisms ranging from Schizosaccharomyces pombe to
human, and overall architecture of LSD1 contains three main domains: an N-terminal SWIRM (Swi3p/Rsc8p/Moira) domain, an AO (Amine Oxidase) domain, and a central protruding tower domain (Figure 7) [48, 50]. A recombinant form of human LSD1 with an N-terminal deletion of first 150 residues was used for the determination of crystal structure of LSD1 by several groups [48, 50]. The first 150 residues were predicted to be unstructured due to the susceptibility to proteolytic cleavage, and functionally they contain a nuclear localization signal [50]. The amine oxidase domain (N-terminal residues 271-417 and C-terminal residues 523-833), a catalytic center of LSD1, is structurally similar to that of polyamine oxidases. This domain structurally splits into two halves that consist of a substrate-binding half and an FAD-binding half, forming a globular domain. The active site of LSD1 consists of a considerably large cavity as compared to that of other amine oxidases. The active site is located in the center of this domain at the interface between the substrate- and FAD-binding subdomains. The rims of the active sites are lined with negatively charged residues, which is believed to guide the positively charged substrate binding into the active site [50, 51]. The N-terminal SWIRM domain (residues 172-270) consists of a six-helical bundle and a 310 helix. The exact function of the SWIRM domain in LSD1 has not been established yet, but presumably it contributes to the binding of H3 tail by forming a groove for the C-terminal segment of the H3 tail. Structurally N-terminal SWIRM domain and the C-terminal amine oxidase domain closely pack against each other, creating the binding
interface [50, 51, 52]. It should be noted that despite the large size of the active site of LSD1, it is not large enough to hold the first 21 residues of the histone H3 tail—a sufficient length for detectable LSD1 enzymatic activity. Consequently, it is likely that the surface groove formed between the SWIRM domain and the amine oxidase domain provides an additional binding site for the C-terminal portion of H3 tail. Lastly, the tower domain of LSD1 (residues 418-522) consists of an antiparallel coiled coil in which two $\alpha$-helices pack together in a left-handed superhelix. The tower domain projects away from the catalytic center of LSD1, which is a distinct feature of LSD1 as compared to other amine oxidases and a recently identified LSD2. The tower domain is known to be essential for the demethylation activity of LSD1 by acting as an adaptor to recruit other proteins such as CoREST (Co-repressor for RE1 Silencing Transcription factor) that endows LSD1 with the ability to demethylate nucleosomal substrates. [50, 51, 53, 54]. Although LSD1 and LSD2 share similar identity in the catalytic domain, due to the lack of the tower domain in LSD2, the biology of LSD2 is proposed to differ from that of LSD1. Consistently, the deletion of the tower domain of LSD1 resulted in the decreased demethylation activity of LSD1 toward nucleosomes, indicating its important role as a binding platform for other co-regulatory proteins.
(a) Hydride transfer mechanism

(b) Nucleophilic mechanism
(c) Single electron transfer mechanism

Scheme 2: Hypothetical mechanisms by which LSD1 achieves its oxidation reaction with H3K4 dimethylated substrates [49]. (a) In the hydride transfer mechanism, a hydride is directly transferred to the flavin. (b) In the nucleophilic mechanism, the flavin forms a covalent bond with the methylated substrate. (c) Single electron transfer mechanism includes a transfer of a proton and electron. The imine intermediate produced by either mechanisms described above then hydrolyzes to produce formaldehyde byproduct and the demethylated lysine side chain on histone H3.
Figure 7: (a) Domain and (b) crystal structure of LSD1 with an N-terminal deletion of first 150 residues (NΔ150). LSD1 tower domain (residues 418-522) is colored in blue, amino oxidase domain (N-terminal residues 271-417 and C-terminal residues 523-833) in green, and SWIRM domain (residues 172-270) in purple, respectively. (c) Surface structure of AOD (green) and SWIRM domain (purple). The 16-mer peptide substrate (red) is located at the center of the AOD. The C-terminal of the histone H3 tail is expected to locate at the binding interface between the AOD and SWIRM domain indicated by a pink arrow. All structural figures were generated using PyMoL (PDB file: 2IW5).
1.3.2 The Role of LSD1 in Gene Regulation

LSD1 was originally identified as a component of the BRAF histone deacetylase (HDAC) transcriptional compressor complex containing the REST corepressor, CoREST, known to be important for repression of neuronal genes in non-neuronal cells [42]. Consistent with the role of LSD1 in removal of an active methylation mark on H3K4, thereby promoting repression of gene expression, LSD1 has been found in various corepressor complexes including CoREST and HDACs; consequently, in association with the transcription factor repressors, LSD1 plays a widespread role in various biological processes. In hematopoiesis, the LSD1-CoREST-HDACs complex associates with growth factor independence (Gfi) proteins via SNAG repression domain, controlling hematopoietic differentiation [55]. This complex is also involved in the progression of herpes simplex viruses (HSV) infection, promoting the viral gene expression in the establishment of latency [56]. The mechanism underlying Blimp-1-mediated gene suppression in plasma cell differentiation includes interaction with LSD1 in complex with HDAC and CoREST [57]. It is likely that the LSD1-CoREST complex is recruited by different factors to be utilized for repression in various contexts. Very recently LSD1 has been reported to be present in the NuRD chromatin-remodeling complex also including HDACs and targets metathesis in breast cancer by coupling to chromatin-remodeling Swi/Snf type APTases important for transcriptional repression [58]. In this complex, the metastasis-associated (MTA) proteins function in place of
CoREST and play an important role in directing LSD1 demethylation activity toward nucleosomal substrates. Taken together, these studies suggest that LSD1, much like HDACs, plays an extensive role in establishing repressive chromatin environment as a histone demethylase, and the demethylation by LSD1 and its interaction with CoREST/HDACs imply that LSD1 most likely coordinates with histone deacetylation to play a role in transcription repression. This level of complexity is consistent with the histone code hypothesis as described earlier.

In addition to the role of LSD1 in transcriptional repression, LSD1 functions as a transcriptional activator in association with the androgen receptor (AR) and thus demethylates H3K9, suggesting that the substrate specificity can be modulated by its interacting proteins. Removal of the repressive H3K9 methylation mark by LSD1 and subsequent AR-dependent transcription of target genes have been associated with metastatic prostate cancer and tumor recurrence during therapy [59]. Similarly, LSD1 is also highly expressed in breast cancer cells and has been broadly recruited to approximately 80% of the promoter of estrogen receptor (ER) target genes [60, 61]. Consistently, the dynamic regulation of ER function through post-translational modification by histone demethylation has been widely studied. ER functions as a ligand-dependent transcription factor that regulates gene transcriptions via direct recruitment to target gene chromatin. The transcription function of ER is shown to be influenced by several co-regulatory proteins that are dynamically associated with
histone-modifying enzymes including LSD1 [62]. More details of recent studies for the regulation of ER by LSD1 will be discussed in Chapter 5. It is therefore possible that LSD1 plays a crucial role in transcriptional activation for various nuclear receptors through the demethylation activity.

LSD1 can act on non-histone substrates such as p53 and DNA methyltransferase 1 (Dnmt1). p53, the tumor suppressor, associates with LSD1 to repress p53-mediated transcriptional activation and inhibit the role of p53 in promoting either cell cycle arrest or apoptosis [63]. Recently, DNA methyltransferase 1 (Dnmt1), the major enzyme responsible for maintaining DNA methylation patterns during DNA replication, was identified as a novel substrate for LSD1 [64]. As such, substantial research has been performed to understand the mechanisms and targets of LSD1, indicating that LSD1 has profound effects on numerous biological processes either as a transcriptional repressor or a transcriptional activator, but the exact function of LSD1 as a component of various transcriptional complexes remained to be investigated, especially regarding its interactions with other proteins in complexes.

1.3.3 Associated Factors of LSD1

Through in vitro studies using recombinant human LSD1 and peptide substrates corresponding to the N-terminal sequence of histone H3, it has been shown that LSD1 can demethylate mono- or dimethylated peptide substrates without the aid of other
coregulatory proteins for catalysis [65]. However, LSD1 alone is unable to demethylate nucleosomal substrates in vivo. As described above, LSD1 is associated with a number of co-regulatory proteins to both positively and negatively regulate gene transcription in native environment. Furthermore, some of co-regulatory proteins have shown to directly modulate the specificity and activity of LSD1. For example, the addition of purified HDAC1 and BHC80 has shown no effect on LSD1 catalytic activity, whereas the addition of CoREST has restored the ability of LSD1 to demethylate nucleosomal substrates. Interestingly HDACs in the LSD1 complex are likely to function upstream of LSD1/CoREST by generating a hypoacetylated substrate that is better recognized by LSD1/CoREST (Figure 8-(a)) [66]. BHC80 containing a PHD finger recognizes the demethylated state and prevents remethylation [66]. Both LSD1 and the NuRD complex similarly function in transcriptional repression as the LSD1-CoREST complex. The MTA protein in the NuRD complex contains a SANT domain, which presumably serves as a bridge between LSD1 and nucleosomes, allowing for the demethylation activity of LSD1 (Figure 8-(b)).

In contrast to the role of HDACs/CoREST or HDACs/MTA for LSD1 demethylation activity on H3K4, as briefly described above, the presence of AR is thought to dictate the substrate specificity of LSD1 to H3K9 (Figure 8-(c)). The precise molecular mechanism in which LSD1 switches its specificity from H3K4 to H3K9 is still unclear, but it is most likely that co-regulatory proteins of LSD1 or other localized
histone marks contribute to it. It is possible that due to the close proximity of the tower domain of LSD1 to the catalytic center, binding to other proteins through the tower domain may alter the conformation of the catalytic domain, affecting its substrate specificity and activity. As such, protein-protein interactions between LSD1 and its binding proteins appear to be important to regulate LSD1 activity and its substrate specificity. Our group has begun to decode the biochemical and biophysical basis for coregulatory proteins in LSD1 biology. Particularly, we are interested in study of the association of LSD1 with CoREST, the prominent binding protein of LSD1, to understand how the LSD1-interacting protein can contribute to the modulation of LSD1 demethylation activity in cellular environments.
Figure 8: Key histone demethylase complexes. LSD1 plays an important role in transcriptional regulation by associating with a variety of proteins. (a) LSD1-CoREST complex. (b) LSD1-NuRD complex. (c) LSD1-nuclear hormone receptor (AR/ER) complex. Abbreviations: AR/ER: androgen receptor/estrogen receptor; BRAF35: BRCA2-associated factor 35; BRCA2: breast cancer 2 susceptibility protein; BHC80: BRAF and histone deacetylase complex 80; HDAC1/2: histone deacetylase 1/2; MBD: methyl-CpG-binding domain protein 2; Mi2: myositis autoantigen 2; MTA: metastasis associated protein; RbAp: retinoblastoma-associated protein (Adapted from [54]).
1.3.3.1 Overview of CoREST

CoREST is a 66 kDa protein that functions as a corepressor to the silencer REST (RE1 silencing transcription factor/ neural restrictive silencing factor). The REST protein was originally thought to play an important role in the regulation of neuronal gene expression because it binds to RE1 sites identified in several neuron-specific genes [67]. However, the observations on how this protein regulates gene expression by recruiting various chromatin-modifying enzymes have revealed that the REST protein regulates gene expression throughout the body [68]. REST mediates gene expressions by recruiting two separate corepressor complexes, mSin3 and CoREST; its N-terminus interacts with the mSin3 complex, whereas the C-terminal region of REST associates with the CoREST complex [68, 69]. As such, CoREST is a component of corepressor complexes with other proteins such as HDACs, LSD1, BRAF35, and BHC80. In the complex, CoREST plays an important role in targeting histone H3K4 demethylation by LSD1 in the context of nucleosomes. It consists of an ELM2 (Egl-27 and MTA1 homology 2) domain, two SANT (Swi3/Ada2/NCoR/Transcription factor IIIB) domains, and the linker region between the two SANT domains (Figure 9-(a)) [69, 70, 71]. Two SANT domains resemble the DNA-binding domains of Myb-related DNA-binding proteins. Particularly, N-terminal SANT domain of CoREST (SANT1) has shown to be necessary for interaction with HDAC1 [70]. Removal of this region causes loss of associated HDAC1 and histone deacetylase activity. However, the significant role of SANT1
domain for LSD1 demethylation activity has not known yet. Presumably, the SANT1 domain may act as a bridging sequence between LSD1 and its substrates because of the structural similarity with the SANT2 domain; thus, the SANT1 and SANT2 domains may independently facilitate LSD1-mediated demethylation. The linker region of CoREST between two SANT domains consists of a short $\alpha$-helix (residues 316-325) and a long $\alpha$-helix (residues 330-363) that are arranged in a configuration of the letter L (Figure 9). When bound to LSD1, the short $\alpha$-helix of the linker region packs against the substrate binding subdomain of AOD of LSD1, whereas the long $\alpha$-helix forms a parallel coiled coil with LSD1 tower domain. The SANT2 domain of CoREST links to the long $\alpha$-helix of the linker through a flexible loop and is located at the tip of the long stalk formed by helices of LSD1 and CoREST. Owing to the functional importance of the linker region and SANT2 domain of CoREST in LSD1 activity, only these domains have been co-crystallized with LSD1.
Figure 9: (a) Domain structure of CoREST. The essential domains of CoREST, the linker region and SANT2 domain, for LSD1 demethylation activity are highlighted in red and pink, respectively. (b) Crystal structure of C-terminal segment of CoREST including the linker region (red) and SANT2 domain (pink). (c) LSD1 in complex with the C-terminal segment of CoREST. All structural figures were generated using PyMoL (PDB file: 2IW5).
1.3.3.2 The Importance of CoREST in LSD1 Activity

Structural analysis and accumulating evidence have shown that the tower domain of LSD1 represents a surface of CoREST binding. In a similar manner, the linker region of CoREST itself has shown to be sufficient for LSD1 binding, indicating that the helical coiled-coil interaction may be an integral driving force for the binding interaction between LSD1 and CoREST. Though the linker region of CoREST is capable of binding to LSD1, it is insufficient to stimulate the demethylation activity of LSD1 toward nucleosomes, which implies that CoREST SANT domains are crucial for facilitating LSD1 demethylation activity. Particularly, the sequence of CoREST SANT2 domain is highly conserved in the DNA binding residues in v-Myb, which allows us to predict the role of the SANT domain in facilitating the binding with nucleosomal DNA. There are two supportive but conflicting data regarding the role of the SANT1 domain of CoREST. In one study, CoREST fragment containing the SANT1 domain was shown to be insufficient for demethylation activity of LSD1, suggesting a requirement of CoREST SANT2 domain [66]. Additionally, the binding interaction between CoREST SANT2 and DNA was examined by NMR spectroscopy, and their dissociation constant was determined to be approximately 84 μM, and the mutants of CoREST SANT2 domain have shown to fail to bind to nucleosomal DNA [65]. The other study has shown that either the CoREST SANT1 or SANT2 domain promotes LSD1-mediated demethylation activity [71]. Nonetheless, these studies have proposed an essential role of the SANT
domain as a molecular bridge that connects LSD1 to its nucleosomal substrates. The interaction of LSD1 with CoREST through the helical interaction also contributes to the stability of LSD1. Knockdown of CoREST by RNAi led to a reduction of CoREST, which was followed by a reduction of LSD1 expression. When the proteasome inhibitor was added, the LSD1 protein level appeared to be restored, which indicates that the absence of CoREST causes a proteasomal degradation of LSD1 [66]. As such, the C-terminal segment of CoREST plays a significant role in LSD1 demethylation activity and stability.

Despite the lack of the structural information of CoREST in complex with HDACs, previous studies have shown that CoREST also forms a stable histone deacetylase complex, mediating the deacetylation activity of HDAC1 and 2. In contrast to the importance of the C-terminal segment of CoREST in LSD1 activity, not only is the N-terminal segment of CoREST sufficient for HDAC1 association, but also it confers nucleosomal deacetylation to HDAC1. Given the significance of CoREST in mediating both LSD1 and HDAC enzymatic activities, understanding the exact mechanism by which the two enzymes are coupled will provide insights into identification novel means to treat human diseases such as cancer with synergetic effects.
1.3.4 Strategies to Inhibit the Demethylation Activity of LSD1

Because of the implication of LSD1 function and mechanism in the development of human diseases, LSD1 has become a novel therapeutic target. Subsequently, from the knowledge of LSD1 mechanism, there have been increasing efforts to identify or design mechanism-based LSD1 inhibitors for the treatment of cellular pathologies (Figure 10). The AOD of LSD1 shares considerable sequence similarity with those of the FAD-dependent MAO (monoamine oxidase) enzymes responsible for oxidizing arylalkylamine neurotransmitters such as dopamine and serotonin [72], so initially small molecule MAO inhibitors were used to examine whether they may serve as LSD1 inhibitors. MAO inhibitors such as tranylcypromine and phenelzine has shown to potentially inhibit nucleosomal demethylation of H3K4, and the treatment of embryonal carcinoma cells with tranylcypromine led to a global increase in H3K4 methylation level [73]. A series of tranylcypromine analogues were further expanded, which has shown to inhibit LSD1 with a $K_i$ value ranging from 477 to 22 µM with growth inhibition of prostate cancer cells and noticeable effects on cell differentiation of promyelocytic leukemia cells [74, 75]. Due to the sequence homology of LSD1 with amine oxidases such as spermine oxidase and N$^1$-acetylpolyamine oxidase, biguanide and bisguanidine polyamine analogues were examined as inhibitors of LSD1 as well. These analogues showed noncompetitive LSD1 inhibition at < 2.5 µM and were able to induce re-expression of multiple silenced tumor suppressor genes in colon cancer [76].
dimethylated lysine peptide substrate is replaced with residues conjugated with propargylamine or aziridine, the lysine-derivatized peptide exhibited a clear time- and concentration-dependent inhibition of LSD1 [77, 78]. As such, the inhibition assays with mechanism-based LSD1 inactivators have hinted at the therapeutic potential of LSD1 inhibition in the treatment of cellular pathologies. Still, achieving apparent selectivity of those mechanism-based inhibitors for LSD1 remains a significant challenge due to the similarity of LSD1 to other FAD–dependent amine oxidases.
Small molecule MAO inhibitors [72, 73]

- Pargyline
- Tranylcypromine
- Phenelzine

Polyamine Analogs [76]

Substrate Analogs [77, 78]

Figure 10: Mechanism-based inactivators of LSD1.
Figure 11: Mechanism-based inhibitors of LSD1 in complex with FAD. The structures show the covalent adducts involving (a) FAD and tranylcypromine [74] and (b) FAD and propargyl-Lys derivatized peptide inhibitor [78].
1.4 Thesis Objectives

The recent discovery of histone demethylases has initiated a rapid increase in our knowledge about the regulation of histone modifications. LSD1 is a unique histone lysine demethylase in that it belongs to the FAD-dependent family of amine oxidases and functions as either a transcriptional activator or repressor. Also it is noteworthy that its substrate specificity can be modulated by its interacting proteins. Several recent studies have established that LSD1 can be implicated in the regulation of variety of genes and serves as an important link to the development of cellular pathologies such as cancer progression; thus, regulation of LSD1 activity has become a promising therapeutic target. This further suggests that there is a pressing need to better understand the function of LSD1. However, little is known about the molecular details of the mechanistic action and interaction of LSD1 with important protein cofactors such as CoREST. Here CoREST can serve as a model system for understanding how LSD1-interacting proteins contribute to the modulation of LSD1 activity. As described earlier, hence, given the significance of CoREST in directing demethylation of LSD1 to specific nucleosomal substrates, the overall goal of this thesis is to extend our understanding of the mechanism of LSD1 regarding its interaction with CoREST by characterizing the molecular basis of their interaction and to suggest a new means to modulate LSD1 activity by disrupting the potentially critical protein-protein interactions. Towards these goals, we have established the biophysical nature of the binding interaction between
LSD1 and CoREST at molecular level, made efforts to identify key interaction sites between two proteins, and attempted to disrupt their binding to inhibit LSD1 demethylation activity *in vitro* and *in vivo*. Initially, in Chapter 2, we discuss the expression and effective purification of LSD1 and CoREST. In Chapter 3, we describe thermodynamic characterization of the binding interaction between LSD1 and CoREST using isothermal titration calorimetry to provide a fundamental understanding of their interaction. Chapter 4 presents various attempts including site-directed mutagenesis study and peptide fragment analysis to define a ‘hot spot’ - a key interaction region – at the binding interface. Because we have found that the binding energy is evenly spread across the binding interface, in Chapter 5, we perform initial experiments to disrupt the binding interaction between LSD1 and CoREST using the linker region of CoREST, designated as a Linker peptide, as a competitive modulator and evaluate the potential of the Linker peptide to inhibit LSD1 demethylation activity in cellular models of ERα-positive breast cancer cell line (MCF7) where LSD1 is known to co-localize with ERα and affect ERα-transcription activities. The addition of the Linker peptide has increased global methylation levels on H3K4 and H3K9 and further decreased ERα-transcriptional activities of target genes as well as the proliferation of breast cancer cells. Chapter 6 describes our efforts to establish a system to express the Linker peptide as an LSD1/CoREST interaction inhibitor in MCF7 cells in a controlled manner, which will be xenotransplanted into nude mice. This xenograft study is still ongoing at the moment.
We will observe the tumor growth in xenograft mouse model in the presence or absence of the expression of the Linker peptide.

We believe that this research on modulating LSD1 enzymatic activity by interrupting a critical protein-protein interaction may offer a new means to inhibit LSD1 activity independent of targeting the amine oxidase catalytic mechanism and further provide a distinctively new avenue of therapeutic intervention for breast cancer via sequestering LSD1 and in turn disrupting ERα-transcriptional activation.
2. Expression and Purification of LSD1 and CoREST

2.1 Background

This chapter presents preliminary works on the expressions and purifications of LSD1 and CoREST to study their binding interaction. As described in Chapter 1, LSD1 is a flavin-containing oxidase that catalyzes the removal of methyl groups from histone and non-histone substrates. A truncation mutant of LSD1 lacking its first N-terminal 150 residues has exhibited to retain full demethylase activity against methylated H3K4 peptide substrates [65]. The C-terminal segment of CoREST (residues 286-482) has shown to be a functionally essential region for LSD1 demethylase activity. Here we have prepared truncated forms of LSD1 and CoREST as described.

2.2 Expression and Purification of LSD1

The gene encoding the truncated form of human LSD1 in pET-151/D-TOPO vector was originally used for expression and purification. However, the low levels of expressed proteins were observed, which subsequently produced the low yield of LSD1 at less than 1 mg per liter of E. coli culture (Figure 12). In order to enhance the protein expression level, we developed codon-optimized genes encoding the same region of LSD1 in pET-15b. The clone optimized for codon usage in E. coli has successfully produced an increased LSD1 yield to almost 5 mg of protein per liter of E. coli culture.
(Figure 13). In addition to increased amount of LSD1, the purity was improved by an additional purification step with an ion-exchange chromatography.
Figure 12: Representative AKTA FPLC profiles and pictures of 10 % SDS-PAGE gels showing LSD1 purification on (a) chelating sepharose Ni (II) affinity chromatography and (b) Hiprep 26/60 Sephacryl S200 gel filtration chromatography. A blue line represents UV absorbance at 280 nm as a function of retention volume (mL), and a green line represents a buffer gradient. Red arrows indicate the peaks where LSD1 eluted from the column. Details are given in Experimental Sections.
Figure 13: Representative AKTA FPLC profiles and pictures of 15 % SDS-PAGE gel showing LSD1 purification on (a) chelating sepharose Ni (II) affinity chromatography, (b) Hiprep 26/60 Sephacryl S200 gel filtration chromatography, and (c) Q-sepharose anion exchange chromatography. A blue line represents UV absorbance at 280 nm as a function of retention volume (mL), and a green line represents a buffer gradient. Red arrows indicate the peaks where LSD1 was eluted off from the column. Further purification of LSD1 with an additional ion exchange chromatography improved the purity of the protein. Details are given in Experimental Sections.
2.3 Expression and Purification of CoREST$^{286-482}$

Initially the gene encoding CoREST lacking its first N-terminal 52 residues on pET151/D-TOPO vector developed by Dr. Dawn M. Z. Schmidt was used for expression and purification (Experimental details are described in Appendix). However, given the significance of the functional region of CoREST including the linker region and SANT2 domain, whose crystal structure has been previously determined by several groups, in directing demethylation of LSD1 to specific nucleosomal substrates, a truncated version of CoREST lacking its first N-terminal 285 residues (designated as CoREST$^{286-482}$) was expressed, purified, and used for further experiments in this thesis.

2.3.1 Gene Cloning for CoREST$^{286-482}$

A vector encoding a truncated CoREST$^{286-482}$ was constructed using a pET28b vector as described in Experimental Section. The gene of our interest was extracted from an original template, CoREST$^{53-482}$ by PCR at different annealing temperatures of primers, and as shown in Figure 14, all samples that were run on 1% agarose gel showed the correct sized PCR products, which were then digested with the relevant enzymes, generating compatible sticky ends to assist insertion into a expression vector, a pET28b.
Figure 14: (a) Picture of 1% agarose gel of PCR product of the extraction of the gene encoding CoREST\textsuperscript{286-482}. PCR was performed at different annealing temperatures. (b) A vector map of pET28b indicates the portion of CoREST\textsuperscript{286-482} gene.
2.3.2 Expression and Purification of CoREST\textsuperscript{286-482}

CoREST\textsuperscript{286-482} was overexpressed in several \textit{E. coli} strains in order to determine the optimal yield. The cDNA encoding CoREST\textsuperscript{286-482} was cloned into the pET28b vector as shown in Figure 14-(b), which contains an C-terminal His\textsubscript{6} tag, and the resulting plasmid was transformed into electrocompetent BL21 (DE3), BL21 Star (DE3), and C43 (DE3) \textit{E. coli} strains. The cells were grown under different temperatures (23 °C and 37 °C) with 0.5 mM IPTG induction. Of the strains used, BL21 Star (DE3) \textit{E. coli} cells, which carry an \textit{rne} gene mutation encoding a truncated RNaseE gene that loses the ability to degrade mRNA, exhibited the best expression of CoREST\textsuperscript{286-482} (Figure 15). Growing cells at higher temperatures (37 °C) led to better expression but yielded poor solubility. The supernatant (soluble fraction) was separated from insoluble aggregates called inclusion bodies by centrifugation. The solubility test showed that some CoREST\textsuperscript{286-482} was expressed as inclusion bodies. Despite the low level of the expression, in order to yield a better soluble protein, cells were grown at 23 °C instead of 37 °C because inducing protein expression at low temperature can alleviate aggregation problem and also allow sufficient time for protein to fold into its native conformation. The low yield of the protein expression at 23 °C was circumvented by using TB (Terrific Broth) media to support higher cell density.

CoREST\textsuperscript{286-482} was then purified via two steps with Ni (II)-affinity chromatography and cation-exchange chromatography to homogeneity as judged by SDS-PAGE gel
analysis (Figure 16). The purified protein was sent for sequencing to Duke Proteomics Facility. Its identification as the C-terminal portion of CoREST was accomplished by proteolytic digest with trypsin followed by analysis of tandem mass spectroscopy (Figure 17).
Figure 15: Expression of CoREST\textsuperscript{286-482} in (a) BL21 (DE3) \textit{E. coli} strain, (b) BL21 Star (DE3) \textit{E. coli} strain, and (c) C43 (DE3) \textit{E. coli} strain. Pre: before IPTG induction; 1h, 2h, 4h: time points after IPTG induction; S: soluble fraction; IS: insoluble fraction. Red arrows indicate the location of CoREST\textsuperscript{286-482} on each gel.
Figure 16: Representative AKTA FPLC profiles and pictures of 12 % SDS-PAGE gel showing CoREST\textsuperscript{286-482} purification on (a) chelating sepharose Ni (II) affinity chromatography and (b) CM-sepharose cation exchange chromatography. A blue line represents UV absorbance at 280 nm as a function of retention volume (mL), and a green line represents a buffer gradient. Red arrows indicate the peaks where CoREST\textsuperscript{286-482} was eluted off from the column.
Figure 17: A graphic representation of amino acid sequence of CoREST\textsuperscript{286-482} identified by tandem mass spectrometry. The sequences in the blue box represent the ones corresponding to CoREST\textsuperscript{286-482}. The yellow regions represent the sequences identified and matched with original sequence of CoREST, and the green color represents the oxidized methionines.
2.4 Enzymatic Assay of LSD1 in the Presence of CoREST^{286-482}

The demethylation activity of LSD1 against several methylated peptides of varying lengths has been previously reported [79, 80], and showed that the first 21 amino acids of H3 including dimethylated K4 was sufficient for the detectable LSD1 activity (Figure 18). The first 30-amino acid substrate showed a similar enzymatic activity measured with the 21-amino acid substrate, whereas with the first 16-amino acid substrate or the much shorter substrate, the activity was hardly observed [80], indicating that the substrate recognition by LSD1 is not confined to the lysine 4 residue but rather LSD1 senses other modifications on neighboring residues for catalysis in a long peptide substrate. As such, we examined LSD1 enzymatic activity on the peptide consisting of the first 21 N-terminal amino acid residues of H3 with dimethylated K4 by utilizing a horseradish peroxidase (HRP) coupled assay (Scheme 3). The HRP converts a byproduct of LSD1 enzymatic activity, H_{2}O_{2}, to O_{2} while reducing amplex red into fluorescent resorufin, which can be monitored using a fluorimeter [81]. Then the kinetic parameters for LSD1 activity were evaluated by using the GraFit 6.0 software (Erithacus Software, West Sussex, UK); kinetic parameters and representative data are shown in Table 4 and Figure 19, respectively. Derived kinetic values are in reasonable agreement with previously reported values [79, 80]. The incubation of LSD1 with CoREST^{286-482} for 2 h at 4 °C increased the initial velocity of the catalytic activity of LSD1 roughly by 1.6-fold, while decreasing catalytic efficiency ($k_{cat}/K_{m}$). Thus, CoREST^{286-482} has little impact on the
catalytic efficiency of LSD1 towards peptide substrates. However, it is possible that the catalytic efficiency may be increased when nucleosomal substrates are used because CoREST is known to stimulate the catalytic activity of LSD1 toward nucleosomal substrates.

Figure 18: Sequence of the first N-terminal 30 amino acids of histone 3 (H3). The numbers above residues indicate the length of peptides.
Scheme 3: Enzymatic scheme for coupled assay to detect LSD1 activity. HRP (horseradish peroxidase) converts a byproduct of LSD1 enzymatic activity, H₂O₂, to O₂ while reducing amplex red into fluorescent resorufin.
Table 4: Kinetic parameters of the catalytic activity of LSD1 with the H3K4diMe peptide substrate (21 amino acids). 

<table>
<thead>
<tr>
<th></th>
<th>Initial Velocity (µM/s)</th>
<th>$K_M$ (µM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSD1 only</td>
<td>0.028 ± 0.002</td>
<td>2.300 ± 0.305</td>
<td>0.057 ± 0.003</td>
</tr>
<tr>
<td>LSD1-CoREST$^{286-482}$</td>
<td>0.046 ± 0.001</td>
<td>8.833 ± 1.559</td>
<td>0.092 ± 0.001</td>
</tr>
</tbody>
</table>

$^a$ 50 mM Tris buffer at pH 7.85 and 25 °C, 19.9 mM peptide substrate used.

Figure 19: Initial velocity curves of the catalytic activity of LSD1 using the first 21 amino acid residues of H3 with dimethylated K4 as a substrate. The assay was performed in 50 mM Tris buffer, pH 7.85 at 25 °C, and the final concentration of both proteins was 0.5 µM. Data were fitted to the Michaelis-Menten equation. The black curve represents the activity of LSD1, whereas the green curve represents the activity of LSD1 in the presence of CoREST$^{286-482}$.
2.5 Discussion and Future Directions

In this chapter, we have presented the expressions and purifications of LSD1 and CoREST\textsuperscript{286-482}. Both proteins were overexpressed in BL21 Star (DE3) \textit{E. Coli} strain with 0.5 mM IPTG induction and overnight growth under 23 °C. LSD1 was then purified via Ni (II)-affinity chromatography, gel-filtration chromatography, and anion-exchange chromatography, while CoREST\textsuperscript{286-482} was able to be purified via only two steps, Ni (II)-affinity chromatography and cation-exchange chromatography. The presence of CoREST\textsuperscript{286-482} increased the initial velocity of LSD1 enzymatic activity by almost two-fold but had little impact on catalytic efficiency of LSD1 towards peptide substrates.

We believe that kinetic parameters of LSD1 enzymatic activity towards nucleosomal substrate in the presence of CoREST\textsuperscript{286-482} may be changed given the role of CoREST\textsuperscript{286-482} in directing LSD1 demethylation activity towards nucleosome; thus, purification of nucleosomes and following kinetic studies are important future goals that our lab is pursuing. Also, kinetic parameters of LSD1 activity in the presence of CoREST containing all functional domains have not been reported yet. Kinetic parameters of LSD1 activity with a full-length CoREST are expected to be similar with the ones with truncated form of CoREST; nonetheless it is necessary to study kinetic parameters of LSD1 activity with a full-length of CoREST to confirm the role of N-terminal portion of CoREST on nucleosomes.
2.6 Experimental Section

Reagents and Materials. Media, antibiotics, and all other buffer reagents were purchased from Sigma, New England Biolabs, MP Biomedicals, and BD Biosciences. Chemically competent BL21 Star (DE3) *Escherichia coli* cells were purchased from Invitrogen. Chromatographic protein purifications were carried out on an AKTA FPLC (GE Healthcare), and the protein concentration was determined by UV absorbance spectroscopy.

Expression and Purification of LSD1. The gene encoding a truncated form of LSD1 (residues 151-852) was codon-optimized by Dr. Helena M. Gaweska, a previous lab member in the McCafferty group. The codon-optimized gene was used for expression and purification as previously described [81, 82] except that the gene was cloned into the pET15b vector instead of the pET 151-D/TOPO vector for expression. The resulting plasmid was transformed into chemically competent BL21 Star (DE3) *E. coli* cells, which were grown on a LB agar plate supplemented with ampicillin overnight at 37 °C. Streaks of the colonies on the plate were used to grow 6 L of cells in TB media with shaking (200 rpm) at 23 °C. When the cell density reached an OD_{600} of 0.6, 0.5 mM IPTG was added to the flasks to induce LSD1 expression. The cells were allowed to grow overnight and collected by centrifugation at 4225 × g. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, and 0.4 mM PMSF (pH 7.8). LSD1 was then purified via nickel-affinity chromatography,
HiPrep 26/60 Sephacryl S200 gel filtration chromatography (GE Life Sciences), and anion-exchange chromatography (Q-Sepharose Fast Flow, GE Life Sciences). The final concentration of LSD1 was determined by absorption spectroscopy at 458 nm [82] and the protein was stored at -20 °C in 80 % glycerol.

*Extraction of the Gene Encoding CoREST Residues* 286-482. Primers were designed to extract CoREST286-482 from CoREST53-482. Initial experiment started with amplifying the gene of interest encoding CoREST286-482 using PCR. Primers were designed to contain *Nde*I (CATATG) and *BamH*I (GGATCC) restriction sites at N- and C-terminals respectively to allow for facile ligation into the pET28b vector. The PCR reaction mix and conditions were as follows:

<table>
<thead>
<tr>
<th>PCR reaction mix</th>
<th>PCR condition (30 cycles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mM dNTP mix</td>
<td>Initial denaturation</td>
</tr>
<tr>
<td>CoREST53-482 template</td>
<td>95 °C</td>
</tr>
<tr>
<td>Forward primer</td>
<td>Denaturation</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>95 °C</td>
</tr>
<tr>
<td>10X pfu ultraHF reaction buffer</td>
<td>Annealing primers</td>
</tr>
<tr>
<td>pfu turbo polymerase</td>
<td>49-60 °C</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Elongation</td>
</tr>
<tr>
<td>Sterile water</td>
<td>72 °C</td>
</tr>
<tr>
<td>Total</td>
<td>Final elongation</td>
</tr>
<tr>
<td></td>
<td>72 °C</td>
</tr>
<tr>
<td></td>
<td>Store</td>
</tr>
<tr>
<td></td>
<td>4 °C</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2 µL</td>
<td></td>
</tr>
<tr>
<td>1 µL</td>
<td>5 min</td>
</tr>
<tr>
<td>2 µL</td>
<td>1 min</td>
</tr>
<tr>
<td>2 µL</td>
<td>1 min</td>
</tr>
<tr>
<td>5 µL</td>
<td>1:40 min</td>
</tr>
<tr>
<td>1 µL</td>
<td>10 min</td>
</tr>
<tr>
<td>1 µL</td>
<td>∞</td>
</tr>
<tr>
<td>36 µL</td>
<td></td>
</tr>
<tr>
<td>50 µL</td>
<td></td>
</tr>
</tbody>
</table>

The PCR product and pET28b vector were then digested with *Nde*I and *BamH*I enzymes, and ligated at 16°C overnight together.
Expression and Purification of CoREST\textsuperscript{286-482}. A pET28b vector containing the gene of interest encoding CoREST\textsuperscript{286-482}, whose sequence was verified, was used for expression and purification. The vector was transformed into electrocompetent BL21 Star (DE3) E. coli cells, which were used to grow for purification. Streaks of the colonies on a LB agar containing kanamycin were used to grow 6 L of bacteria. The cells were grown as described above for LSD1 purification. CoREST\textsuperscript{286-482} was purified via nickel-affinity chromatography with a linear gradient from 50 to 500 mM imidazole in 50 mM sodium phosphate and 300 mM NaCl (pH 7.4), which was followed by purification via cation-exchange chromatography (CM-Sepharose Fast Flow, Sigma) with a linear gradient from 100 to 800 mM NaCl in 50 mM sodium phosphate (pH 7.4). The concentration of CoREST\textsuperscript{286-482} was measured spectrophotometrically using the extinction coefficient of 16,950 cm\textsuperscript{-1}M\textsuperscript{-1} at 280 nm.

Enzymatic Assay. Steady-state kinetic assays to determine kinetic parameters for LSD1 activity on the dimethylated H3K4 21-mer peptide substrate with and without CoREST\textsuperscript{286-482} were performed by employing a fluorescence assay as previously described [81, 82]. All assays were performed at 25 °C, and the product was monitored by a fluorescence plate reader (Molecular Devices SpectraMax Germini EM) at 560 nm excitation and 590 nm emission wavelength. The experiment was performed in duplicate.
3. Characterization of the Binding Interaction between LSD1 and CoREST\textsuperscript{286-482}

3.1 Background

Chromatin is dynamically regulated by various post-translational modifications, which have been the major focus of studies in the epigenetic field to date. Histone modifications regulate gene transcriptions and chromatin structure in a residue-specific manner by the complexity of proteins involved. Mostly histone-modifying enzymes do not work alone but interact with other biological entities such as DNA or other proteins to perform their functions. Hence, in order to fully understand intricacies of gene regulation, it is necessary to elucidate the mechanisms by which protein complexes function in their native transcriptional complexes.

LSD1 has been found in numerous transcriptional complexes including CoREST [54, 65], HDACs [70], metastasis tumor antigen (MTA) [58], androgen receptor (AR) [59], and estrogen receptor (ER) [60]; thus analysis of these macromolecular complexes can lead to significant insights into the function of LSD1. Particularly, our group has worked on the association of LSD1 with CoREST in the regulation of LSD1 enzymatic activity. CoREST is known to endow LSD1 with the ability to associate with and demethylate nucleosomal substrates. Recent studies by other groups have shown that the C-terminal fragment of CoREST (residues 286-482), including the linker region and SANT2 domain, is essential for LSD1-catalyzed demethylation of H3K4 within nucleosomes and
physically interacts with LSD1 [52, 65, 66, 71]. However, little is known about the biophysical nature of the interaction and the molecular details of the interaction sites between LSD1 and CoREST. In this chapter, on the basis of the importance of the functional region of CoREST (CoREST$^{286-482}$) in stimulating LSD1 activity, we describe thermodynamic characterization of the binding interaction between LSD1 and CoREST$^{286-482}$ to provide a fundamental understanding of their interaction.

### 3.1.1 Protein-Protein Interactions

Protein-protein interactions are essential for life by mediating an enormous number of cellular processes, including DNA replication, transcription and translation, signal transduction, and metabolic pathways [83]. Therefore, identifying and characterizing protein-protein interactions and assemblies are major themes for understanding the mechanisms of various biological processes on a molecular level. Despite the fact that protein interactions are considerably diverse, protein-protein interactions can be classified by common properties they share into several types depending on their binding strength (permanent versus transient), specificity (specific versus non-specific), and the similarity between interacting subunits (homo- versus hetero-oligomeric) [84, 85]. However, often the boundaries between these classes are not distinct and rather protein interactions can be considered continuous interactions.
Various histone-modifying enzymes also perform their functions through protein-protein interactions by forming large complexes in a highly specific manner, which allows them to gain several advantages such as more efficient recognition of substrates and the ability to modulate the activity or function of a protein by altering its partner binding composition. LSD1 and CoREST are such core proteins in transcriptional complexes in which the interaction of LSD1 with CoREST facilitates the demethylation activity of LSD1 in the context of nucleosomes. To understand the mechanisms of protein-protein interaction at the molecular level, various different biophysical experimental techniques have been developed to determine the intrinsic thermodynamics of binding, including isothermal titration calorimetry (ITC) that directly measures $\Delta H$, fluorescence polarization (FP) and surface plasmon resonance that detect the increase in size of protein-protein interactions, fluorescence resonance energy transfer (FRET) that detects a change in fluorescence spectrum when a fluorescent protein binds to its binding protein [86].

### 3.1.2 Overview of Isothermal Titration Calorimetry and Thermodynamics

Isothermal titration calorimetry (ITC) has been used extensively to study protein-protein and protein-ligand binding interactions [87, 88, 89]. When two proteins or protein and ligand bind, ITC measures the heat change produced by binding and the
changes in the thermodynamic potentials (ΔG, ΔH, ΔS); thus, ITC is the only technique capable of determining both the binding affinity and two thermodynamic terms directly - ΔH, and ΔS- in a single experiment and measurements are also conducted quickly within a couple of hours on average [87, 89]. Other methods to investigate protein-protein or ligand interactions such as SPR, fluorescence polarization, and equilibrium dialysis require proteins to be either immobilized to the surface or radioactively or fluorescently labeled, which may interfere with the binding interaction. In a typical ITC experiment, it is not necessary to have proteins with functional groups to monitor the binding process; hence, the binding interaction can be observed using native proteins without any of their modifications. Also ITC directly measures sub-millimolar to nanomolar binding constants (10^2 to 10^9 M^-1) [87, 88, 89, 90].

An ITC instrument consists of two lollipop shaped cells, a sample cell and a reference cell, which are composed of highly efficient thermal conducting materials such as gold surrounded by an adiabatic jacket [90]. Sensitive thermoelectric devices measure the temperature difference between two cells and between the cells and jacket. The temperature difference between the cells and the jacket is continuously monitored to maintain an identical temperature, which is used as a baseline [87, 90]. A feedback control system monitors the difference in temperature (ΔT) between two cells. In an ITC experiment, one of the proteins is placed in the sample cell and its binding partner in a syringe. The reference cell contains buffer or water without protein samples. The syringe
also serves to stir the solution in the sample cell to achieve proper mixing. At specific
time intervals, a small volume of the protein in the syringe is injected into the cell, giving
rise to characteristic titration heat effects; heat is taken up or released depending on
whether the binding interaction is endothermic or exothermic. For exothermic reactions,
the temperature in the sample cell will be increased, and the feedback system will be
deactivated to maintain constant temperature between two cells. For endothermic
reactions, the feedback system will increase the power to the sample cell to maintain the
temperature. Therefore, the change in power required to maintain a constant
temperature is recorded as a series of injections is made (Figure 20). As the course of
injection is completed, the binding sites on protein in the sample cell are gradually
saturated and only the heat of dilution of the other protein sample in the syringe
(background) is observed.
Figure 20: Schematic representation of isothermal titration calorimetry (ITC) instrument. The sample cell is filled with the protein (blue) and the injection syringe, which stirs the solution to assure proper mixing, is filled with the partner protein (pink). At specific time intervals, as the protein in the syringe is injected into the cell, heat is generated or absorbed. The feedback system will either increase or decrease power to the sample cell to maintain equal temperature with the reference cell. The heat per unit time (µcal/sec) supplied to the sample cell is monitored in the ITC experiment (Adapted from [90]).
The ITC signal is dependent on several variables such as concentrations and conditions of samples, volume of samples, and binding affinity, so care must be taken in all aspects of the experiment. Initial consideration of the experiment is typically made on the preparation of appropriate concentration of the interacting molecules so that the heat associated with the given binding is detectable and yields a curved thermogram. The binding curve is characterized by the unitless value, sometimes denoted as $c$, which is the product of association constant ($K_a$), the concentration of the protein in the sample cell ([M] indicating the concentration of macromolecules in the sample cell), and the stoichiometry of the reaction ($n$) [90]. For an accurate determination of the binding constant, the $c$ value should be in the range between 1 and 1000, but more ideally between 10 and 100. As the $c$ value increases, a binding curve has a rectangular shape with a sharp transition, whereas low $c$ values lose the characteristic sigmoidal shape showing broad transitions [87, 88, 90]. The choice of buffer is also a critical consideration to take into because the heat of ionization of the buffer can affect the binding enthalpy [87, 90]. The use of a buffer with a low ionization enthalpy such as sodium phosphate ($\Delta H_{\text{ion}} = 0.9$ kcal/mol at 25 °C) is desirable [91]. Additionally both molecules in the sample cell and the syringe should be extensively dialyzed in the same buffer to achieve complete chemical equilibrium so that the impact of buffer mismatch is minimized. Prior to loading the sample cell and the syringe, both molecules should be degassed to remove air bubbles that may cause irregular background noise. The sample cell and syringe
should be afterwards filled with the degassed samples without introducing any air bubbles.

A binding interaction involves \( n \) binding sites, typically indicated as stoichiometry of the binding interaction, which is described by the reactions [92]:

\[
A + nB \xleftrightarrow{K} AB_n
\]

where \( A \) and \( B \) represents two binding proteins in the sample cell and syringe respectively, \( n \) the number of binding sites, and \( K \) the equilibrium binding constant. The heat associated during each injection (\( \Delta Q \)) is proportional to the change in concentration of the complex (\( \Delta [AB] \)), the number of binding sites (\( n \)), the molar enthalpy of binding (\( \Delta H^\circ \)), and the volume of the sample cell (\( V_o \)). The derivation of the equation for the enthalpy of the binding for the formation of a 1:1 complex shown below is adapted from [92, 93]:

\[
\Delta Q = n \Delta [AB] \Delta H^\circ V_o
\]  
(1)

The incremental heat associated as a function of the protein in syringe (B) is expressed as:

\[
\frac{dQ}{d[B]_t} = n \frac{d[AB] \Delta H^\circ V_o}{d[B]_t}
\]  
(2)

Since the bound and free concentrations of both proteins are unknown, the binding constant can be expressed as a function of the total protein concentrations, \([A]_t = [A] + [AB] \) and \([B]_t = [B] + n[AB] \), respectively, where \( t \) represents total:
Isolating $[AB]$ from Eq. (3) produces a quadratic equation:

$$[AB] = \frac{1}{2} \left( \left( [A]_t + \frac{[B]_t}{n} + \frac{1}{nK} \right) - \sqrt{\left( [A]_t + \frac{[B]_t}{n} + \frac{1}{nK} \right)^2 - 4[AB]_k/n} \right)$$  \hspace{1cm} (4)

The incremental change in $[AB]$ with respect to addition of the second protein $B$ in the syringe can be calculated as

$$\frac{d[AB]}{d[B]_t} = \frac{1}{2} \left[ \frac{[B]_t}{n[A]_t} - 1 + \frac{1}{nK[A]_t} \right]$$ \hspace{1cm} (5)

Substituting Eq. (5) into (2) results in the following expression, Eq. (6), which is used to fit the binding curve obtained from the ITC experiment:

$$\frac{dQ}{d[B]_t} = \frac{1}{2} \Delta H^o V_o \left[ 1 - \frac{[B]_t}{[A]_t} - n + \frac{1}{K[A]_t} \right]$$ \hspace{1cm} (6)

Therefore, Eq. (6) allows us to evaluate the relationship of the enthalpy associated with the binding interaction as a function of the concentration of the second protein $B$ in the
syringe for the formation of a 1:1 complex. Knowing the total concentrations of protein A and B, fitting the binding curve to Eq. (6) yields the stoichiometry (n), accurate equilibrium binding constant \( K_{eq} \), and the enthalpy of the binding \( \Delta H^\circ \). Then the entropy of the binding \( \Delta S^\circ \) and the Gibbs free energy of binding \( \Delta G^\circ \) can be calculated from the known thermodynamic relationships, Eq. (7):

\[
\Delta G^\circ = \Delta H^\circ - T \Delta S^\circ = -RT \ln K_{eq}
\]

where R is the universal gas constant and T is the absolute temperature.

The information obtained through thermodynamic data of the binding interaction reflects various types of forces that drive the binding interaction, including enthalpic contributions of bond formation and entropic effects such as loss of degrees of freedom upon binding that contribute to the total binding free energy. The Gibbs free energy of binding \( \Delta G^\circ \) determines the stability of the formation of a complex [94]. The binding interactions between biological molecules most likely occur only when it is coupled with a negative value of free energy of binding that indicates a spontaneous and thermodynamically favorable process. The understanding of this free energy of binding originates from the contributions of the enthalpy and entropy of the binding interactions. The enthalpic contribution to the free energy reflects the changes in non-covalent bonds during the binding event. It must be the results of the loss of protein-solvent hydrogen bonds and Van der Waals interactions, formation of protein-protein (or ligand) bonds, salt bridges and Van der Waals contacts, and solvent reorganization.
near protein surfaces [94]. These interactions may produce either favorable or unfavorable contributions. The changes in the binding entropy can be affected by contributions associated with the proteins and solvent. It is well known that the entropy of binding observed most likely can be ascribed to the contributions from conformational changes upon binding; a linear correlation of the change in conformational entropy with the change in total entropy of binding has been studied [95]. Also burial of hydrophobic surface upon binding and the subsequent release of water to solvent can result in significant change in entropy [96].

3.2 Calorimetric Titration of CoREST\textsuperscript{286-482} against LSD1

Previous studies have demonstrated the interaction between CoREST\textsuperscript{286-482} and LSD1 [52, 65, 71], but little biophysical information about the interaction exists. Initial experiments to study this binding interaction were performed using a MicroCal VP-ITC microcalorimeter at 25 °C. Because of the lack of information of the binding affinity between LSD1 and CoREST until we characterized it, the determination of the $c$ value for LSD1 and CoREST binding interaction was estimated based on the information from previous studies that LSD1 and CoREST complex was able to survive extensive washing step [52], which suggests that their binding is most likely a tight interaction. Thus, assuming that the binding interaction between LSD1 and CoREST has a high affinity with approximately $10^7$–$10^8$ M of the binding constant ($K_a$), aliquots (1.5 mL) of LSD1 at 3
μM was chosen as a starting point, which would produce 30–300 of the ε value. This predicted affinity could require that approximately 10 times higher concentration of CoREST (30 μM) needed to be used in the syringe.

3.2.1 Analysis of the Thermodynamic Parameters for the Interaction between LSD1 and CoREST<sub>286-482</sub>

Calorimetric data were analyzed using the Origin 5.0 software. In order to correct data for dilution, average heats observed in the last 10–15 injections were subtracted from binding data. The data was then fit to a one site-binding model (Eq. (6)) to give stoichiometry (n), association constant (K<sub>a</sub>), and change in binding enthalpy (ΔH). Free energy of association (ΔG) and change in entropy (TΔS) were calculated from the known thermodynamic relationships.

ITC experiments between LSD1 and CoREST<sub>286-482</sub> yielded a thermogram shown in Figure 21, displaying raw power output versus time of CoREST<sub>286-482</sub> injection into the cell containing LSD1 (top panel) and the corresponding binding isotherm where the enthalpy per mole of CoREST<sub>286-482</sub> as a function of LSD1 is plotted (bottom panel). Each injection of CoREST<sub>286-482</sub> gave rise to exothermic heats of binding, and each peak became smaller as the binding sites on LSD1 were saturated with CoREST<sub>286-482</sub>. Overall, we observed a small amount of heats produced after each injection, which is most likely because low concentrations of proteins were used for titrations to generate appropriately
measurable binding curve for high-affinity binding interaction. However, such amount of heat is useful for most protein-protein interactions of biological importance. Thermodynamic parameters for the interaction are summarized in Table 5. The dissociation constant ($K_d$) for LSD1-CoREST$^{286-482}$ interaction was determined to be $15.9 \pm 2.07$ nM, indicating a tight binding between LSD1 and CoREST$^{286-482}$. The thermodynamic parameters of binding of LSD1 to CoREST$^{286-482}$ showed overall favorable enthalpic ($-21.3 \pm 0.19$ kcal/mol) and unfavorable entropic ($-10.7 \pm 1.40$ kcal/mol) contributions near room temperature, which suggests that the binding between LSD1 and CoREST$^{286-482}$ is thermodynamically favorable with the free energy of binding of $-10.6 \pm 1.38$ kcal/mol. This observation of the favorable enthalpic change and unfavorable entropic change is commonly observed in biomolecular interactions. An increase in enthalpy by tighter binding almost certainly affects the entropy by the restriction of motion of the interacting molecules, which is referred to as enthalpy-entropy compensation. The high binding affinity is likely attributed to various non-covalent interactions between LSD1 and CoREST$^{286-482}$ upon binding. Based on the known crystal structure of the LSD1-CoREST$^{286-482}$ complex, several candidates for the formation of hydrogen bonds and salt bridges at the interface are found; interactions between residues of F315, D320, D339, K353, N356, K360, R371 of CoREST$^{286-482}$ and Q419, K421, K424, Q438, D495, E505, E512 of LSD1 may contribute directly to the enthalpic gain. The binding interface is also populated with several nonpolar residues, which are
presumably involved in hydrophobic interaction. In addition to those interactions located between helical regions of two proteins, the N-terminal region of CoREST\textsuperscript{286-482} wraps around the bottom of the anti-parallel helices of the LSD1 tower domain\textsuperscript{418-522}, contributing to the overall strength of the binding (Figure 22). The formation of various covalent bonds consequently reflects the unfavorable change in entropy; the effect of the restricted motions of residues upon binding is greater than the one of the release of water molecule from the binding interface to the bulk solution. The stoichiometry is close to 1, suggesting a simple 1:1 complex.
Table 5: Thermodynamic parameters for binding of LSD1 to CoREST<sub>286-482</sub> at 25 °C. <sup>b</sup>

<table>
<thead>
<tr>
<th></th>
<th>( K_d (\text{nM}) )</th>
<th>( \Delta H ) (kcal/mol)</th>
<th>( T \Delta S ) (kcal/mol)</th>
<th>( \Delta G ) (kcal/mol)</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSD1-CoREST&lt;sub&gt;286-482&lt;/sub&gt;</td>
<td>15.9 ± 2.07</td>
<td>-21.3 ± 0.19</td>
<td>-10.7 ± 1.40</td>
<td>-10.6 ± 1.38</td>
<td>1.14 ± 0.00</td>
</tr>
</tbody>
</table>

<sup>b</sup> 50 mM sodium phosphate and 1 mM DTT, pH 7.4
Figure 22: (a) Surface representation of the complex between LSD1 and CoREST. Two main domains responsible for the binding interaction are shown: LSD1 tower domain (residues 418-522, blue) and CoREST linker region (residues 293-380, red). The dark blue and dark red surfaces represent the residues involved in the binding interaction. (b) Possible residues involved in the hydrogen bond and salt bridges are represented by dashed lines. (c) Nonpolar residues are populated at the binding interface. Nonpolar residues of each protein are represented as red and blue surface diagrams, respectively. All structural figures were generated using PyMoL (PDB file: 2IW5).
3.2.2 Surface Plasmon Resonance (SPR) Measurement

The tight binding interaction between LSD1 and CoREST\textsuperscript{286-482} was confirmed by surface plasmon resonance measurements provided at the Duke Human Vaccine Institute Biomolecular Interaction Analysis Facility under the direction of Dr. S. Munir Alam. As shown in Figure 23, the increase in response units (RU) over time represents the amount of LSD1 bound to CoREST\textsuperscript{286-482} that is proportional to the association rate constant ($k_{on}$, $5.63 \times 10^4 \pm 1.93 \times 10^3$ M\textsuperscript{-1}s\textsuperscript{-1}) of the binding interaction. After association, buffer was injected to dissociate the bound LSD1, which results in the decrease in RU over time and yields a dissociation rate constant ($k_{off}$, $7.78 \times 10^{-5} \pm 1.27 \times 10^{-5}$ s\textsuperscript{-1}). Some noises observed in the dissociation phase after 300 seconds, which might affect uncertainty in the determination of the dissociation rate constant ($k_{off}$), were removed. These rate constant values suggest a dissociation constant ($K_d$) of $1.4 \pm 6.61$ nM, slightly lower than the dissociation constant ($15.9 \pm 2.07$ nM) measured using isothermal titration calorimetry. The small but significant differences between binding constants obtained from two techniques reflect a small degree of enhanced binding during SPR measurement. This difference in binding constants may also be attributed to a slow off rate due to a columning phenomenon frequently observed in SPR experiments. Nonetheless, these techniques independently confirm that the interaction between LSD1 and CoREST \textsuperscript{286-482} is tight, and proceeds with a 1:1 stoichiometry.
Figure 23: Surface plasmon resonance (SPR) measurement curves obtained during and after injection of LSD1 on chip surfaces with immobilized CoREST\textsuperscript{286-482}. Sensograms show the association of LSD1 to and dissociation from CoREST\textsuperscript{286-482}. The association rate constant ($k_{on}$, $5.63 \times 10^4 \pm 1.93 \times 10^3 \text{M}^{-1}\text{s}^{-1}$) and dissociation rate constant ($k_{off}$, $7.78 \times 10^{-5} \pm 1.27 \times 10^{-5} \text{s}^{-1}$) were used to calculate a dissociation binding constant ($K_d$) of $1.4 \pm 6.61$ nM.
3.3 Calorimetric Titration of CoREST\textsuperscript{286-482} against LSD1 in Different Buffers

The observed calorimetric enthalpy may not solely arise from the binding interaction, since several other events contribute to the heats of binding. One such event is proton transfer between a protein-protein complex and buffer due to a shift in protein $pK_a$ on complex formation and buffer ionization; thus, the heat effect of protonation/deprotonation contributes to the overall heats of binding [88, 92]. Repeating the experiment at the same pH in buffers with different ionization enthalpies allows for the determination of the number of proton released/accepted by buffer solutions. Such proton transfer can be parsed by the following expression:

$$\Delta H_{\text{cal}} = \Delta H_{\text{intrinsic}} + n_H \Delta H_{\text{ion}}$$

where $\Delta H_{\text{cal}}$ is the sum of calorimetric enthalpy, $\Delta H_{\text{intrinsic}}$ is the enthalpy of binding in absence of protonation effects, $n_H$ is the number of protons transferred during binding, and $\Delta H_{\text{ion}}$ is the enthalpy of buffer ionization. If $n_H$ is not zero, then the binding interaction is coupled to protonation/deprotonation. The slope of the plot of $\Delta H_{\text{cal}}$ versus $\Delta H_{\text{ion}}$ gives the information about a number of protons transferred; a negative slope of the plot indicates a net release of protons from protein to the buffer, while a positive slope indicates a net uptake of protons to protein from the buffer [92].

In order to assess the contribution of heat of ionization upon binding, titrations were performed in five buffers with different heats of ionization (sodium phosphate
with $\Delta H_{\text{ion}} = 0.9$ kcal/mol, PIPES with $\Delta H_{\text{ion}} = 2.7$ kcal/mol, HEPES with $\Delta H_{\text{ion}} = 5.7$ kcal/mol, ACES with $\Delta H_{\text{ion}} = 7.5$ kcal/mol, and Tris with $\Delta H_{\text{ion}} = 11.4$ kcal/mol) [91]. Table 6 shows thermodynamic parameters for the binding between LSD1 and CoREST$^{286-482}$ in various buffers at pH 7.4. A plot of $\Delta H_{\text{cal}}$ versus $\Delta H_{\text{ion}}$ (Figure 24) yielded a straight line with a positive slope, indicating that $0.83 \pm 0.09$ protons are absorbed by the LSD1-CoREST$^{286-482}$ complex from the buffer at pH 7.4. Presumably this protonation event involves either lysine or arginine in the LSD1-CoREST$^{286-482}$ complex near neutral pH. $\Delta H_{\text{intrinsic}}$ (i.e. when $\Delta H_{\text{ion}} = 0$) was determined to be $-19.8 \pm 0.58$ kcal/mol, suggesting that the binding interaction is intrinsically exothermic and enthalpy-driven at room temperature.
Table 6: Thermodynamic parameters for binding of LSD1 to CoREST<sup>286-482</sup> in various buffers at 25 °C.<sup>c</sup>

<table>
<thead>
<tr>
<th>Buffer</th>
<th>$K_d$(nM)</th>
<th>$\Delta H_{cal}$(kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>$T\Delta S$(kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>$\Delta G$(kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphate</td>
<td>17.3 ± 3.98</td>
<td>-19.5 ± 0.26</td>
<td>-8.93 ± 2.06</td>
<td>-10.6 ± 2.44</td>
<td>1.11 ± 0.01</td>
</tr>
<tr>
<td>PIPES</td>
<td>5.84 ± 1.63</td>
<td>-17.4 ± 0.32</td>
<td>-6.19 ± 1.75</td>
<td>-11.2 ± 3.17</td>
<td>0.91 ± 0.01</td>
</tr>
<tr>
<td>HEPES</td>
<td>11.7 ± 4.20</td>
<td>-15.2 ± 0.38</td>
<td>-4.33 ± 1.57</td>
<td>-10.9 ± 3.93</td>
<td>1.01 ± 0.01</td>
</tr>
<tr>
<td>ACES</td>
<td>21.2 ± 7.82</td>
<td>-12.6 ± 0.41</td>
<td>-2.48 ± 0.93</td>
<td>-10.1 ± 3.77</td>
<td>0.93 ± 0.02</td>
</tr>
<tr>
<td>Tris</td>
<td>8.54 ± 3.50</td>
<td>-10.9 ± 0.28</td>
<td>-0.12 ± 0.05</td>
<td>-10.8 ± 4.45</td>
<td>0.92 ± 0.01</td>
</tr>
</tbody>
</table>

<sup>c</sup> 50 mM sodium phosphate (PIPES, HEPES, ACES, Tris) and 1 mM DTT, pH 7.4

Figure 24: Plot of the calorimetric enthalpy ($\Delta H_{cal}$) obtained from binding of LSD1 to CoREST<sup>286-482</sup> as a function of the enthalpy of buffer ionization ($\Delta H_{ion}$) at pH 7.4. The slope of the plot indicates the number of protons transferred between the complex and the buffer upon binding, and the y-intercept indicates the intrinsic $\Delta H_{cal}$ without a protonation effect. At pH 7.4, 0.83 ± 0.09 protons are absorbed to the complex from the buffer.
3.4 Calorimetric Titration of CoREST\textsuperscript{286-482} against LSD1 in Different Temperatures

In addition to the contribution of a proton transfer to the binding interaction, another important contribution to binding enthalpy comes from solvent reorganization upon binding [88, 92]. The thermodynamic parameter most closely associated with solvent reorganization is the change in the molar heat capacity of solution ($\Delta C_p$), the temperature derivative of the enthalpy:

$$C_p = \frac{\partial \Delta H}{\partial T}$$

Thus, the change in heat capacity was obtained by measuring $\Delta H$ over the temperature range from 15 °C to 35 °C in 50 mM sodium phosphate buffer with 1 mM DTT (pH 7.4) (Table 7). At low temperature (10 °C), the signals generated upon binding were too small to fit to a binding model. Titrations at a high temperature (55 °C) yielded an irregular binding curve, suggesting protein denaturation. A plot of $\Delta H$ versus temperature yielded a straight line with a negative slope ($\Delta C_p$), -0.80 ± 0.01 kcal/mol·K (Figure 25). The major contributions to $\Delta C_p$ originate from hydrophobic, conformational, and vibrational effects [96], but hydrophobic interactions are dominant contributors, which is supported by a strong correlation between $\Delta C_p$ and the amount of buried surface area on forming a complex [88, 94]. Accordingly, for the biomolecular binding reactions, $\Delta C_p$ is almost always negative. The crystal structure of the LSD1-CoREST\textsuperscript{286-482}
complex shows that the binding interface between LSD1 and CoREST^{286-482} is mostly populated by nonpolar residues that are presumably involved in hydrophobic interactions during complex formation (Figure 22-(c)).

Also the effect of NaCl on the change in heat capacity $\Delta C_p$ was studied over the same temperature range, using varying concentrations of NaCl (10 mM NaCl and 100 mM NaCl) in 50 mM sodium phosphate buffer with 1 mM DTT (pH 7.4) (Table 8 and Table 9). As the concentration of NaCl in buffer increased, larger binding heats ($\Delta H$) were observed, but $\Delta C_p$ values did not vary significantly (Figure 26). The $\Delta C_p$ values with the addition of 10 mM NaCl and 100 mM NaCl were determined to be -1.04 ± 0.01 kcal/mol·K and -0.84 ± 0.01 kcal/mol·K, respectively. The charge shielding effect by the addition of NaCl seems not to affect the binding affinity, but the heat of ionization of NaCl presumably affects the total heats of binding, resulting the larger $\Delta H$ observed.
Table 7: Thermodynamic parameters for binding of LSD1 to CoREST\textsuperscript{286-482} at different temperatures.\textsuperscript{d}

<table>
<thead>
<tr>
<th>T (°C)</th>
<th>$K_d$ (nM)</th>
<th>$\Delta H$ (kcal mol$^{-1}$)</th>
<th>$T\Delta S$ (kcal mol$^{-1}$)</th>
<th>$\Delta G$ (kcal mol$^{-1}$)</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>21.7 ± 10.9</td>
<td>-4.24 ± 0.17</td>
<td>5.86 ± 2.95</td>
<td>-10.1 ± 5.07</td>
<td>0.99 ± 0.02</td>
</tr>
<tr>
<td>25</td>
<td>3.18 ± 2.54</td>
<td>-12.1 ± 0.39</td>
<td>-0.49 ± 0.39</td>
<td>-11.6 ± 9.27</td>
<td>1.01 ± 0.02</td>
</tr>
<tr>
<td>30</td>
<td>3.70 ± 1.74</td>
<td>-16.1 ± 2.27</td>
<td>-4.42 ± 2.08</td>
<td>-11.7 ± 5.50</td>
<td>1.03 ± 0.01</td>
</tr>
<tr>
<td>35</td>
<td>8.03 ± 4.10</td>
<td>-20.2 ± 0.57</td>
<td>-8.79 ± 4.49</td>
<td>-11.4 ± 5.82</td>
<td>0.97 ± 0.02</td>
</tr>
</tbody>
</table>

\textsuperscript{d} 50 mM sodium phosphate and 1 mM DTT, pH 7.4

Figure 25: Temperature dependence of the binding enthalpy ($\Delta H_{cal}$) change for the interaction between LSD1 and CoREST\textsuperscript{286-482}. The slope of the plot yields the binding heat capacity change ($\Delta C_p$), which is equal to $-0.80 ± 0.01$ kcal/mol·K.
Table 8: Thermodynamic parameters for binding of LSD1 to CoREST<sup>286-482</sup> at different temperatures with the addition of 10 mM NaCl in buffer.<sup>e</sup>

<table>
<thead>
<tr>
<th>T (°C)</th>
<th>$K_d$ (nM)</th>
<th>$\Delta H$ (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>$T\Delta S$ (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>$\Delta G$ (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>$23.8 \pm 11.1$</td>
<td>$-5.07 \pm 0.19$</td>
<td>$4.97 \pm 2.33$</td>
<td>$-10.0 \pm 4.68$</td>
<td>$1.09 \pm 0.02$</td>
</tr>
<tr>
<td>25</td>
<td>$18.2 \pm 6.73$</td>
<td>$-15.4 \pm 0.41$</td>
<td>$-4.84 \pm 1.80$</td>
<td>$-10.6 \pm 3.95$</td>
<td>$0.95 \pm 0.02$</td>
</tr>
<tr>
<td>30</td>
<td>$15.5 \pm 2.48$</td>
<td>$-19.9 \pm 0.23$</td>
<td>$-9.07 \pm 1.43$</td>
<td>$-10.8 \pm 1.71$</td>
<td>$1.02 \pm 0.01$</td>
</tr>
<tr>
<td>35</td>
<td>$12.2 \pm 2.80$</td>
<td>$-26.3 \pm 0.40$</td>
<td>$-15.1 \pm 3.48$</td>
<td>$-11.2 \pm 2.58$</td>
<td>$1.01 \pm 0.01$</td>
</tr>
</tbody>
</table>

<sup>e</sup> 50 mM sodium phosphate, 10 mM NaCl, and 1 mM DTT, pH 7.4

Table 9: Thermodynamic parameters for binding of LSD1 to CoREST<sup>286-482</sup> at different temperatures with the addition of 100 mM NaCl in buffer.<sup>f</sup>

<table>
<thead>
<tr>
<th>T (°C)</th>
<th>$K_d$ (nM)</th>
<th>$\Delta H$ (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>$T\Delta S$ (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>$\Delta G$ (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>$10.5 \pm 3.78$</td>
<td>$-19.7 \pm 0.45$</td>
<td>$-9.20 \pm 3.32$</td>
<td>$-10.5 \pm 3.78$</td>
<td>$1.02 \pm 0.01$</td>
</tr>
<tr>
<td>25</td>
<td>$9.13 \pm 3.47$</td>
<td>$-25.6 \pm 0.58$</td>
<td>$-14.6 \pm 5.56$</td>
<td>$-11.0 \pm 4.18$</td>
<td>$0.98 \pm 0.01$</td>
</tr>
<tr>
<td>30</td>
<td>$4.78 \pm 3.30$</td>
<td>$-31.4 \pm 0.10$</td>
<td>$-19.8 \pm 13.7$</td>
<td>$-11.6 \pm 8.00$</td>
<td>$1.00 \pm 0.01$</td>
</tr>
<tr>
<td>35</td>
<td>$4.07 \pm 1.18$</td>
<td>$-36.5 \pm 0.45$</td>
<td>$-24.7 \pm 7.17$</td>
<td>$-11.8 \pm 3.42$</td>
<td>$1.08 \pm 0.01$</td>
</tr>
</tbody>
</table>

<sup>f</sup> 50 mM sodium phosphate, 100 mM NaCl, and 1 mM DTT, pH 7.4
Figure 26: Temperature dependence of the binding enthalpy change ($\Delta H_{\text{cal}}$) for the interaction between LSD1 and CoREST\textsuperscript{296-482} with different NaCl concentrations in buffers at pH 7.4. Blue diamond marks (•) indicate the enthalpy change with no NaCl in buffer, pink square (□) with 10 mM NaCl, and green circle (○) with 100 mM NaCl. With addition of buffer additive, NaCl, larger binding heats ($\Delta H_{\text{cal}}$) were observed, but the $\Delta C_P$ values were not changed significantly. The $\Delta C_P$ values with the addition of 10 mM NaCl and 100 mM NaCl were determined to be -1.04 ± 0.04 kcal/mol·K and -0.84 ± 0.1 kcal/mol·K, respectively.
3.4.1 Calculation of the Solvent Accessible Surface Area (SASA)

As addressed above, the largest contribution to the $\Delta C_p$ for the binding interaction arises from dehydration of protein surfaces with the burial of some nonpolar and polar surface areas upon binding [94, 97]. Thus, the calculation of the change in solvent-accessible surface area ($\Delta ASA$) between free proteins and a complex allows us to estimate the empirical heat capacity change by the following equation [94]:

$$\Delta C_p = 0.36 \text{ cal/mol} \cdot \text{K} \cdot \text{Å}^2 (\Delta ASA_{\text{nonpolar}}) - 0.25 \text{ cal/mol} \cdot \text{K} \cdot \text{Å}^2 (\Delta ASA_{\text{polar}})$$

where $\Delta ASA_{\text{nonpolar}}$ is the change in nonpolar accessible surface area, $\Delta ASA_{\text{polar}}$ is the change in polar accessible surface area, and 0.36 and -0.25 are empirical constants for nonpolar and polar surface area, respectively. Here, the solvent-accessible surface area is defined as the surface traced out by the center of a solvent probe (mostly water with a radius of 1.4 Å) as it moves over the surface of the protein [94]. The calculation of $\Delta ASA_{\text{nonpolar}}$ and $\Delta ASA_{\text{polar}}$ was performed using GetArea [98]. The values of the changes in polar, nonpolar, and total accessible surface area are summarized in Table 10.

Accordingly, the value of the change in heat capacity predicted by the equation above is -1.16 kcal/mol·K, which is somewhat larger than the experimentally determined value of -0.80 kcal/mol·K, but which is in at least qualitative agreement. This result allows us to conclude that the nonpolar surface area contributes more to the buried surface, in agreement with the negative change in heat capacity. Consistently, the amount of buried nonpolar surface area upon binding was estimated about 75% of the interfacial surface.
Table 10: Change in solvent accessible surface area (ΔASA) upon binding between LSD1 and CoREST^{286-482}.

<table>
<thead>
<tr>
<th></th>
<th>ΔASA (Å²)</th>
<th>ΔASA (Å²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Polar</td>
<td>Nonpolar</td>
</tr>
<tr>
<td>LSD1</td>
<td>12020.1</td>
<td>19688.0</td>
</tr>
<tr>
<td>CoREST^{286-482}</td>
<td>4316.8</td>
<td>7536.7</td>
</tr>
<tr>
<td>Complex</td>
<td>15035.7</td>
<td>23108.2</td>
</tr>
</tbody>
</table>

° GetAREA [98] was used for estimation
3.5 Calorimetric Titration of CoREST\textsuperscript{286-482} against LSD1 as a Function of pH.

The effect of pH on the thermodynamic parameters of the binding interaction was studied over the pH range 6.0 to 9.0 in a buffer system of constant ionic strength. At pH 6.0 and 7.0 the buffer consisted of 50 mM sodium phosphate, 50 mM glycine, and 1 mM DTT, whereas 40 mM sodium phosphate, 30 mM glycine, and 1 mM DTT were used for experiments at pH 8.0 and 9.0. The pH range was selected based on previous reports of the stability and activity of LSD1 [80]. Table 11 demonstrates that while changing pH produces no significant change in affinity, there is a trend towards a more favorable enthalpic contribution and a larger entropic penalty as pH increases. These changes in enthalpic and entropic contributions to free energy as the pH increases can be rationalized by changes in the protonation state of residues that may cause slight perturbations from the optimal binding interaction between LSD1 and CoREST\textsuperscript{286-482}. At pH 9.0, a binding isotherm was not observed. High pH presumably alters the structure of LSD1 or CoREST\textsuperscript{286-482} or both, changing their binding interaction. Previous work has shown that the catalytic activity of LSD1 is diminished above pH 9.5 in accord with our observations at pH 9.0 [82].
Table 11: Thermodynamic parameters for binding of LSD1 to CoREST\textsuperscript{286-482} at different pH (25 °C).\textsuperscript{h}

<table>
<thead>
<tr>
<th>pH</th>
<th>$K_\text{d}$ (nM)</th>
<th>$\Delta H$ (kcal mol\textsuperscript{-1})</th>
<th>$T\Delta S$ (kcal mol\textsuperscript{-1})</th>
<th>$\Delta G$ (kcal mol\textsuperscript{-1})</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0</td>
<td>4.18 ± 0.96</td>
<td>-8.56 ± 0.01</td>
<td>2.87 ± 0.66</td>
<td>-11.4 ± 2.62</td>
<td>0.90 ± 0.004</td>
</tr>
<tr>
<td>7.0</td>
<td>6.75 ± 1.49</td>
<td>-12.9 ± 0.15</td>
<td>-1.54 ± 0.34</td>
<td>-11.4 ± 2.51</td>
<td>1.00 ± 0.01</td>
</tr>
<tr>
<td>8.0</td>
<td>7.66 ± 2.91</td>
<td>-15.6 ± 0.32</td>
<td>-4.48 ± 1.71</td>
<td>-11.1 ± 4.22</td>
<td>1.00 ± 0.01</td>
</tr>
<tr>
<td>nb</td>
<td>nb</td>
<td>nb</td>
<td>nb</td>
<td>nb</td>
<td>nb</td>
</tr>
</tbody>
</table>

\textsuperscript{h} 50 mM sodium phosphate, 50 mM glycine, and 1 mM DTT, pH 6.0; 50 mM sodium phosphate, 50 mM glycine, and 1 mM DTT, pH 7.0; 40 mM sodium phosphate, 30 mM glycine, and 1 mM DTT, pH 8.0; 40 mM sodium phosphate, 30 mM glycine, and 1 mM DTT, pH 9.0. nb = no binding curve observed
3.6 Effect of the Disulfide Bond Formation upon the Binding Interaction

A cysteine from the LSD1 tower domain\textsuperscript{418-522} (residue 491) and another from CoREST\textsuperscript{286-482} (residue 379) are present at the binding interface: these residues can presumably participate in disulfide bonds. Concomitant protein oxidation leading to the mixed disulfide bonds between LSD1 and CoREST\textsuperscript{286-482} would confound interpretation of thermodynamic data. In order to study the effect of disulfide bond formation upon binding, titrations were performed at 25 °C (pH 7.4) with buffer containing only 50 mM sodium phosphate without addition of DTT. The thermodynamic parameters measured under these conditions were in reasonable agreement with those obtained with the buffer containing 1mM DTT (Table 12). On the basis of these data we concluded that derived thermodynamic parameters do not include contributions from disulfide bond formation.

<table>
<thead>
<tr>
<th>T (°C)</th>
<th>$K_d$ (nM)</th>
<th>$\Delta H$ (kcal mol$^{-1}$)</th>
<th>$T\Delta S$ (kcal mol$^{-1}$)</th>
<th>$\Delta G$ (kcal mol$^{-1}$)</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>10.0 ± 2.45</td>
<td>-12.0 ± 0.15</td>
<td>-1.06 ± 0.26</td>
<td>-10.9 ± 2.67</td>
<td>0.98 ± 0.01</td>
</tr>
</tbody>
</table>

\textsuperscript{1} 50 mM sodium phosphate, pH 7.4
3.7 Discussion and Future Directions

This chapter describes thermodynamic parameters for the binding interaction between LSD1 and CoREST\textsuperscript{286-482} determined calorimetrically. ITC experiments demonstrate that the binding between LSD1 and CoREST\textsuperscript{286-482} is a tight interaction with a dissociation constant ($K_d$) in the nanomolar range. The high binding affinity is attributed to various non-covalent interactions between LSD1 and CoREST\textsuperscript{286-482} upon binding as shown in Figure 22. In addition to the major helical interactions between two proteins, the N-terminal region of CoREST\textsuperscript{286-482} wraps around the bottom of the anti-parallel helices of the LSD1 tower domain\textsuperscript{418-522}, which may contribute to the overall strength of the binding.

Most thermodynamic data of the binding between LSD1 and CoREST\textsuperscript{286-482} showed the enthalpies of interaction greater than the free energies of association and the unfavorable entropies of association, which is commonly observed in biomolecular interactions. In the binding interaction between LSD1 and CoREST\textsuperscript{286-482}, the favorable change in enthalpy reflects a net increase in the number or strength of bonds formed between two proteins. Several candidates for the formations of hydrogen bonds and salt bridges at the interface are found, which may contribute directly to the enthalpic gain. By calculating the change in heat capacity of the binding interaction obtained from the slope of the linear temperature dependency of the change in enthalpy, we were able to confirm that hydrophobic interaction is also populated at the interface. The entropy of
association, hence, is affected by their tight association, which yields the unfavorable change in entropy.

As shown in Figure 22, the structural analysis shows that the binding interface is mainly located in the triple-helical bundle that consists of three α-helices: one helix from the linker region\(^{293-380}\) of CoREST (pink-colored) and two anti-parallel helices from the LSD1 tower domain\(^{418-522}\) (blue-colored). This type of interaction is typically classified as a coiled-coil interaction, and is found in various proteins including intermediate filaments, cell surface receptors, molecular motors, and transcription factors [99]. Like the interaction between the LSD1 and CoREST\(^{286-482}\), trimeric coiled-coil interactions have been reported in various studies of the oligomerization domains of hemagglutinin membrane glycoprotein [100], C-type mannose-binding protein [101], mechanisms of Laminin chain assembly [102], the crystal structure of GCN4-pI\(_Q\)I [103], and the crystal structure of the \textit{de novo} designed V\(_{\Lambda}\)Ld [104].

Coiled-coil interactions are usually characterized by a heptad repeat sequence \((abcdefg)_n\) [105, 106, 107, 108], with hydrophobic residues at \(a\) and \(d\) position whose interaction is well-known as a main driving force for the stability of helical conformations, and polar or charged residues at \(e\) and \(g\) position, providing further stability of the coiled-coil interaction through ionic interactions (Figure 27). However, unlike this classic periodicity, all the hydrophobic residues of three α-helices in the LSD1-CoREST\(^{286-482}\) complex are not assigned at \(a\) and \(d\) position (Figure 28). The
presence of polar residues such as asparagine, glutamine, glutamic acid, and lysine is observed at those positions, suggesting their role in the formation of hydrogen bonds and ionic interactions. This structural feature presumably compensates for the destabilizing effect of desolvation [109, 110].

Sequence alignment analysis using the DALI database [111] revealed that many proteins overlap with either two α-helical structures of LSD1418-522 or one helix of the linker region293-380 of CoREST, suggesting that the helical structure is common to other proteins. Although their sequences do not match with high scores, helical conformations can have structurally similar characteristics. Notably, the trimeric helical features formed by two proteins were not scored as hits by DALI analysis, indicating their unique structure. This analysis suggests that although the helical structures of either LSD1 tower domain or CoREST linker region can be used as a structural motif for existing proteins, the trimeric helical structure formed is less well preceded. Thus, small molecules or peptidomimetics targeting the interaction between LSD1 and CoREST may exhibit selectivity and binding affinities of the strengths needed to compete for CoREST binding.
Figure 27: A heptad repeat model of a coiled-coil interaction. Typically the coiled-coil is defined by the heptad amino acid repeat sequence \((abcdefg)_n\). A 3.5 residue periodicity along the coil axis aligns \(a\) and \(d\) position hydrophobic residues, stabilizing the interaction interface, while residues at the \(e\) and \(g\) positions are typically oppositely charged and flank the hydrophobic core, forming salt bridges between helical chains.
Figure 28: (a) Trimeric coiled-coil interaction between LSD1 tower region (blue) and CoREST linker region (red) (side view and bottom view). (b) The side chains at \( a \) and \( d \) position of a heptad repeat model are represented as sticks. The binding interface between two proteins exhibits a unique trimeric coiled-coil structure, which has been less precedent. All structural figures were generated using PyMoL (PDB file: 2IW5).
The results of the thermodynamic study of binding between LSD1 and CoREST\textsuperscript{286-482} led us to question the role of full length CoREST and the effects of other corepressor proteins such as HDACs 1/2 associated with the LSD1/CoREST complex on binding. Previous studies by Shiekhattar and coworkers examined the effect of full length CoREST in regulating the activity of LSD1 [71]. Addition of full length CoREST increased the demethylase activity of LSD1, and the deletional mapping analysis showed the necessity of both the SANT1 and SANT2 domains for the nucleosomal demethylation, although the stimulatory activity of SANT1 was weak. The SANT1 domain may serve to bridge between LSD1 and its substrates because of the structural similarity with the SANT2 domain. Thus, the SANT1 and SANT2 domains may independently facilitate LSD1-mediated demethylation. Also, the activity of HDAC1 increased by addition of full length CoREST owing to the role of an ELM2 domain of CoREST that is known to mediate the deacetylase activity of HDACs 1/2 [70]. Based on these previous studies, full length CoREST seems to have similar activity to the truncated CoREST\textsuperscript{286-482}. These results let us presume that the ELM2 domain and SANT1 domain of CoREST may be located away from the stalk formed between the LSD1 tower domain\textsuperscript{418-522} and the linker region\textsuperscript{293-380} of CoREST, and that consequently the binding affinity between LSD1 and full length CoREST should not be significantly different from the binding affinity obtained between LSD1 and the truncated CoREST\textsuperscript{286-482}. However, the exact role or function of either domain on the binding affinity can not be rationalized
until a crystal structure of LSD1 and full length CoREST is determined. Whether the presence of HDACs 1/2 affects the binding affinity between LSD1 and the full length of CoREST may depend on the location of the ELM2 domain in the complex because of the physical association of CoREST with the ELM2 domain. Assuming that the N-terminal region of CoREST, including the ELM2 domain and SANT1 domain, is not associated with the binding region of the LSD1-CoREST complex, the presence of HDACs 1/2 may not affect the binding affinity significantly as well. However, it is also possible that HDACs 1/2 is sufficiently close to the binding region of LSD1 and full length CoREST due to its spatial occupancy upon binding to the ELM2 domain, to result in weakening the binding affinity.

In recent studies [46], LSD2 has been identified as a flavin-dependent histone demethylase, with enzymatic activity and substrate specificity profile similar to those of LSD1. Unlike LSD1, however, LSD2 does not contain the tower domain, which is essential for binding to CoREST, and thus it is not able to interact with CoREST, distinguishing it from LSD1. In addition, LSD2 differs from LSD1 because it contains a CW-type zinc-finger domain at its N-terminus between residues 130-200. The role of the zinc-finger domain in LSD2 remains unclear, but we speculate with some degree of confidence that this sequence may facilitate direct binding of LSD2 to nucleosomal DNA.
In summary, our thermodynamic study has verified the tight binding interaction between CoREST, which is both enthalpically and energetically favorable. In the next chapter, studies to provide valuable insights into key residues involved in the tight binding interaction will be followed.

3.8 Experimental Section

Reagents and Materials. Media, antibiotics, and all other buffer reagents were purchased from Sigma, New England Biolabs, MP Biomedicals, and BD Biosciences. Chemically competent BL21 Star (DE3) *Escherichia coli* cells were purchased from Invitrogen. Chromatographic protein purifications were carried out on an AKTA FPLC (GE Healthcare), and the protein concentration was determined by UV absorbance spectroscopy.

Expression and Purification of LSD1. The gene encoding a truncated form of LSD1 (residues 151-852) was codon-optimized by Dr. Helena M. Gaweska, a previous lab member in the McCafferty group. The codon-optimized gene was used for expression and purification as previously described [81, 82] except that the gene was cloned into the pET15b vector instead of the pET 151-D/TOPO vector for expression. The resulting plasmid was transformed into chemically competent BL21 Star (DE3) *E. coli* cells, which were grown on a LB agar plate supplemented with ampicillin overnight at 37 °C. Streaks of the colonies on the plate were used to grow 6 L of cells in TB media with shaking (200
rpm) at 23 °C. When the cell density reached an OD$_{600}$ of 0.6, 0.5 mM IPTG was added to the flasks to induce LSD1 expression. The cells were allowed to grow overnight and collected by centrifugation at 4225 \times g. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, and 0.4 mM PMSF (pH 7.8). LSD1 was then purified via nickel-affinity chromatography, HiPrep 26/60 Sephacryl S200 gel filtration chromatography (GE Life Sciences), and anion-exchange chromatography (Q-Sepharose Fast Flow, GE Life Sciences). The final concentration of LSD1 was determined by absorption spectroscopy at 458 nm [82] and the protein was stored at -20 °C in 80 % glycerol.

Expression and Purification of CoREST$^{286-482}$. A pET28b vector containing the gene of interest encoding CoREST$^{286-482}$, whose sequence was verified, was used for expression and purification. The vector was transformed into electrocompetent BL21 Star (DE3) E. coli cells, which were used to grow for purification. Streaks of the colonies on a LB agar containing kanamycin were used to grow 6 L of bacteria. The cells were grown as described above for LSD1 purification. CoREST$^{286-482}$ was purified via nickel-affinity chromatography with a linear gradient from 50 to 500 mM imidazole in 50 mM sodium phosphate and 300 mM NaCl (pH 7.4), which was followed by purification via cation-exchange chromatography (CM-Sepharose Fast Flow, Sigma) with a linear gradient from 100 to 800 mM NaCl in 50 mM sodium phosphate (pH 7.4). The concentration of
CoREST\textsuperscript{286-482} was measured spectrophotometrically using the extinction coefficient of 16,950 cm\(^{-1}\)M\(^{-1}\) at 280 nm.

\textit{Isothermal Titration Calorimetry (ITC).} ITC experiments were performed using a MicroCal VP-ITC microcalorimeter (MicroCal, Northampton, MA) with a cell volume of 1.4346 mL. Protein samples were dialyzed against the buffer containing 50 mM sodium phosphate and 1 mM DTT (pH 7.4) and vacuum-degassed for at least 30 min before loading into the calorimeter. Approximately 1.5 mL of the degassed LSD1 and buffer were placed in the sample cell and reference cell respectively, and CoREST\textsuperscript{286-482} was loaded into the syringe injector. In most ITC experiments, aliquots (3 \(\mu\)L) of CoREST\textsuperscript{286-482} at 30 \(\mu\)M concentration were titrated sequentially against LSD1 at 3 \(\mu\)M concentration in the sample cell at 25 °C. Each injection lasted for 6 s, and there was a delay of 300 s between injections. During the titration the stirring speed was 310 rpm and total 40 to 60 injections of CoREST\textsuperscript{286-482} were titrated into LSD1. A one site-binding model (Origin 5.0 software, MicroCal software, Inc) was used to fit the data. The titrations were performed under different experimental conditions such as various buffers, temperature range of 15 °C to 35 °C, different pH, and changes of buffer additives.

\textit{Surface Plasmon Resonance (SPR) Measurements.} All SPR measurements were made using a BIAcore 3000 instrument, and data analyses were performed using the BIAevaluation 4.1 software (BIAcore). The CoREST\textsuperscript{286-482} was immobilized (~2200 RU) on a BIAcore CM5 (research grade) chip using standard amine coupling chemistry with
reagents obtained from BIAcore. In a parallel flow cell, LSD1 was immobilized (~2500 RU). During the screening experiments with 50 mM sodium phosphate and 1 mM DTT (pH 7.4) as a running buffer, LSD1 was injected over the immobilized CoREST$^{286-482}$ for 2 min at a flow rate of 30 µL/minute to monitor the binding interaction. The surfaces were regenerated by injecting 10 µL glycine (pH 2.0) at 50 µL/min flow rate. The response from the LSD1 surface was used to subtract out the background (non-specific) signal. The $K_d$ value of the LSD1-CoREST$^{286-482}$ interaction was estimated by injecting LSD1 in buffer containing 50 mM sodium phosphate and 1 mM DTT (pH 7.4) for 3 min at various concentrations ranging from 6.25 nM to 62.5 nM. A global fitting of binding curves at different concentrations with the Langmuir binding model was used to measure the association ($k_{on}$) and dissociation rate constant ($k_{off}$), and the apparent binding affinity constant ($K_d$).

**Calculation of the Solvent Accessible Surface Area (SASA).** Solvent accessible surface area (SASA) of the LSD1-CoREST$^{286-482}$ complex was estimated by GetArea [98] with the radius of the probe set to 1.4 Å. The structural data file of LSD1 in complex with CoREST$^{286-482}$ was obtained from the Protein Data Bank (PDB entry code 2IW5) and used to calculate the parameters.
4. Determination of the Key Residues of the Binding Interface between LSD1 and CoREST

4.1 Background

4.1.1 Targeting Protein-Protein Interactions

Although many cellular processes rely on properly functioning enzymatic reactions through their active sites, protein-protein interactions are essential components of various regulatory pathways; the binding of two or more proteins in cells can have a wide range of effects such as modulation or initiation of signal transduction and regulating gene transcriptions. Often protein-protein interactions hold the potential to cause pathological conditions, which indicates that modulation of such protein-protein interactions affords possibilities for treatment of human diseases.

To date, efforts to develop novel drugs have traditionally centered on targeting enzyme active sites with considerably less effort to target protein-protein interactions. This is because the enzyme active site is relatively small so that it can be easily exploited to design small molecule inhibitors. Targeting protein-protein interactions with small molecules has been more challenging due to some reasons including relatively large interfacial size and often noncontiguous binding interaction [112]. However, recently more successful approaches to target protein-protein interactions by utilizing peptides or peptidomimetics to mimic the binding interface have been reported. For example, β-peptides are capable of folding into stable secondary structures and also resisting
proteolysis. They have been used to disrupt the interaction of the HIV-1 envelope glycoprotein with the cell receptor to prevent HIV-1 infection [113]. The tumor suppressor p53 is inactivated in majority of human cancers by binding to oncogenic proteins such as hDM. Thus disturbing this interaction has been recognized as a novel strategy of cancer therapy. Consistently p53-derived peptides and peptidomimetics such as β-peptide and peptoids have been used to inhibit their binding interaction and activate the p53 pathway as a result [114, 115, 116]. The Bcl-2 family (Bcl-2 or Bcl-xL) serves to prevent apoptosis by inhibiting the function of pro-apoptotic members of the Bcl-2 family such as Bax and Bak via binding to their BH3 (Bcl-2-homology 3) domain; thus, the disruption of their interaction between Bak and Bcl-2 by BH3-derived peptides and peptidomimetics has restored biological functions of pro-apoptotic Bcl-2 family members in tumor cells [117, 118, 119].

In order to target such protein-protein interactions, it is necessary to identify their recognition sites, which aids in designing drugs with high selectivity. Intriguingly, it has been recognized that for many protein complexes only relatively small regions of the binding surface, often referred to as “hot spots” contribute significantly to the major part of the free energy of binding [85, 120, 121]. Amino acid residues in hot spots are most likely to be clustered together at the center of the protein-protein interface and are surrounded by amino acid residues that are energetically less important on stability and probably serve to occlude solvent. An interesting feature of hot spots is their structural
adaptivity. Hot spots are most likely conserved through multiple binding partner proteins; thus, they adapt to present the same residues in different structural contexts [121]. The understanding of hot spots seems to be helpful not only for the study of a single protein-protein interaction, but also for determining binding sites for other binding partners.

LSD1 overexpression and its following enzymatic activity have been implicated in various pathological problems such as cancer progression and viral gene expression by associating with various binding partners. As discussed, one of its important binding proteins, CoREST, is known to endow LSD1 with the ability to associate with and demethylate nucleosomal substrates. We expect that targeting the binding interaction between LSD1 and CoREST will provide a powerful means to inhibit the demethylation activity of LSD1 by orphanizing LSD1 from nucleosomal contexts. Thus prior to disrupting their binding interaction, in this chapter, we attempted to define the key residues involved in the binding interaction between LSD1 and CoREST\textsuperscript{286-482} by performing alanine-scanning mutagenesis and peptide fragment analysis, which may help design peptidomimetics to inhibit LSD1 function.
4.2 Calorimetric Titration of CoREST Linker$^{293-380}$ and CoREST SANT2$^{381-450}$ against LSD1

The crystal analysis shows that the binding interface is centered on the helical regions of two proteins, suggesting that the CoREST linker region (residue 293-380, designated as CoREST Linker$^{293-380}$) is important for binding to LSD1. In order to define the role of the CoREST Linker$^{293-380}$ in complex formation, we dissected CoREST$^{286-482}$ into two separate domains, CoREST Linker$^{293-380}$ and CoREST SANT2 domain (residue 381-450, designated as CoREST SANT2$^{381-450}$). Prior to the titrations of each domain against LSD1, both CoREST Linker$^{293-380}$ and CoREST SANT2$^{381-450}$ were expressed and purified as described in Experimental section. Even though the high amount of proteins were expressed at 37 °C, the proteins were found to be almost completely insoluble. When the culture temperature was shifted to 23 °C, their solubility was somewhat increased. This is most likely because low-temperature growth minimizes protein aggregation by allowing the proteins to spend more time in the folding pathway. Both proteins were purified to homogeneity via Ni (II) - affinity chromatography and cation exchange chromatography (Figure 29).
Figure 29: Expression and purification of (a) CoREST Linker\textsuperscript{293-380} and (b) CoREST SANT2\textsuperscript{381-450}. Both proteins were overexpressed in BL21 Star (DE3) E.coli stain. The first and fourth panels show the expression of each protein at 37 °C, the second and fifth show the expression of each protein at 23 °C. The third and sixth panels show purified proteins indicated by arrows.
Titrations of each domain against LSD1 were performed at 25 °C (pH 7.4) (Figure 30). Table 13 shows that although diminished enthalpic and entropic change were observed in the binding interaction between LSD1 and CoREST Linker\textsuperscript{293-380} compared to the those for binding between LSD1 and CoREST\textsuperscript{286-482}, the interaction is still characterized by a favorable enthalpic contribution near room temperature with a smaller entropic penalty at pH 7.4. There is no significant difference in binding affinity and the free energy of the binding, with $K_d$ of 7.78 ± 3.50 nM and $\Delta G$ of -11.1 ± 5.00 kcal/mol. The binding of LSD1 and CoREST Linker\textsuperscript{293-380} still occurs in a 1:1 molar ratio. Due to the close proximity of CoREST Linker\textsuperscript{293-380} to CoREST SANT\textsuperscript{281-450}, removing CoREST SANT\textsuperscript{281-450} may alter the conformation of CoREST Linker\textsuperscript{293-380} itself and thus affect solvent reorganization and many individual non-covalent interactions. It is also possible that the differences in thermodynamic parameters between the binding of LSD1 to CoREST\textsuperscript{286-482} and to CoREST Linker\textsuperscript{293-380} are attributed to different conditions of proteins themselves such as the presence of impurities, but nonetheless, the binding between LSD1 and CoREST Linker\textsuperscript{293-380} showed a tight interaction. Interestingly, the titration of CoREST SANT\textsuperscript{281-450} against LSD1 showed no binding under the same experimental condition, indicating no or only very weak binding between them. This result indicates that the residues crucial for the interaction with LSD1 lie in CoREST Linker\textsuperscript{293-380}, and the lack of the binding interaction observed between LSD1 and CoREST SANT\textsuperscript{281-450} supports the importance of CoREST Linker\textsuperscript{293-380} for LSD1 binding.
Table 13: Thermodynamic parameters for binding of LSD1 to CoREST Linker$^{293-380}$ at 25 °C.$^i$

<table>
<thead>
<tr>
<th></th>
<th>$K_r$ (nM)</th>
<th>$\Delta H$ (kcal mol$^{-1}$)</th>
<th>$T\Delta S$ (kcal mol$^{-1}$)</th>
<th>$\Delta G$ (kcal mol$^{-1}$)</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSD1-CoREST$^{286-482}$</td>
<td>15.9 ± 2.07</td>
<td>-21.3 ± 0.19</td>
<td>-10.7 ± 1.40</td>
<td>-10.6 ± 1.38</td>
<td>1.14 ± 0.00</td>
</tr>
<tr>
<td>LSD1-CoREST Linker$^{293-380}$</td>
<td>7.78 ± 3.50</td>
<td>-13.5 ± 0.37</td>
<td>-2.42 ± 1.09</td>
<td>-11.1 ± 5.00</td>
<td>0.96 ± 0.01</td>
</tr>
</tbody>
</table>

$^i$ 50 mM sodium phosphate and 1 mM DTT, pH 7.4

Figure 30: (a) A calorimetric titration of LSD1 with CoREST Linker$^{293-380}$ and (b) a titration of LSD1 with CoREST SANT2$^{381-450}$ in 50 mM sodium phosphate and 1 mM DTT, pH 7.4. The binding interaction between LSD1 and CoREST Linker$^{293-380}$ yielded the qualitatively similar thermodynamic parameters as compared to the parameters of the binding interaction between LSD1 and CoREST$^{286-482}$, while no binding isotherm curve was observed between LSD1 and CoREST SANT2$^{381-450}$.
4.3 Site-Directed Mutagenesis Study

4.3.1 Calorimetric Titration of CoREST\textsuperscript{286-482} Mutants against LSD1

As described in previous studies, the important regions of CoREST involved in the binding interaction with LSD1 were localized in the helical region of CoREST, known as CoREST Linker\textsuperscript{293-380}. In order to identify the effects of specific residues of CoREST Linker\textsuperscript{293-380} on the interactions with LSD1, ITC studies were combined with site-directed mutagenesis. Alanine scanning mutagenesis was employed to substitute specific amino acids with alanine by altering the codon sequences using site-directed \textit{in vitro} mutagenesis procedures. Alanine and glycine are commonly used for mutagenesis, but glycine often introduces conformational flexibility in protein; thus, alanine is the most preferred amino acid for mutagenesis. Alanine scanning mutagenesis has been considered slow and labor-intensive, but nonetheless it is the most reliable method for mapping functional epitopes [120].

The crystal structure of LSD1 in complex with CoREST\textsuperscript{286-482} was inspected first to analyze details of their binding interactions, and residues of CoREST\textsuperscript{286-482}, especially localized in the trimeric coiled-coil region with association of two helices of the LSD1 tower domain, were chosen and mutated to alanine: Q318A, D320A, D339A, L342A, K346A, I349A, K353A, N356A, K360A, L363A, R371A (Figure 31). Table 14 shows the summary of the thermodynamic parameters determined calorimetrically for the binding interaction between LSD1 and CoREST\textsuperscript{286-482} mutants. Interestingly, the mutations of
CoREST$^{286-482}$ show no significant effects on the binding affinity and the changes in enthalpy and entropy. Some of the mutants exhibited approximately 4-fold decreased binding affinities, but nonetheless most CoREST$^{286-482}$ mutants still revealed tight binding interactions toward LSD1 with dissociation binding constants ($K_d$) in the nanomolar ranges with thermodynamic parameters similar to the ones for binding between LSD1 and wild-type CoREST$^{286-482}$. Binding of the CoREST$^{286-482}$ mutants to LSD1 is still characterized by an enthalpically driven process and an overall entropy loss.

Consistently, the binding free energy ($\Delta G$) was not drastically changed as well. As described earlier, in protein-protein interactions, hot spots are defined as small clusters of residues that result in a change of binding free energy greater than 2.0 kcal/mol [122]. The binding free energies of all CoREST$^{286-482}$ mutants exhibit small differences less than 2.0 kcal/mol. The loss of entropy most likely results from the significant loss of the conformational degrees of freedom of residues upon the binding interaction between LSD1 and CoREST$^{286-482}$ mutants, which is compensated by an enthalpy gain due to the formation of favorable binding interactions frequently observed in protein-protein interactions. This indicates that the binding energy for the interaction between LSD1 and CoREST$^{286-482}$ is uniformly distributed across the interface rather than being concentrated in a few residues. In fact, the residues chosen for mutational studies here are all known to be involved in several hydrogen bond formations and salt bridges as well as hydrophobic interactions, which are presumably the main contributions to the favorable
enthalpic gain. As such, a singly mutated residue of CoREST does not noticeably change a thermodynamic profile of the binding interaction toward LSD1.
Figure 31: (a) A side view and (b) bottom view of LSD1 tower domain (blue) in complex with CoREST^{286-482} mutants. The mutated residues of CoREST^{286-482} are represented as sticks in orange color. All structural figures were generated using PyMoL (PDB file: 2IW5).
Table 14: Thermodynamic parameters for binding of LSD1 to CoREST<sup>286-482</sup> mutants at 25 °C.<sup>k</sup>

<table>
<thead>
<tr>
<th>CoREST&lt;sup&gt;286-482&lt;/sup&gt; mutants</th>
<th>$K_d$ (nM)</th>
<th>$\Delta H$ (kcal mol$^{-1}$)</th>
<th>$\Delta T$ (kcal mol$^{-1}$)</th>
<th>$\Delta G$ (kcal mol$^{-1}$)</th>
<th>Stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>15.9 ± 2.07</td>
<td>-21.3 ± 0.19</td>
<td>-10.7 ± 1.40</td>
<td>-10.6 ± 1.38</td>
<td>1.14 ± 0.00</td>
</tr>
<tr>
<td>Q318A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>9.25 ± 2.90</td>
<td>-20.0 ± 0.35</td>
<td>-9.07 ± 2.85</td>
<td>-10.9 ± 3.42</td>
<td>1.01 ± 0.01</td>
</tr>
<tr>
<td>D320A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>7.96 ± 2.87</td>
<td>-10.3 ± 0.22</td>
<td>0.74 ± 0.27</td>
<td>-11.0 ± 4.05</td>
<td>1.05 ± 0.01</td>
</tr>
<tr>
<td>D339A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>20.7 ± 7.45</td>
<td>-15.9 ± 0.62</td>
<td>-5.36 ± 1.98</td>
<td>-10.5 ± 3.86</td>
<td>0.95 ± 0.02</td>
</tr>
<tr>
<td>L342A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>11.3 ± 1.12</td>
<td>-36.7 ± 0.21</td>
<td>-25.9 ± 2.57</td>
<td>-10.8 ± 1.07</td>
<td>0.99 ± 0.003</td>
</tr>
<tr>
<td>K346A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>9.61 ± 2.40</td>
<td>-20.0 ± 0.35</td>
<td>-9.10 ± 2.30</td>
<td>-10.9 ± 2.75</td>
<td>1.02 ± 0.01</td>
</tr>
<tr>
<td>I349A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>1.98 ± 0.87</td>
<td>-18.2 ± 0.18</td>
<td>-6.31 ± 2.80</td>
<td>-11.9 ± 5.27</td>
<td>1.01 ± 0.005</td>
</tr>
<tr>
<td>K353A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>10.9 ± 1.38</td>
<td>-20.9 ± 0.17</td>
<td>-9.98 ± 1.28</td>
<td>-10.9 ± 1.39</td>
<td>0.99 ± 0.004</td>
</tr>
<tr>
<td>N356A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>8.58 ± 4.12</td>
<td>-13.7 ± 0.34</td>
<td>-2.72 ± 1.33</td>
<td>-11.0 ± 5.36</td>
<td>0.96 ± 0.01</td>
</tr>
<tr>
<td>K360A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>68.4 ± 10.3</td>
<td>-14.1 ± 0.24</td>
<td>-4.33 ± 0.66</td>
<td>-9.77 ± 1.47</td>
<td>0.99 ± 0.01</td>
</tr>
<tr>
<td>L363A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>52.1 ± 15.1</td>
<td>-10.1 ± 0.28</td>
<td>-0.15 ± 0.04</td>
<td>-9.96 ± 2.88</td>
<td>1.07 ± 0.02</td>
</tr>
<tr>
<td>R371A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>30.1 ± 15.3</td>
<td>-9.09 ± 0.35</td>
<td>1.17 ± 0.60</td>
<td>-10.3 ± 5.24</td>
<td>0.96 ± 0.04</td>
</tr>
<tr>
<td>N356AK360A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>14.1 ± 4.37</td>
<td>-18.8 ± 0.39</td>
<td>-8.06 ± 2.57</td>
<td>-10.7 ± 3.40</td>
<td>0.98 ± 0.01</td>
</tr>
<tr>
<td>K353AK360A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>55.2 ± 13.2</td>
<td>-33.1 ± 0.90</td>
<td>-23.2 ± 5.59</td>
<td>-9.90 ± 2.37</td>
<td>1.05 ± 0.02</td>
</tr>
</tbody>
</table>

<sup>k</sup> 50 mM sodium phosphate and 1 mM DTT, pH 7.4
4.4 Peptide Fragment Analysis

In addition to the site-directed mutagenesis study, we have utilized a fragment-based approach to identify important regions of CoREST\textsuperscript{293-380} involved in the binding interaction with LSD1. Toward this end, five 20-mer peptides corresponding to the sequences of CoREST Linker\textsuperscript{293-380} were synthesized using microwave Fmoc-solid phase peptide chemistry (Table 15 and Figure 32) and used as potential inhibitors to disrupt the binding interaction between LSD1 and GST-tagged CoREST\textsuperscript{286-482} in an \textit{in vitro} pull-down assay. Those peptide fragments covering CoREST Linker\textsuperscript{293-380} may tuck themselves into the binding interface of LSD1 and prevent the binding interaction with CoREST\textsuperscript{286-482}. Thus, in the assay, the peptide disrupting the binding interaction at a lower concentration is expected to competitively bind to LSD1 with higher binding affinity. LSD1 forming a complex with the peptide is washed away and not visualized in Western blotting analysis (Scheme 4).
Table 15: 20-mer peptides corresponding to CoREST Linker<sup>293-380</sup> sequences.

<table>
<thead>
<tr>
<th>Peptides</th>
<th>Residues</th>
<th>Sequences</th>
<th>Mass Calculated (M&lt;sup&gt;+&lt;/sup&gt;, m/z)</th>
<th>Mass Observed (M&lt;sup&gt;+&lt;/sup&gt;, m/z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoR1</td>
<td>300-319</td>
<td>TQAKNRAKRKPFGMFLSQ</td>
<td>2572.00</td>
<td>2570.30</td>
</tr>
<tr>
<td>CoR2</td>
<td>315-334</td>
<td>FLSQEDVEAVSANATAATTV</td>
<td>2280.45</td>
<td>2281.02</td>
</tr>
<tr>
<td>CoR3</td>
<td>330-349</td>
<td>AATTVLRQLDMELVSVKRQI</td>
<td>2528.98</td>
<td>2528.74</td>
</tr>
<tr>
<td>CoR4</td>
<td>345-364</td>
<td>VKRQIQNIKQTNSALKEKLD</td>
<td>2611.01</td>
<td>2611.15</td>
</tr>
<tr>
<td>CoR5</td>
<td>360-379</td>
<td>KEKLDGGIEPYRLEVIQKC</td>
<td>2571.99</td>
<td>2571.64</td>
</tr>
</tbody>
</table>

Figure 32: Five 20-mer peptides corresponding to the sequence of CoREST Linker<sup>293-380</sup> are synthesized, which are represented in different colors. Cyan represents CoR1, yellow CoR2, pink CoR3, green CoR4, and brown CoR5 as described in Table 15. Structural figure was generated using PyMoL (PDB file: 2IW5).
Scheme 4: GST pull-down assay. CoREST\textsuperscript{286-482} is GST- tagged, which is captured on glutathione agarose beads. Peptides (indicated by P in light-orange color) corresponding to the sequences of CoREST Linker\textsuperscript{293-380} are added to GST-CoREST\textsuperscript{286-482}, which is followed by the addition of LSD1. The mixture is incubated for 1 h at room temperature. Peptides are expected to disrupt the binding interaction by competitively binding to LSD1. The LSD1-peptide complex is washed away and any bound LSD1 to GST-CoREST\textsuperscript{286-482} is eluted, resolved by SDS-PAGE, and analyzed by Western blotting. As a control, no CoREST\textsuperscript{286-482} is loaded onto the glutathione agarose beads; thus, LSD1 is washed away.
However, as shown in Figure 33, the short fragments of the CoREST Linker$_{293-380}$ do not exhibit potent inhibitory activity to disrupt the binding interaction; they suppressed the binding interaction at excessive concentrations of around 500 µM. The results of the pull-down assay suggest that the 20-mer short fragments of the CoREST Linker may not be sufficient in length to bind to LSD1, and as described above, due to the uniformly distributed binding energy over the binding interface, the entire region of the CoREST Linker$_{293-380}$ may be required for efficient inhibition of the binding interaction. Peptide-fragment analysis has supported the phenomenon observed in the site-directed mutagenesis study.
Figure 33: Western blot probing LSD1 after binding to GST-CoREST\textsuperscript{286-482} with potential peptide inhibitors listed in Table 15. Mostly those peptides showed weak bindings to LSD1 with poor inhibition activity at excessive concentration of 500 µM.
4.5 Spot Synthesis

Peptide array covering to the sequences of CoREST Linker$^{293-380}$ was prepared by using the Spot technology on cellulose support to investigate the key regions of CoREST$^{293-380}$ for binding to LSD1. CoREST Linker$^{293-380}$ was segmented into 12-mer peptides overlapping by ten amino acids to insure full coverage. The Spot synthesis was initially developed by Ronald Frank and co-workers in 1990 [123]. The basic principle follows standard Fmoc-peptide chemistry and involves the delivery of small volumes of activated amino acids directly to distinct points on a cellulose membrane support, creating a pattern of small spots [124]. Over several coupling cycles, large numbers of peptides are built up on the membrane support and then used to study protein-protein interactions, especially to identify peptides that bind to proteins [124, 125]. The presence of bound protein to peptides can be detected via Western blot analysis (Scheme 5). Table 16 shows the list of peptides synthesized using Spot technique. The spot G5 serves as a control where no peptide is developed except the spacer β-Ala.
Scheme 5: Spot synthesis technique mainly involves three phases: preparation of the cellulose membrane, stepwise couplings of the amino acids, and the cleavage of the side-chain protection groups. The presence of bound LSD1 to peptides can be detected via Western blot analysis.
Table 16: Sequences of peptides synthesized on the membrane. Reactivity was determined by visual observation: +++ strong, ++ average, + weak reactivity.

<table>
<thead>
<tr>
<th>Spot position</th>
<th>Residues</th>
<th>1 (N)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12 (C)</th>
<th>Reactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>293-304</td>
<td>V</td>
<td>K</td>
<td>K</td>
<td>E</td>
<td>K</td>
<td>H</td>
<td>S</td>
<td>T</td>
<td>Q</td>
<td>A</td>
<td>K</td>
<td>N</td>
<td>+++</td>
</tr>
<tr>
<td>B1</td>
<td>295-306</td>
<td>K</td>
<td>E</td>
<td>K</td>
<td>H</td>
<td>S</td>
<td>T</td>
<td>Q</td>
<td>A</td>
<td>K</td>
<td>N</td>
<td>R</td>
<td>A</td>
<td>+++</td>
</tr>
<tr>
<td>C1</td>
<td>297-308</td>
<td>K</td>
<td>H</td>
<td>S</td>
<td>T</td>
<td>Q</td>
<td>A</td>
<td>K</td>
<td>N</td>
<td>R</td>
<td>A</td>
<td>K</td>
<td>R</td>
<td>+++</td>
</tr>
<tr>
<td>D1</td>
<td>299-310</td>
<td>S</td>
<td>T</td>
<td>Q</td>
<td>A</td>
<td>K</td>
<td>N</td>
<td>R</td>
<td>A</td>
<td>K</td>
<td>R</td>
<td>K</td>
<td>P</td>
<td>+++</td>
</tr>
<tr>
<td>E1</td>
<td>301-312</td>
<td>Q</td>
<td>A</td>
<td>K</td>
<td>N</td>
<td>R</td>
<td>A</td>
<td>K</td>
<td>R</td>
<td>K</td>
<td>P</td>
<td>P</td>
<td>K</td>
<td>+++</td>
</tr>
<tr>
<td>F1</td>
<td>303-314</td>
<td>K</td>
<td>N</td>
<td>R</td>
<td>A</td>
<td>K</td>
<td>R</td>
<td>K</td>
<td>P</td>
<td>P</td>
<td>K</td>
<td>G</td>
<td>M</td>
<td>+++</td>
</tr>
<tr>
<td>G1</td>
<td>305-316</td>
<td>R</td>
<td>A</td>
<td>K</td>
<td>R</td>
<td>K</td>
<td>P</td>
<td>P</td>
<td>K</td>
<td>G</td>
<td>M</td>
<td>F</td>
<td>L</td>
<td>+++</td>
</tr>
<tr>
<td>A2</td>
<td>307-318</td>
<td>K</td>
<td>R</td>
<td>K</td>
<td>P</td>
<td>P</td>
<td>K</td>
<td>G</td>
<td>M</td>
<td>F</td>
<td>L</td>
<td>S</td>
<td>Q</td>
<td>+++</td>
</tr>
<tr>
<td>B2</td>
<td>309-320</td>
<td>K</td>
<td>P</td>
<td>P</td>
<td>K</td>
<td>G</td>
<td>M</td>
<td>F</td>
<td>L</td>
<td>S</td>
<td>Q</td>
<td>E</td>
<td>D</td>
<td>+++</td>
</tr>
<tr>
<td>C2</td>
<td>311-322</td>
<td>P</td>
<td>K</td>
<td>G</td>
<td>M</td>
<td>F</td>
<td>L</td>
<td>S</td>
<td>Q</td>
<td>E</td>
<td>D</td>
<td>V</td>
<td>E</td>
<td>+++</td>
</tr>
<tr>
<td>D2</td>
<td>313-324</td>
<td>G</td>
<td>M</td>
<td>F</td>
<td>L</td>
<td>S</td>
<td>Q</td>
<td>E</td>
<td>D</td>
<td>V</td>
<td>E</td>
<td>A</td>
<td>V</td>
<td>+++</td>
</tr>
<tr>
<td>E2</td>
<td>315-326</td>
<td>F</td>
<td>L</td>
<td>S</td>
<td>Q</td>
<td>E</td>
<td>D</td>
<td>V</td>
<td>E</td>
<td>A</td>
<td>V</td>
<td>S</td>
<td>A</td>
<td>++</td>
</tr>
<tr>
<td>F2</td>
<td>317-328</td>
<td>S</td>
<td>Q</td>
<td>E</td>
<td>D</td>
<td>V</td>
<td>E</td>
<td>A</td>
<td>V</td>
<td>S</td>
<td>A</td>
<td>N</td>
<td>A</td>
<td>+</td>
</tr>
<tr>
<td>G2</td>
<td>319-330</td>
<td>E</td>
<td>D</td>
<td>V</td>
<td>E</td>
<td>A</td>
<td>V</td>
<td>S</td>
<td>A</td>
<td>N</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>+</td>
</tr>
<tr>
<td>A3</td>
<td>321-332</td>
<td>V</td>
<td>E</td>
<td>A</td>
<td>V</td>
<td>S</td>
<td>A</td>
<td>N</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>+</td>
</tr>
<tr>
<td>B3</td>
<td>323-334</td>
<td>A</td>
<td>V</td>
<td>S</td>
<td>A</td>
<td>N</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>V</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>C3</td>
<td>325-336</td>
<td>S</td>
<td>A</td>
<td>N</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>V</td>
<td>L</td>
<td>R</td>
<td>+++</td>
</tr>
<tr>
<td>D3</td>
<td>327-338</td>
<td>N</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>V</td>
<td>L</td>
<td>R</td>
<td>Q</td>
<td>L</td>
<td>+++</td>
</tr>
<tr>
<td>E3</td>
<td>329-340</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>V</td>
<td>L</td>
<td>R</td>
<td>Q</td>
<td>L</td>
<td>D</td>
<td>M</td>
<td>+++</td>
</tr>
<tr>
<td>F3</td>
<td>331-342</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>V</td>
<td>L</td>
<td>R</td>
<td>Q</td>
<td>L</td>
<td>D</td>
<td>M</td>
<td>E</td>
<td>L</td>
<td>++</td>
</tr>
<tr>
<td>G3</td>
<td>333-344</td>
<td>T</td>
<td>V</td>
<td>L</td>
<td>R</td>
<td>Q</td>
<td>L</td>
<td>D</td>
<td>M</td>
<td>E</td>
<td>L</td>
<td>V</td>
<td>S</td>
<td>++</td>
</tr>
<tr>
<td>A4</td>
<td>335-346</td>
<td>L</td>
<td>R</td>
<td>Q</td>
<td>L</td>
<td>D</td>
<td>M</td>
<td>E</td>
<td>L</td>
<td>V</td>
<td>S</td>
<td>V</td>
<td>K</td>
<td>++</td>
</tr>
<tr>
<td>B4</td>
<td>337-348</td>
<td>Q</td>
<td>L</td>
<td>D</td>
<td>M</td>
<td>E</td>
<td>L</td>
<td>V</td>
<td>S</td>
<td>V</td>
<td>K</td>
<td>R</td>
<td>Q</td>
<td>+++</td>
</tr>
<tr>
<td>C4</td>
<td>339-350</td>
<td>D</td>
<td>M</td>
<td>E</td>
<td>L</td>
<td>V</td>
<td>S</td>
<td>V</td>
<td>K</td>
<td>R</td>
<td>Q</td>
<td>I</td>
<td>Q</td>
<td>+++</td>
</tr>
<tr>
<td>D4</td>
<td>341-352</td>
<td>E</td>
<td>L</td>
<td>V</td>
<td>S</td>
<td>V</td>
<td>K</td>
<td>R</td>
<td>Q</td>
<td>I</td>
<td>Q</td>
<td>N</td>
<td>I</td>
<td>+++</td>
</tr>
<tr>
<td>E4</td>
<td>343-354</td>
<td>V</td>
<td>S</td>
<td>V</td>
<td>K</td>
<td>R</td>
<td>Q</td>
<td>I</td>
<td>Q</td>
<td>N</td>
<td>I</td>
<td>K</td>
<td>Q</td>
<td>+++</td>
</tr>
<tr>
<td>F4</td>
<td>345-356</td>
<td>V</td>
<td>K</td>
<td>R</td>
<td>Q</td>
<td>I</td>
<td>Q</td>
<td>N</td>
<td>I</td>
<td>K</td>
<td>Q</td>
<td>T</td>
<td>N</td>
<td>+++</td>
</tr>
<tr>
<td>G4</td>
<td>347-358</td>
<td>R</td>
<td>Q</td>
<td>I</td>
<td>Q</td>
<td>N</td>
<td>I</td>
<td>K</td>
<td>Q</td>
<td>T</td>
<td>N</td>
<td>S</td>
<td>A</td>
<td>+++</td>
</tr>
<tr>
<td>A5</td>
<td>349-360</td>
<td>I</td>
<td>Q</td>
<td>N</td>
<td>I</td>
<td>K</td>
<td>Q</td>
<td>T</td>
<td>N</td>
<td>S</td>
<td>A</td>
<td>L</td>
<td>K</td>
<td>+++</td>
</tr>
<tr>
<td>B5</td>
<td>351-362</td>
<td>N</td>
<td>I</td>
<td>K</td>
<td>Q</td>
<td>T</td>
<td>N</td>
<td>S</td>
<td>A</td>
<td>L</td>
<td>K</td>
<td>E</td>
<td>K</td>
<td>+++</td>
</tr>
<tr>
<td>C5</td>
<td>353-364</td>
<td>K</td>
<td>Q</td>
<td>T</td>
<td>N</td>
<td>S</td>
<td>A</td>
<td>L</td>
<td>K</td>
<td>E</td>
<td>K</td>
<td>L</td>
<td>D</td>
<td>+</td>
</tr>
<tr>
<td>D5</td>
<td>355-366</td>
<td>T</td>
<td>N</td>
<td>S</td>
<td>A</td>
<td>L</td>
<td>K</td>
<td>E</td>
<td>K</td>
<td>L</td>
<td>D</td>
<td>G</td>
<td>G</td>
<td>+</td>
</tr>
<tr>
<td>E5</td>
<td>357-368</td>
<td>S</td>
<td>A</td>
<td>L</td>
<td>K</td>
<td>E</td>
<td>K</td>
<td>L</td>
<td>D</td>
<td>G</td>
<td>G</td>
<td>I</td>
<td>E</td>
<td>+</td>
</tr>
<tr>
<td>F5</td>
<td>359-370</td>
<td>L</td>
<td>K</td>
<td>E</td>
<td>K</td>
<td>L</td>
<td>D</td>
<td>G</td>
<td>G</td>
<td>I</td>
<td>E</td>
<td>P</td>
<td>Y</td>
<td>++</td>
</tr>
<tr>
<td>A6</td>
<td>361-372</td>
<td>E</td>
<td>K</td>
<td>L</td>
<td>D</td>
<td>G</td>
<td>G</td>
<td>I</td>
<td>E</td>
<td>P</td>
<td>Y</td>
<td>R</td>
<td>L</td>
<td>+++</td>
</tr>
<tr>
<td>B6</td>
<td>363-374</td>
<td>L</td>
<td>D</td>
<td>G</td>
<td>G</td>
<td>I</td>
<td>E</td>
<td>P</td>
<td>Y</td>
<td>R</td>
<td>L</td>
<td>P</td>
<td>E</td>
<td>+++</td>
</tr>
<tr>
<td>C6</td>
<td>365-376</td>
<td>G</td>
<td>I</td>
<td>E</td>
<td>P</td>
<td>Y</td>
<td>R</td>
<td>L</td>
<td>P</td>
<td>E</td>
<td>V</td>
<td>I</td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td>D6</td>
<td>367-378</td>
<td>I</td>
<td>E</td>
<td>P</td>
<td>Y</td>
<td>R</td>
<td>L</td>
<td>P</td>
<td>E</td>
<td>V</td>
<td>I</td>
<td>Q</td>
<td>K</td>
<td>+++</td>
</tr>
<tr>
<td>E6</td>
<td>369-380</td>
<td>P</td>
<td>Y</td>
<td>R</td>
<td>L</td>
<td>P</td>
<td>E</td>
<td>V</td>
<td>I</td>
<td>Q</td>
<td>K</td>
<td>C</td>
<td>N</td>
<td>+++</td>
</tr>
</tbody>
</table>
According to the chemiluminescent probing result by Western blotting, the helical region of CoREST\textsuperscript{286-482} seems to be mainly involved in the binding interaction with LSD1 as expected. The reactivity signals were not ascribed to non-specific bindings between LSD1 antibody and peptides because no signals were detected when the membrane was incubated with LSD1 antibody in the absence of LSD1 protein. It looks like the first N-terminal region of CoREST\textsuperscript{293-380} (residues 293-315) shows relatively strong binding to LSD1, but since crystal structure starts with CoREST residue 308, there may be some discrepancies in results between the signals on the spot membrane and the crystal structure reflecting the spot membranes probed. The short \(\alpha\)-helix of CoREST Linker\textsuperscript{293-380} (residues 315-324), which packs against the substrate binding subdomain of AOD of LSD1, and the end of the long helical portion of CoREST Linker\textsuperscript{293-380} (residues 363-370) exhibited relatively weak binding, while the long helical portion of CoREST Linker\textsuperscript{293-380} (residues 325-362) showed relatively strong reactivity toward LSD1 binding (Figure 34). The Spot synthesis technique allows us to probe a large number of peptides at the same time by a simple method, but several factors may influence the signal intensity and binding of peptides to LSD1, and thus it is difficult to decide quantitatively between binding and non-binding. For example, the synthesis may lead to different amounts of the correct full-length peptide in the spots, which may results in different reactivity signals. However, nonetheless this experiment hinted that the residues in CoREST Linker\textsuperscript{293-380}, particularly the residues in the long helical region, are evenly
involved in the tight binding interaction toward LSD1 as we observed from studies of site-directed mutagenesis and peptide-fragment analysis. Thus, we decided to utilize the entire linker region of CoREST for \textit{in vitro} and \textit{in vivo} inhibition assays, which will be discussed in Chapter 5 (Figure 35).
Figure 34: (a) Reactivity of 12-mer peptides corresponding to CoREST Linker$^{293-380}$ with LSD1 and (b) crystal structure reflecting the reactivity of the peptides on the spot membrane. Sequences colored in red and light red represent relatively strong and weak reactivities toward LSD1 binding. The regions labeled as a, b, and c exhibited relatively strong reactivities in binding to LSD1. Structural figure was generated using PyMoL (PDB file: 2IW5).
Figure 35: A putative model of the potential crystal structure of LSD1 (blue) in complex with the linker region of CoREST (orange). The entire linker region of CoREST may be required for complete inhibition of the binding interaction between LSD1 and CoREST\textsuperscript{286-482}. Structural figure was generated using PyMoL (PDB file: 2IW5).
4.6 Discussion and Future Directions

In this chapter, we have attempted to define essential residues of the binding interaction between LSD1 and CoREST\textsuperscript{286-482}, especially exploiting CoREST Linker\textsuperscript{293-380} residues. Initially we mutated some CoREST Linker\textsuperscript{293-380} residues, presumably involved in the binding with LSD1, to alanine and determined the effect of the mutations on the binding interaction with LSD1 by combining with ITC. Interestingly enough, none of the mutations had large effect on the affinity and the free energy of binding in interaction with LSD1; the mutations did not lead to a significant drop in the binding constant, typically 10-fold or higher, and a large change of binding energy greater than 2 kcal/mol. This indicates that the binding energy for the interaction is evenly distributed across the interface rather than being concentrated in a few residues. This phenomenon was observed in the peptide fragment analysis. We synthesized peptides corresponding to CoREST Linker\textsuperscript{293-380}; this Linker domain was segmented into 20-mer peptides overlapping by five amino acids to ensure full coverage. In order to examine the capability of the peptides to bind to LSD1 and subsequently prevent the interaction between LSD1 and CoREST\textsuperscript{286-482}, a GST pull-down assay was employed. However, most of these peptides exhibited weak bindings to LSD1 and thus prevented the binding interaction at extremely excessive concentration. The insignificant binding of the peptides to LSD1 could be attributed to a few factors. First, it may be necessary for the entire CoREST Linker\textsuperscript{293-380} in order to bind to LSD1, and targeting a smaller region may
not be optimal. Also, it is possible that the peptides do not contain α-helical conformation and hence may not have the favorable secondary structure for binding to LSD1. Indeed, such small portions of protein have shown little or no helical character when excised from the stabilizing protein context [126]. In the Spot assay, although we were not able to analyze the binding of peptides to LSD1 quantitatively, qualitative information was provided that the entire CoREST Linker\(^{293-380}\), particularly the long helical region of CoREST Linker\(^{293-380}\) is necessary for binding to LSD1. Taken together, we believed that the central binding determinant of CoREST in binding to LSD1 lies along entire CoREST Linker\(^{293-380}\), which creates a triple helical bundle by interacting with two antiparallel helices of LSD1.

Based on the known structure of LSD1 in complex with CoREST\(^{286-482}\), the binding interface between LSD1 and CoREST\(^{286-482}\) is predicted to be approximately 2592.86 Å, which is considered a very large interface for the heterocomplex as compared to the 1400 ~ 1600 Å of average of the binding interface [97, 112]. Therefore, targeting such a large binding interface with small molecules or short peptide segments would not be possible as mentioned; thus, in order for the peptides examined in our experiment to be used to target the binding interaction, it may be necessary to test the combination of two or more peptides or large peptides covering the entire helical interface. Thus, they presumably enhance their ability to prevent the interactions between LSD1 and CoREST\(^{286-482}\) by competitively binding to the large surface of LSD1.
Alternatively, since the CoREST interaction lies within its helical segment, the peptides may need to be synthetically modified to enforce the α-helical conformation. Recently much research has been focused on the synthetic strategies for restoring their helical conformation. Peptide stapling has been reported as one of such synthetic strategies; the staple consists of an all-hydrocarbon macrocyclic bridge connecting adjacent turns of the helix [127]. The stapling is incorporated into a peptide by introducing two α-methyl, α-alkenyl amino acids during solid-phase peptide synthesis, which is followed by ruthenium-mediated ring-closing olefin metathesis (RCM) [127]. This strategy has been applied to various protein-protein interactions to modulate their biological functions, including p53 tumor suppressor pathway and anti-apoptotic pathway of BCL-2 family [119, 128]. The incorporation of β-amino acids with α-residues, referred to as α/β-peptides or foldamers, also has drawn considerable interest. Since β-residues have extra carbon, the backbone of α/β-peptides is longer than that of peptides that consist of α-amino acids, and thus α/β-peptides most likely form a helical structure [129, 130]. Additionally, other strategies for the stabilization of α-helices involve replacement of one of the main chain intramolecular hydrogen bonds between the C=O of the ith amino acid residue and the NH of the i+4th amino acid residue with a covalent linkage (C=X-Y-N) and a i, i+7 cysteine cross-linker by utilizing bisarylmethylene bromides [131, 132]. It is of note that such peptidomimetics have shown increased cellular uptake and proteolytic resistance as well as improved helicity of peptides, thus
it is expected that they may be cell-permeable. By exploiting these methods, the helicity of CoREST peptides will be enhanced and thus their ability to bind to LSD1 and eventually disrupt the binding interaction may be improved. However, prior to the design and synthesis of peptidomimetics to efficiently target the binding interaction between LSD1 and CoREST\textsuperscript{286-482}, it is necessary to validate whether the disruption of the binding interaction can propose a new means to inhibit the demethylation activity of LSD1 in cells and thus circumvent simultaneous inhibition of other FAD-dependent amine oxidases that current small molecule LSD1 inhibitors do. Toward this end, we decided to use the entire linker region of CoREST (designated as a Linker peptide for the rest of thesis) as a molecular probe for further \textit{in vitro} and \textit{in vivo} inhibition assays, which will be discussed in the following chapter.

\textbf{4.7 Experimental Section}

\textit{Reagents and Materials.} Media, antibiotics, and all other buffer reagents were purchased from Sigma, New England Biolabs, MP Biomedicals, and BD Biosciences. Antibody (Ab) that recognizes LSD1 (ab17721) was purchased from Abcam. Chemically competent BL21 Star (DE3) \textit{E. coli} cells were purchased from Invitrogen. Amino acids used for peptide synthesis were purchased from AnaSpec and Novabiochem, and organic solvents including DMF, DCM, piperidine, and TFA were obtained from Sigma, EMD, and BDH chemicals. Chromatographic protein purifications were carried out on
an AKTA FPLC (GE Healthcare), and the protein concentration was determined by UV absorbance spectroscopy. Direct binding studies were performed using isothermal titration calorimetry on a Microcal VP-ITC instrument.

Expression and Purification of LSD1. The gene encoding a truncated form of LSD1 (residues 151-852) was codon-optimized by Dr. Helena M. Gaweska, a previous lab member in the McCafferty group. The codon-optimized gene was used for expression and purification as previously described [81, 82] except that the gene was cloned into the pET15b vector instead of the pET 151-D/TOPO vector for expression. The resulting plasmid was transformed into chemically competent BL21 Star (DE3) E. coli cells, which were grown on a LB agar plate supplemented with ampicillin overnight at 37 °C. Streaks of the colonies on the plate were used to grow 6 L of cells in TB media with shaking (200 rpm) at 23 °C. When the cell density reached an OD_{600} of 0.6, 0.5 mM IPTG was added to the flasks to induce LSD1 expression. The cells were allowed to grow overnight and collected by centrifugation at 4225 × g. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, and 0.4 mM PMSF (pH 7.8). LSD1 was then purified via nickel-affinity chromatography, HiPrep 26/60 Sephacryl S200 gel filtration chromatography (GE Life Sciences), and anion-exchange chromatography (Q-Sepharose Fast Flow, GE Life Sciences). The final concentration of LSD1 was determined by absorption spectroscopy at 458 nm [82] and the protein was stored at -20 °C in 80 % glycerol.
Expression and Purification of CoREST\textsuperscript{286-482}. A pET28b vector containing the gene of interest encoding CoREST\textsuperscript{286-482} was used for expression and purification. The vector was transformed into electrocompetent BL21 Star (DE3) \textit{E. coli} cells, which were used to grow for purification. Streaks of the colonies on a LB agar containing kanamycin were used to grow 6 L of bacteria. The cells were grown as described above for LSD1 purification. CoREST\textsuperscript{286-482} was purified via nickel-affinity chromatography with a linear gradient from 50 to 500 mM imidazole in 50 mM sodium phosphate and 300 mM NaCl (pH 7.4), which was followed by purification via cation-exchange chromatography (CM-Sepharose Fast Flow, Sigma) with a linear gradient from 100 to 800 mM NaCl in 50 mM sodium phosphate (pH 7.4). The concentration of CoREST\textsuperscript{286-482} was measured spectrophotometrically using the extinction coefficient of 16,950 cm\textsuperscript{-1}M\textsuperscript{-1} at 280 nm.

Expression and Purification of CoREST Linker\textsuperscript{293-380} and CoREST SANT2\textsuperscript{381-450}. The coding sequence of CoREST Linker\textsuperscript{293-380} was extracted and amplified by using PCR with a forward primer (5\textquoteleft GCG CAT ATG GTC AAA AAA GAA AAA CAT AGC ACA CAA GCT AAA-3\textacute{'}\prime) and a reverse primer (5\textquoteleft -GCG CTC GAG TTA ATT ACA TTT CTG AAT GAC CTC TGG AGG-3\textacute{'}\prime) under the following conditions: an initial denaturation step for 5 min at 95 °C, 30 cycles of denaturation for 1 min at 95 °C, annealing for 1 min at a gradient of 54-65 °C, elongation for 1:40 min at 70 °C, and followed by a final elongation step for 10 min at 70 °C. The primers were designed to contain NdeI and XhoI restriction sites at the N- and C-termini respectively to allow for facile ligation into the pET28b
vector. The vector was transformed into chemically competent BL21 Star (DE3) E. coli cells, and the cells were grown as described above for LSD1 purification. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, 0.1 % Triton X-100, 0.4 mM PMSF, and 1X protease inhibitor cocktail solution (pH 7.8). CoREST linker was then purified via nickel-affinity chromatography with a linear gradient from 50 to 500 mM imidazole in 50 mM sodium phosphate and 300 mM NaCl (pH 7.8), which was followed by purification via cation-exchange chromatography (SP-Sepharose Fast Flow, Sigma) with a linear gradient from 100 to 800 mM NaCl in 50 mM sodium phosphate (pH 7.8). The protein concentration was measured spectrophotometrically at 280 nm with its extinction coefficient of 1,490 cm⁻¹M⁻¹.

The gene encoding CoREST SANT2 was extracted, amplified by PCR with a forward primer (5′-GCG CAT ATG GCA CGT TGG ACT ACA GAA GAG CAG CTT-3′) and a reverse primer (5′-GCT CTC GAG TTA ACT GGG CCC ATT GGT CTC TTC TTT ACC-3′), and ligated into the pET 28b vector. The vector was transformed into chemically competent BL21 Star (DE3) E. coli cells, and cells were grown overnight as described above for LSD1 purification. CoREST SATN2 was purified via nickel-affinity chromatography with a linear gradient from 50 to 500 mM imidazole in 50 mM MES and 300 mM NaCl (pH 6.0) and cation-exchange chromatography (SP-Sepharose Fast Flow, Sigma) with a gradient from 100 to 800 mM NaCl in 50 mM MES (pH 6.0).
The protein concentration was measured spectrophotometrically at 280 nm using its extinction coefficient of 13,980 cm⁻¹M⁻¹.

*Site-Directed Mutagenesis of CoREST²⁸⁶⁻⁴⁸².* To provide structural information on the interaction between LSD1 and CoREST²⁸⁶⁻⁴⁸², site-specific mutants were designed to locate the binding interaction sites. The singly mutated genes corresponding to Q318A, D320A, D339A, L342A, K346A, I349A, K353A, N356A, K360A, L363A, and R371A of CoREST²⁸⁶⁻⁴⁸² were made with the QuickChange site-directed mutagenesis kit (Stratagene) according to the manufacturer’s instructions using pET28b-CoREST²⁸⁶⁻⁴⁸² as a template. Then the sequence-verified gene was used for expression and purification of each CoREST mutant.

**PCR reaction mix**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mM dNTP mix</td>
<td>2 μL</td>
</tr>
<tr>
<td>CoREST²⁸⁶⁻⁴⁸² template</td>
<td>1 μL</td>
</tr>
<tr>
<td>Forward primer</td>
<td>2 μL</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>2 μL</td>
</tr>
<tr>
<td>10X pfu ultraAD reaction buffer</td>
<td>5 μL</td>
</tr>
<tr>
<td>pfu turbo polymerase</td>
<td>1 μL</td>
</tr>
<tr>
<td>Sterile water</td>
<td>37 μL</td>
</tr>
<tr>
<td>Total</td>
<td>50 μL</td>
</tr>
</tbody>
</table>

**PCR condition (30 cycles)**

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>95 °C</td>
<td>10 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95 °C</td>
<td>1 min</td>
</tr>
<tr>
<td>Annealing primers</td>
<td>54-65 °C</td>
<td>2 min</td>
</tr>
<tr>
<td>Elongation</td>
<td>72 °C</td>
<td>10 min</td>
</tr>
<tr>
<td>Final elongation</td>
<td>72 °C</td>
<td>10 min</td>
</tr>
<tr>
<td>Store</td>
<td>4 °C</td>
<td>∞</td>
</tr>
</tbody>
</table>

*Purification of CoREST²⁸⁶⁻⁴⁸² Mutants.* CoREST²⁸⁶⁻⁴⁸² mutants were purified as described for the purification of wild-type CoREST²⁸⁶⁻⁴⁸² because the pIs of CoREST mutants²⁸⁶⁻⁴⁸² were not significantly changed from the pI of wild-type CoREST²⁸⁶⁻⁴⁸².
<table>
<thead>
<tr>
<th>Wild-type CoREST&lt;sup&gt;286-482&lt;/sup&gt;</th>
<th>Molecular weight (g/mol)</th>
<th>pI</th>
<th>Charge at pH 7.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q318A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24586</td>
<td>8.83</td>
<td>3.7</td>
</tr>
<tr>
<td>D320A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24599</td>
<td>9.05</td>
<td>4.7</td>
</tr>
<tr>
<td>D339A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24599</td>
<td>9.05</td>
<td>4.7</td>
</tr>
<tr>
<td>L342A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24601</td>
<td>8.83</td>
<td>3.7</td>
</tr>
<tr>
<td>K346A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24586</td>
<td>8.55</td>
<td>2.7</td>
</tr>
<tr>
<td>I349A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24601</td>
<td>8.83</td>
<td>3.7</td>
</tr>
<tr>
<td>K353A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24586</td>
<td>8.55</td>
<td>2.7</td>
</tr>
<tr>
<td>N356A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24600</td>
<td>8.83</td>
<td>3.7</td>
</tr>
<tr>
<td>K360A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24586</td>
<td>8.55</td>
<td>2.7</td>
</tr>
<tr>
<td>L363A CoREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24601</td>
<td>8.83</td>
<td>3.7</td>
</tr>
<tr>
<td>R371A COREST&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24558</td>
<td>8.55</td>
<td>2.7</td>
</tr>
<tr>
<td>N356AK360A COR&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24529</td>
<td>8.17</td>
<td>1.7</td>
</tr>
<tr>
<td>K353AK360A COR&lt;sup&gt;286-482&lt;/sup&gt;</td>
<td>24543</td>
<td>8.55</td>
<td>2.7</td>
</tr>
</tbody>
</table>

**Isothermal Titration Calorimetry (ITC).** ITC experiments were performed using a MicroCal VP-ITC microcalorimeter (MicroCal, Northampton, MA) as previously described. Approximately 1.5 mL of the degassed LSD1 and buffer were placed in the sample cell and reference cell respectively, and CoREST<sup>286-482</sup> mutants were loaded into the syringe injector. In most ITC experiments, aliquots (3 µL) of CoREST<sup>286-482</sup> mutants at 30 µM concentration were titrated sequentially against LSD1 at 3 µM concentration in the sample cell at 25 °C. Each injection lasted for 6 s, and there was a delay of 300 s between injections. During the titration the stirring speed was 310 rpm and total 40 to 60 injections of CoREST<sup>286-482</sup> mutant were titrated into LSD1. A one site-binding model (Origin 5.0 software, MicroCal software, Inc) was used to fit the data.
Peptide Synthesis and Purification. 20-mer peptides corresponding the sequences of CoREST Linker\textsuperscript{293-380} were synthesized using Fmoc solid phase peptide chemistry with Rink amide SS resin (Advanced Chemtech) on a 0.1 mmol scale with a CEM Liberty microwave peptide synthesizer. The steps for the creation of the peptides on the resin included deprotection of the Fmoc protective group with 20 \% piperidine in DMF, washing with DMF, coupling of activated amino acids in HOBT and HBTU dissolved in DMF, acetylation of unreacted free amino acids with DIEA in NMP, and washing of the resin containing peptides before entering the next cycle. These steps were repeated until the desired sequence was obtained. The peptides were manually cleaved from the resin with a TFA/H\textsubscript{2}O/TIS (95:2.5:2.5) mixture for 3 h at room temperature with occasional stirring. Excess TFA was removed by rotary evaporation, and the peptides were precipitated with ice-cold diethyl ether, suspended in water, and lyophilized to afford the desired crude peptides. Crude peptides were then purified on a Vydac reverse phase octadecylsilica preparative column and the mass was verified by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectroscopy (Applied Biosystems).

Expression and Purification of GST-tagged CoREST\textsuperscript{286-482}. The gene encoding a truncated form of CoREST (residues 286-482) was cloned into the pDEST15 vector that contains glutathione S-transferase (GST)-tag gene by Gateway cloning system, and the resulting plasmid was used for expression and purification. The BL21 (DE3) \textit{E. coli} cells
were transformed with the pDEST15 vector containing the CoREST<sup>286-482</sup> gene, which were grown on a LB agar plate supplemented with ampicillin overnight at 37 °C. Streaks of the colonies on the plate were used to grow 4 L of cells in TB media with shaking (200 rpm) at 23 °C. Gene expression was induced with 0.5 mM of IPTG and cell were allowed to grow overnight and collected by centrifugation at 4225 × g. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, and 0.4 mM PMSF (pH 7.8). GST-tagged CoREST<sup>286-482</sup> was then purified via GSTrapTM 4B column (GE Life Sciences) with elution buffer containing 10 mM reduced glutathione and 50 mM Tris (pH 8.0) and HiPrep 26/60 Sephacryl S200 gel filtration chromatography (GE Life Sciences) with buffer containing 140 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, and 1.8 mM KH<sub>2</sub>PO<sub>4</sub> (pH 7.3). The final concentration of GST-tagged CoREST<sup>286-482</sup> was determined by absorption spectroscopy at 280 nm with its extinction coefficient of 68,300 cm<sup>-1</sup>M<sup>-1</sup> and the protein was stored at -20 °C in 80 % glycerol.

*GST Pull-down In Vitro Inhibition Assay to Disrupt the Binding Interaction between LSD1 and CoREST<sup>286-482</sup>.* Purified GST-CoREST<sup>286-482</sup> was incubated with glutathione-sepharose resin (GE Healthcare) in binding buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.3) for 30 min at room temperature, and the resin was washed extensively with the binding buffer. Then the 20-mer Linker peptide at the indicated concentrations was added, which was followed by the addition of LSD1. The
mixture was incubated for 1 h at room temperature. The LSD1 and peptide complexes were washed off with the binding buffer, and any bound LSD1 to CoREST\textsuperscript{286-482} was eluted with the buffer containing 50 mM Tris and 10 mM reduced glutathione (pH 8.0) by incubating for 20 min at room temperature. The eluted LSD1 was resolved by SDS-PAGE, transferred to a PVDF membrane (Bio-Rad), and analyzed by Western blotting. The PVDF membrane was first blocked with 5 % milk in TBST (5 mM Tris, pH 7.5, 15 mM NaCl, 0.27 mM KCl, 0.1 % Tween 20) at room temperature for 2 h. The membrane was then incubated with 0.1 \( \mu \)g/mL of monoclonal rabbit antibody of LSD1 (ab 17721) at 4 \( ^\circ \)C overnight. After three washing in TBST, 0.02 \( \mu \)g/mL a goat anti-rabbit IgG (H+L)-horseradish conjugated antibody (Bio-Rad 170-6515) was applied to the membranes for 1 h at room temperature. The binding of LSD1 was finally detected by an enhanced chemiluminescence reaction (Amersham Life Science).

\textit{SPOT Synthesis.} Array of 12-mer peptides, covering the sequence of CoREST Linker\textsuperscript{293-380}, were prepared on amino functionalized cellulose membrane by using the SPOT technique, as previously described by the SPOT synthesis method [125]. Hydroxyl groups on a commercially available filter paper (Whatman 50) were esterified with carboxyl groups of Fmoc-protected \( \beta \)-Ala. Following Fmoc deprotection, a layer of \( \beta \)-Ala is formed on the cellulose. At distinct spots, a second Fmoc-protected \( \beta \)-Ala was delivered to define the areas for the peptide synthesis. Following acetylation of all free amino groups, subsequent Fmoc deprotection of the second \( \beta \)-Ala on the defined spots
was followed. β-Ala built up on the membrane is expected to serve as a spacer between the membrane and peptides. On each spot, the 0.5 µL of the pre-activated amino acids at 0.3 M was delivered, and acetylation of unreacted free amino groups, and Fmoc deprotection with piperidine were followed to free amino group for the next coupling. After peptide assembly, the side-chain protecting groups were removed by TFA treatment in the presence of TIPS and water. The binding assay of the cellulose-bound peptides was performed by Western blotting as described above.
5. Inhibition of LSD1 Activity by Disrupting the Binding Interaction with CoREST in ERα-Positive Breast Cancer

5.1 Background

In this chapter, we describe initial experiments of how the disruption of the binding interaction between LSD1 and CoREST with the linker region of CoREST, designated as a Linker peptide, affects the demethylation activity of LSD1 and concomitant cellular processes. As discussed earlier, LSD1 is implicated in various transcriptional regulations by forming complexes with other co-regulatory proteins; thus, the biological functions and outcomes of LSD1 are often tightly connected to the characteristics of complexes in which LSD1 resides.

Recent studies have shown that LSD1 is associated with nuclear receptors, mostly androgen receptor (AR) and estrogen receptor (ER) that govern development, cell growth, physiology, and diseases such as cancer. AR is especially important in the production of secondary sexual characteristics as well as in prostate cancer; ER is known to mediate numerous activities of estrogen that is important in normal development and reproduction as well as breast cancer. In this context, LSD1 is most likely to function upstream of these nuclear receptors by demethylating proximal histones, which enables AR/ER-mediated transcriptions [59, 61]. Interestingly, it was reported that LSD1 is associated with both ERα-negative and positive human breast cancer cells; in ERα-negative breast cancer cells, LSD1 is highly expressed and thus LSD1 serves as a
predictive marker for aggressive breast tumor [60], while in ERα-positive breast cancer cells, LSD1 is shown to be significantly recruited to ERα-target genes [61]. Given the significance of LSD1 in ERα-positive/negative breast cancer cells as previously studied, Dr. Julie A. Pollock in our group utilized small molecule inhibitors of LSD1 to examine the role of LSD1 in especially ER-positive breast cancer cells [133]. Similarly in this study, we use ERα-positive breast cancer cells as a cellular model to probe the effect of the disruption of the binding interaction between LSD1 and CoREST on demethylation activity of LSD1 and ERα-mediated functions.

5.1.1 Breast Cancer

Breast cancer is the most common disease in women worldwide, and it is the number one cause of cancer-related deaths in non-smoking women in the United States (Figure 36) [134]. In 2011, approximately 230,000 new cases of invasive breast cancer were reported in the United States, resulting in an estimated 40,000 deaths [135, 136]. It is a heterogeneous disease with a high degree of diversities between and within subtypes as well as cancer-bearing individuals; each subtype varies in clinical behaviors such as different risk factors for incidence, and risk of disease progression and thus therapeutic implications [137].
Figure 36: Estimated cancer cases and deaths by sex in 2011 (United States); (a) both sexes, (b) male, and (c) female. Blue bar indicates total estimated cancer cases, while red bar indicates resultant deaths. Breast cancer is the highest cause of deaths in women [135, 136].
The growth of breast cancer is closely related to steroid hormones. Particularly, the close link between breast cancer and one of well-known steroid hormones, estrogen, was described for the first time by Beatson in 1896. He used ovariectomy to prevent tumor recurrence, which actually led to the induced regression of the primary tumor [138]. Since the ovaries were sources for steroid hormones, this study hinted at steroid hormones, especially estrogens, being major determinants for disease progression in breast cancer.

Steroid hormones are secreted mainly by three steroid producing glands -the adrenal cortex, testes, and ovaries - and during pregnancy by the placenta. They are essential regulators for key physiological processes such as reproduction, sexual characteristics, metabolism, and response to stress [139, 140]. They all are derived from cholesterol and differ only in the ring structure and side chains attached to it (Figure 37). They can be classified into five classes by the receptors to which they bind: glucocorticoids, mineralocorticoids, androgen, estrogen, and progestins (known as progestagens, progestogens). Gluocorticoid promotes gluconeogenesis and the formation of glycogen and enhances the degradation of fat and protein [139, 140]. Mineralocorticoid, primarily aldosterone, acts on the distal tubules of the kidney to increase the reabsorption of Na\(^+\) and the excretion of K\(^+\) and H\(^+\) that lead to an increase in blood volume and blood pressure [139, 140]. Androgen, which mainly includes testosterone, is responsible for the development of male secondary sex characteristics.
Estrogen, mainly secreted from ovaries, is responsible for the development of secondary sex characteristics in female. Additionally the principal function of estrogen is to cause cellular proliferation and growth of tissues related to reproduction [139, 140, 141]. Progesterone, the most important progestin, participates in the ovarian cycle along with estrogens.

These steroid hormones are typically transported through the bloodstream to the cells of various target organs where they carry out the regulation of wide range of physiological functions as described above. Among these steroid hormones, estrogen has been considered a highly implicated sex steroid hormone in breast cancer and thus its action has been studied in clinical breast cancer for more than 20 years [142, 143].
**Figure 37:** Major steroid hormones.
5.1.2 Estrogen and Estrogen Receptor (ER)

As briefly introduced above, estrogen is essential for normal development and growth of breast and other organs important for reproduction. However, paradoxically, estrogen can be harmful at the same time. Women’s risk for breast cancer is also associated with exposure to estrogen because of its role in stimulating breast cell proliferation, which increases the chance of developing breast cancer. It is well known that the biological effects of steroid hormones are largely mediated through their bindings to structurally and functionally distinct steroid hormone receptors (Table 17), which belong to the nuclear receptor superfamily and normally function as transcription factors [144]. As such, estrogen exerts its effect by binding to estrogen receptors. Estrogen receptor is of special interest because its protein level is elevated in malignant breast lesions as opposed to normal tissue and thus it promotes proliferation of cancerous cells [144]. Furthermore, it has proven to be valuable predictive and prognostic factors in clinical management of breast cancer [145], which suggests that inhibition of ER activity can be one of the major strategies for treatment of breast cancer.
Table 17: Steroid hormones and their respective receptors. The biological effects of estrogen are mediated by binding to estrogen receptors.

<table>
<thead>
<tr>
<th>Steroid Hormones</th>
<th>Steroid Hormone Receptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Androgen</td>
<td>Androgen Receptor</td>
</tr>
<tr>
<td>Estrogen</td>
<td>Estrogen Receptor</td>
</tr>
<tr>
<td>Glucocorticoid</td>
<td>Glucocorticoid Receptor</td>
</tr>
<tr>
<td>Progesterone</td>
<td>Progesterone Receptor</td>
</tr>
<tr>
<td>Mineralocorticoid</td>
<td>Mineralocorticoid Receptor</td>
</tr>
</tbody>
</table>
ER contains several functional domains. These include a DNA binding domain that recognizes specific elements in the promoter regions of target genes called estrogen-responsive elements (EREs) [146]. The C-terminal region of ER contains a ligand (estrogen) binding domain, a hormone-dependent transactivation domain that is important for estrogen-mediated transcriptional activity (AF-2: Activation Function-2), a nuclear localization signal, co-regulator binding sites, and another dimerization domain [134, 146, 147]. A hormone-independent N-terminal activation domain (AF-1) has also been identified, which functions either independently or synergistically, depending on the cellular context (Figure 38) [134].

The human ER is now known to exist as two isoforms: ERα and ERβ. They display a high degree of sequence similarities in the DNA- and C-terminal ligand-binding domain and comparable binding affinities for estradiol (E2), but expression levels and ratios of both receptors vary substantially in many tissues, suggesting that they may have distinct functions. For example, ERα is the predominant ER found in breast, uterus, and mammary gland, whereas ERβ is highly expressed in prostate and gastro-intestinal tract [146]. In breast tumors, ERα has been believed to the predominant ER [134, 146, 147].

To date, the crystal structures of the estrogen-bound ligand binding domain (LBD) and DNA-binding domain of ERα have been resolved. Particularly, ERα LBDs contain twelve α-helices, which are arranged in an antiparallel α-helical sandwich fold.
The helix named as H12 functions as a docking module to recruit various co-regulators (Figure 39) [147].
Figure 38: Functional domains of human estrogen receptors (ERs). Both ERα and ERβ share a highly conserved DNA- and C-terminal ligand binding domain (LBD) except the AF-1 domain. No apparent AF-1 domain is present in ERβ. Numbers in boxes indicate the numbers of residues in each domain [147].
Figure 39: Crystal structure of ERα ligand binding domain (LBD). ERα LBD is arranged in an antiparallel α-helical sandwich fold. Particularly, the helix colored in pink functions as a docking module to recruit various co-regulators. Estrogen is represented as sticks in green. Structural figures were generated using PyMoL (PDB file: 1QKU).
Since estrogen is synthesized from cholesterol and thus soluble in lipid, it can easily diffuse across the membrane of target cells and binds to ERα. Target cells affected by estrogen mostly contain ERα; therefore, when estrogen molecules circulate in the bloodstream and move throughout the body, they exert effects only on the cells that contain ERα. According to the classical model of ERα action, in response to binding of 17-β-estradiol (E2), ERα undergoes a series of sequential events; initially the binding of E2 induces allosteric changes within the ERα, promoting dimerization and permitting the ERα to recognize specific DNA elements (EREs) at promoter regions of target genes. This mechanism activates gene transcription and the synthesis of specific proteins that can influence cell behavior in different ways (Figure 40). Some nuclear proteins, defined as co-regulators, can also interact with the ERα dimer to either positively or negatively regulate the expressions of certain genes [134, 146, 147]
Figure 40: A classical model of ERα action. Estrogen (E2) travels through blood stream and passes through the membrane of target cells. When estrogen molecule enters the cell, it binds to its receptors (ERα), which induces conformational changes within ERα and thus promotes dimerization. The estrogen-ERα complex then binds to specific DNA sites, called estrogen-responsive elements (EREs), near genes that are transcriptionally regulated by estrogens. This mechanism activates gene transcription and then synthesis of specific proteins that can influence cell behavior in different ways.
In breast cancer patients, about two thirds of the tumors rely on estrogen for their proliferation and survival. These tumors are referred to as estrogen-dependent. Tumors that do not rely on estrogen for growth are called estrogen-independent.

Normally estrogen dependence can be predicted by the presence of estrogen receptors in tumor cells; tumor cells that express estrogen receptor α – called ERα-positive – are usually estrogen-dependent. Tumors that lack estrogen receptor α – called ERα-negative – are estrogen-independent. ERα-positive breast cancer is potentially amendable to hormone treatments that are designed to either deplete estrogen or block estrogen signaling. Indeed, anti-estrogen therapy is the first successful cancer therapy targeting ERα expression and functions [148]. Although other important factors such as progesterone receptor status may affect hormone treatment, ERα-related pathway has been considered a major determinant of hormone treatments [149]. However, its efficacy is quite limited only to ERα-positive breast cancers. Up to one-third of breast cancer lacks ERα expression at the time of diagnosis, and unfortunately ERα-negative breast cancers are known to be more aggressive through a still unidentified mechanism [150]. These cancers are often considered to be results of tumor progression from ERα-positive breast cancer [151]; a fraction of cancers that were initially ERα-positive lose the ERα status during tumor progression. Given the limited therapeutic options for ERα-negative cancers, the development of novel therapies targeting other molecules, based on alteration of overexpression of a breast cancer oncogene or reexpression of a breast
cancer suppressor gene, is desperately required. Hereafter we will focus on the ERα-positive breast cancer.

5.1.3 Current Efforts to Modulate ERα Activity

Because activated ERα mainly stimulates cell proliferation in breast cancer, inhibiting their actions has become a major therapeutic goal. Significant ongoing research efforts have focused on modulating ERα activity within ERα-related pathways such as: (1) the inhibition of the synthesis of estrogens; (2) identification of small molecules that act by competing with estrogen for binding in the ERα-binding pocket; (3) inhibition of ERα binding to estrogen response elements (EREs); and (4) inhibition of the interaction between ERα and co-regulatory proteins (Figure 41).
Figure 41: Schematic representation of ERα-related pathways targeted by small molecules and peptides used to block ERα action and breast cancer cell growth. Currently small molecules mainly target estrogen synthesis and competition with estrogen for binding in ERα. Alternative approaches include inhibitions of ERα binding to ERE DNA sites and coregulators.
Aromatase inhibitors lower the level of estrogen in the tumor by targeting the synthesis of estrogen. In postmenopausal women, the ovaries produce only a small amount of estrogen, and the primary estrogen source is derived from the conversion of androstenedione to estrone (E1) and of testosterone to estradiol (E2) in the peripheral tissues including skin and breast [152]. Thus, aromatase inhibitors function to block those conversions. These aromatase inhibitors, however, are not given to premenopausal women because they still have such robust estrogen production in the ovaries, so the inhibitors are effective only in women who are past-menopause. Currently three selective aromatase inhibitors are available in the United States, which includes anastrozole, letrozole, and exemestane (Figure 42).

These inhibitors have often been compared with another class of inhibitors, tamoxifen and raloxifene, commonly known as selective estrogen receptor modulators (SERMs) (Figure 42). They are the first antiestrogens identified and have been currently used in clinical trials. SERMs work by binding to ERα so that estrogen molecules themselves cannot bind to ERα; thus, they impair receptor dimerization and DNA-binding [153, 154, 155]. Additionally they accelerate degradation of the ERα, which has been speculated to occur through the ubiquitin-protease pathway [156]. These inhibitors optimally function as estrogens in bone and the cardiovascular system, but as antiestrogen in the breast and uterus [157]. Over time, however, tumors become resistant to aromatase inhibitors and SERMs, requiring new strategies to inhibit ERα action.
Alternative approaches have been developed by targeting other sites in the pathway. These approaches include an identification of small molecules inhibitors of ERα binding to estrogen-responsive element (ERE) DNA sites. Recent studies have shown that the zinc finger motifs in the ERα DNA binding domain are susceptible to chemical inhibition by disulfide benzamide and benzisothiazolone derivatives [158]. Polyamines such as spermidine and spermine and other small molecules including theophylline, 8-[(benzylthio)methyl]-(7CI, 8CI) (TPBM) screened by a high throughput assay also exhibited inhibitory activities of the ERα binding to ERE sites (Figure 42) [159, 160]. These inhibitors block the subsequent transcription and significantly estrogen-stimulated cell proliferation.

Another alternative approach focuses on development of compounds that interfere binding of ERα with co-regulators. A number of coregulators such as SRC-1/NCoA-1, GRIP-1/TIF-2/NCoA-2, CBP/p300 have been identified and shown to be important for ERα transactivation [147]. All of these proteins contain a signature LXXLL motif (L = leucine, X= any amino acid), which is necessary and sufficient to permit the interaction between ERα and co-regulators. Several groups employ various screening methods such as fluorescence polarization assay, FRET assay, cell-based and computer-based screening assay, and phage display to screen small molecules and peptides to disrupt the interaction of ERα with co-regulators [161, 162, 163]. Small molecules such as pyrimidine derivatives and peptide containing the LXXLL motif have shown to function...
as receptor-specific antagonists, blocking E2-ERα induced transcriptions. Also, indirect methods using molecules that inhibit the activities or expressions of co-regulatory proteins can be used, which we are interested in here.
Figure 42: Structures of ERα antagonists. Exemestane, letrozole, and anastrozole belong to aromatase inhibitors [152]; Tamoxifen and raloxifene compete with estrogen for binding to ERα [153, 154, 155]; TPBM (theophylline, 8-[(benzylthio)methyl]-[7CI, 8Cl]) and DIBA (2,2′-dithiobisbenzamide-1) inhibit ERα binding to ERE DNA sites [159, 160].
5.1.4 Epigenetics in Breast Cancer

As introduced briefly above, ERα does not function in isolation, but rather requires specific cellular factors for maximal responses. Accordingly, an increasing number of co-regulatory proteins has been discovered that bind to ERα, which either stimulates or inhibits activation of target genes, denoted as co-activators and co-repressors respectively. For example, one of the first ERα co-activators is SRC family (steroid receptor coactivator) including SRC-1, SRC-2, and SRC-3 that bind to ERα AF-2 pocket in the ligand-binding domain directly. They have shown to enhance the transcriptional activity of ERα and increase estrogen-induced proliferation of tumor cells [164]. The most-studied corepressors in ERα biology are NCoR (nuclear receptor corepressor) and SMRT (silencing mediator for retinoid and thyroid hormone receptors) that function to enhance repression of target genes [165].

Interestingly, the transcriptional outcomes of ERα are tightly regulated by dynamic interactions of histone-modifying enzymes, which are frequently associated with co-regulatory proteins [166]. Some bind to ERα in direct manner, while others affect ERα activity through indirect interactions by binding to the primary co-regulators. Acetylation and deacetylation of lysine residues in a histone tail have been a good example that shows a mechanism by which ERα communicates histone-modifying enzymes. ERα when bound to estrogen recruits coactivators like SRC-1 or CBP/p300 with histone acetyltransferase activity to estrogen-responsive elements, resulting in
acetylation of histones at or near the site of ER\(\alpha\) binding with consequent open chromatin structure facilitating transcription factors binding to local DNA sequences at promoter regions of estrogen-responsive genes [167]. In contrast, ER\(\alpha\) conversely utilizes NCoR/SMRT corepressor complexes that associate with histone deacetylases that lead to a closed chromatin structure for repression of target genes. In addition to histone acetylation and deacetylation, estrogen-ER\(\alpha\) signaling activates mitogen-activated protein kinase (MAPK) cascades that transmit and amplify signals involved in cellular proliferation [168].

As discussed in Chapter 1, histone methylation and demethylation are crucial for regulating chromatin structure and gene regulation. Particularly, it is a key regulatory signal in ER\(\alpha\)-mediated gene expression. The recent study by Garcia-Bassets, I. et al. showed that LSD1, which demethylates both H3K4 and H3K9, is recruited to significant fraction of ER\(\alpha\) target genes and required to demethylate proximal histone to enable ER\(\alpha\)-mediated transcriptions [61]. However, despite the clear evidence showing LSD1 association in the ER\(\alpha\) activity, the mechanism by which ER\(\alpha\) coordinates LSD1 is poorly understood. It is suspected that other local chromatin modifications like acetylation/deacetylation may play a role in their communication or that other co-regulatory proteins of ER\(\alpha\) assist their crosstalk.

As such, given the significance of histone-modifying enzymes in ER\(\alpha\)-transcriptional activity, it is now clear that their mechanisms or expressions can be
attractive targets to design novel drugs or small molecule inhibitors for treatment of ERα-positive breast cancer. Based on the significance of LSD1 in ERα-positive breast cancer, we are interested in understanding the role of LSD1 through its interaction with CoREST in ERα-positive tumors so as to suggest a new means to regulate ERα-transcriptional activity with modulation of the demethylation activity of LSD1.

Intriguingly, our group first investigated the mRNA expression levels of all of the FAD-dependent amine oxidases including LSD1 in cellular models of breast cancer to determine the expression level of enzymes. Among the nine different FAD-dependent amine oxidases, we observed LSD1 and LSD2 are the most highly expressed across all the breast cancer cell lines including ERα-positive breast cancer cells [133], which allows us to believe that modulating LSD1 activity may be useful in the treatment of ERα-positive breast cancer (Figure 43).
Figure 43: A heat map illustrating expression levels of FAD-dependent amine oxidases in breast cancer cell lines. Red and blue colors indicate high and low expression levels, respectively. Among nine FAD-dependent amine oxidases, LSD1 is the most-highly expressed, suggesting that LSD1 can be a useful target in the treatment of breast cancers [133].
5.2 In vitro Inhibition of the Binding Interaction between LSD1 and CoREST\textsuperscript{286-482} with the Linker peptide

Given the significance of the Linker peptide in binding to LSD1 with an affinity similar to CoREST\textsuperscript{286-482}, the Linker peptide is expected to compete with CoREST\textsuperscript{286-482} for binding to LSD1 and thus occlude LSD1 from proper localization to the nucleosome with consequent inhibition of its demethylation activity in cells. Owing to the difficulty of the synthesis of a long peptide consisting of 80 amino acids, we expressed and purified the Linker peptide in \textit{E.coli} cells as described previously. Prior to the examination of the capability of the Linker in cells, we tested it as an inhibitor to disrupt the binding interaction between LSD1 and CoREST\textsuperscript{286-482} \textit{in vitro}. Expressing such a long peptide in cells most likely allows the Linker peptide to be folded into its native conformation as compared to the synthesis by chemical ligation. As described in Chapter 4, we employed the GST pull-down assay to examine the capability of the Linker peptide as an inhibitor. As expected, the Linker is capable of disrupting the binding interaction between LSD1 and CoREST\textsuperscript{286-482} in a dose-dependent manner, indicating that it competitively binds to LSD1 with complete inhibition at the almost same concentration of LSD1 present in the reaction (Figure 44).
Figure 44: Western blot probing for LSD1 binding to GST-CoREST\textsuperscript{286-482}. The Linker peptide indeed disrupts the binding interaction between LSD1 and CoREST\textsuperscript{286-482} in a dose-dependent manner. The eluted LSD1 was immunoblotted with anti-LSD1.
5.3 In vivo Inhibition of the Binding Interaction between LSD1 and CoREST in ERα-Positive Breast Cancer Cell

5.3.1 Overexpression of CoREST\textsuperscript{293-380} in ERα-Positive Breast Cancer Cells

Initially we set out to express the Linker gene in MCF7 cells, ERα-positive breast cancer cells and examined the change in a global histone methylation level. Transfection efficiency of the vector expressing the Linker peptide was quantified by co-transfection of the vector expressing a GFP protein and flow cytometry. 80 % of transfection efficiency was achieved with the 1:2 ratio of DNA to transfection reagent (X-tremeGENE HP, Roche) for 24 hr incubation at 37 °C, however, the efficiencies were not consistent with repeated experiments. Consistently, we did observe some erratic changes in the methylation levels at H3K4 and H3K9, which is not reliable to conclude the increased methylation level observed. Alternatively, instead of expressing the Linker peptide in cells, the purified Linker peptide can be delivered by aid of carrier tags that bring the peptide into cells, which is followed in the next section.

5.3.2 Inhibition of the Binding Interaction with the Fusion Linker protein with HIV-TAT domain

5.3.2.1 Preparation of the Linker peptide with HIV-TAT domain

In order for the Linker peptide to be delivered into ERα-positive MCF7 cells to disrupt the binding interaction between endogenous LSD1 and CoREST, it needs to be
tethered to particular carrier tags that can translocate through the cellular membrane, thereby enhancing its delivery inside the cell. The ability of such carrier tags to translocate across the membrane is typically confined to short sequences of less than 20 amino acids. Such sequences are called cell-penetrating proteins or peptides (CPPs). CPPs commonly used include transportan and model amphipathic peptide (MAP), polyarginine (Arg), TAT, Antp (homeodomain of Antennapedia) and penetratin [169]. Here for delivery of the Linker peptide into cells, we utilized the TAT peptide that was originally derived from HIV1-TAT protein transduction domain [170]. To this end, initially the gene encoding the Linker peptide was cloned into the pTAT-HA plasmid that was kindly provided by Dr. Steven F. Dowdy in University of California, San Diego. This plasmid, as indicated by its name, particularly contains the genes encoding HIV TAT-derived domain and hemagglutinin (HA) tag (Figure 45-(a) and (b)). The TAT domain (YGRKKRRQRRR) is a cluster of basic amino acids and known to be a minimal transduction domain that mediates protein or peptide delivery into mammalian cells [170]. Ionic interaction between the cationic charges of the TAT peptide and anionic charges of the membrane components is believed to initiate the membrane adsorption of the peptide. The HA tag is used for detection of protein by Western blotting and its antibody.

The Linker-TAT peptide was initially purified via Ni (II) affinity chromatography but undesired proteins (indicated by a blue arrow in Figure 45-(c))
were eluted together with the Linker-TAT peptide even in further several purification steps including size-exclusion chromatography and ion-exchange chromatography. The undesired proteins, however, were largely removed by selective precipitation with ammonium sulfate. In a preliminary test with the cell lysate of 1L culture, the ammonium sulfate concentration was increased stepwise, and the precipitated protein was recovered in PBS buffer and analyzed by 15 % SDS PAGE gel. The Linker-TAT peptide was found to be separated from the undesired proteins with 30 % ammonium sulfate concentration even though it showed some disulfide dimerization (Figure 46-(a)). Based on this observation, we applied the ammonium sulfate precipitation method after the purification via Ni (II)-affinity chromatography. The dimer was reduced to the monomer with addition of 1 mM DTT in PBS buffer, and finally the Linker-TAT peptide was purified to near homogeneity (Figure 46-(b)).
Figure 45: Preparation of the Linker-TAT peptide (18 kDa). (a) The gene encoding the Linker was cloned into pTAT-HA expression vector. (b) Sequence of the Linker-TAT peptide. His$_6$ tag is colored in green, TAT domain in purple, HA tag in blue, and Linker in orange. (c) The Linker-TAT peptide (indicated by a pink arrow) was purified via Ni (II) affinity chromatography. However the undesired proteins indicated by a blue arrow coexisted with the Linker-TAT peptide through several purifications steps such as size exclusion chromatography and ion exchange chromatography.
Figure 46: Selective precipitation of the Linker-TAT peptide (indicated by a pink arrow) with addition of ammonium sulfate ($\text{(NH}_4\text{)}_2\text{SO}_4$). (a) In a preliminary test with the cell lysate, the ammonium sulfate concentration was increased stepwise, and each protein present in the lysate started to precipitate at different concentration of ammonium sulfate. Particularly, the Linker-TAT peptide precipitated at 30% ammonium sulfate concentration and was separated from the undesired proteins (indicated by a blue arrow). We applied the ammonium sulfate precipitation method to polish the Linker-TAT peptide after the purification via Ni (II)-affinity chromatography. (b) The disulfide dimerization formed after the salt addition was reduced to the monomer in addition of 1 mM DTT, and finally the Linker-TAT peptide was purified to near homogeneity.
5.3.2.2 *In vitro* Inhibition of the Binding Interaction between LSD1 and CoREST\textsuperscript{286-482} with the Linker–TAT peptide

Before treatment of MCF7 cells with the Linker TAT peptide, in order to assess the effect of the additional TAT domain on the capability of the Linker peptide, the GST pull-down assay as described earlier was performed. The result of Western blotting analysis indicates that the additional TAT domain does not affect the capability of the Linker peptide to disrupt the binding interaction significantly. The Linker-TAT peptide still inhibits the binding interaction between LSD1 and GST-tagged CoREST\textsuperscript{286-482} with almost same concentration of LSD1 present (Figure 47-(a)). As a positive control, a single point mutation of the Linker-TAT peptide might be ideal. However, as studied in Chapter 4, the single point mutations of the linker region of CoREST still showed tight binding interactions to LSD1, and due to the difficulty of the synthesis of large portion of mutations in the Linker peptide, GFP protein, which is similar with the Linker-TAT peptide in size, was used as a positive control. As expected, GFP protein does not affect the binding interaction between LSD1 and CoREST *in vitro* (Figure 47-(b)).
Figure 47: Western blot probing for LSD1 binding to GST-CoREST$^{286-482}$. (a) The Linker peptide with an additional TAT domain still inhibits the binding interaction between LSD1 and CoREST$^{286-482}$ in a dose-dependent manner almost at the same concentration of LSD1 present. (b) GFP-TAT peptide was used as a control did not affect the binding interaction at all as expected.
5.3.3 Examination of the Global Histone Methylation Levels

Given the significance of CoREST in directing demethylation of LSD1 to specific nucleosomal substrate, H3K4, the impact of the disruption of the binding interaction on this modification was examined in cells. Additionally the change in methylation level on histone H3K9 was also studied because ERα binding in response to estrogen stimulation requires active LSD1 on H3K9 even though the role of CoREST in LSD1 demethylation activity on H3K9 has not been studied. ERα-positive MCF7 cells were treated with various concentrations of the cell-permeable Linker-TAT peptide, and the changes in global levels of H3K4 and H3K9 dimethylation were examined by Western blotting analysis. Initially the transduction of the Linker-TAT peptide in a dose-dependent manner was confirmed by Western blotting with primary antibody for the HA tag sequence. Accordingly as shown in Figure 48, a robust increase in a global dimethylation of H3K4 were observed, which indicates that the disruption of the binding interaction between LSD1 and CoREST obviously influences the inhibition of the demethylation activity of LSD1. Interestingly, the disruption of the binding interaction also affects the global level of H3K9 dimethylation. This indicates that CoREST is implicated in the demethylation activity of LSD1 toward H3K9 to some degree in the presence of AR/ER. Green fluorescence protein (GFP) was used as a control because it does not change the dimethylation levels of both modifications. Our data suggest that the Linker-TAT peptide is specifically inhibiting the binding interaction between LSD1 and CoREST.
This study further suggests a new means to regulate LSD1 activity as compared to small molecule inhibitors targeting the active site of LSD1 and other FAD-dependent amine oxidases. Also since LSD1 is only the amino oxidase that contains the tower domain responsible for binding to CoREST, the disruption of the binding interaction would be specific to the inhibition of LSD1 activity.
Figure 48: Global dimethylation levels of H3K4 and H3K9 in ERα-positive MCF7 breast cancer cells after treatment with various concentrations of the Linker-TAT peptide (0-200 nM). GFP-TAT peptide was used as a control. (a) GFP and (b) Linker peptide were transduced into MCF7 cells in a dose-dependent manner by aids of TAT domain. (c) The methylation levels on H3K4 and H3K9 were not changed with treatment of MCF7 cells with GFP-TAT peptide, but (d) significantly increased methylation levels on both marks were observed with treatment with Linker-TAT peptides. This indicates that disruption of the binding interaction affects the global levels of H3K4 and H3K9 dimethylation and that the Linker-TAT peptide is most likely specific to modulate the binding interaction between LSD1 and CoREST. H3 was used as a loading control. Abbreviation: NT; non-treated, BF; PBS buffer-treated.
5.3.4 Examination of the Stability of LSD1

The crystal structure of LSD1 in complex with CoREST describes that CoREST wraps around the LSD1 tower domain, apparently stabilizing the helical conformation of LSD1. The previous study has confirmed that CoREST regulates LSD1 stability in vivo by preventing the proteasomal degradation of LSD1 [66]. Given the significance of CoREST for LSD1 stability, we examined how the addition of the Linker peptide can affect the LSD1 level. Towards this end, initially ERα-positive MCF7 cells were treated with CoREST siRNA to accomplish knockdown of CoREST, and the cells were immunostained with DAPI, which marks the nucleus, and fluorescently labeled LSD1 antibody. As shown in Figure 49, the knockdown of CoREST resulted in a significant reduction of LSD1 level, indicating that the presence or absence of CoREST influences LSD1 protein level. However, when the cells were treated with 200 nM of the Linker-TAT peptide with CoREST siRNA simultaneously, the LSD1 level was not changed considerably and rather it was increased by approximately 63% as compared to the LSD1 level in siRNA-CoREST treated cells. The analysis of Western blotting also showed that the knockdown of CoREST with siRNA led to the decreased expression of LSD1, but the LSD1 level was stable with the addition of the Linker TAT peptide. This study suggests that the interaction with CoREST, particularly through helical region, is critical for LSD1 stability in vivo. The disruption of the binding interaction would cause the degradation of LSD1 level, but the conservation of the helical conformation of LSD1
with the addition of the helical-structured Linker-TAT inhibitor will contribute to the LSD1 enzymatic activity in other complexes in which LSD1 is implicated.
Figure 49: The binding of LSD1 with CoREST through helical interactions is required for LSD1 stability in vivo. (a) Immunostaining of siRNA CoREST-transfected cells with DAPI, which marks the transfected cells (left column), and LSD1 antibody (right column). The knockdown of CoREST led to a reduction of LSD1 expression, but when the Linker-TAT peptide was added, the LSD1 expression was restored. (b) Fluorescence intensity of LSD1 antibody analyzed for LSD1 levels after addition of the Linker-TAT peptide. 63% of LSD1 was restored as compared to the LSD1 levels in siRNA-CoREST treated cells. (c) Western blotting analysis of the knockdown of CoREST. The knockdown of CoREST decreased LSD1 expression, but the addition of the Linker TAT peptide restored the LSD1 expression. Mock is untreated cell and siMED represents a scrambled siRNA. GAPDH was included as a loading control.
5.3.5 Examination of the Effects of the Disruption of the Binding Interaction in ERα-Positive Breast Cancer

5.3.5.1 Decreased Recruitment of LSD1/CoREST/ER to ERα-Dependent Target Gene Promoter

As mentioned above, upon binding to estrogen, ERα recognizes specific DNA elements (EREs) at promoter regions of target genes where it assembles various co-regulatory proteins to activate gene transcriptions. LSD1 is such a co-regulatory enzyme that is associated with ERα to facilitate the activation of ERα-dependent target genes [61]. Initially we confirmed that LSD1 is recruited to the promoter of two representative ERα-dependent target genes, pS2, a marker for hormone-dependent breast cancer and PR, the gene encoding the progesterone receptor by using chromatin immunoprecipitation (ChIP) assays. We also determined whether CoREST is recruited to the genes where ERα and LSD1 are recruited, which has not been reported yet. Using the same assay, we were able to observe the recruitment of CoREST and ERα to pS2 and PR genes (Figure 50); this recruitment was specific to the ERE binding sites and not observed on the distal sites of the promoters. This study indicates that CoREST is also participated in the LSD1 demethylation activity on H3K9 to a certain degree.
Figure 50: ChIP/qPCR recruitment analysis of ERα, CoREST, and LSD1 on the proximal and distal binding sites of the promoters of ERα target genes, (a) pS2 and (b) PR. CoREST and LSD1 associate with ERα-promoter targets in MCF7 cells. Data is presented as ± SEM for triplicate wells. IgG was used as a negative control.
Then we examined whether the disruption of the binding interaction between LSD1 and CoREST affects the recruitments of LSD1/CoREST/ERα to the target genes by treating MCF7 cells with 200 nM Linker-TAT peptide. As shown in Figure 51, the ChIP/qPCR analysis showed that the recruitments of LSD1/CoREST/ERα to the ERα target genes were markedly decreased. This indicates that inhibition of LSD1 enzymatic activity by disrupting its interaction with CoREST may change the chromatin structure and interfere with proper ERα binding to the EREs of target genes. This observation also corresponds to the results of the inhibition of LSD1 activity by small molecule inhibitors targeting its active site [133]; when the cells were treated with LSD1 small molecule inhibitors, we were able to observe the decreased recruitment of ERα to the promoters of ERα-target genes. Although the exact mechanism still remains elusive, the changes in methylation levels clearly affect the recruitment of ERα to the target genes. The treatment of cells with GFP-TAT peptide did not affect the recruitments of those proteins to the pS2 and PR ERα target genes (Figure 52).
Figure 51: Disruption of the binding interaction between LSD1 and CoREST with the Linker-TAT peptide leads to the decreased recruitments of ERα, CoREST, and LSD1 on the proximal binding site of the promoters of ERα target genes, (a) pS2 and (b) PR. Data is presented as ± SEM for triplicate wells. IgG was used as a negative control.
Figure 52: Disruption of the binding interaction between LSD1 and CoREST with GFP-TAT peptide (control peptide) does not affect the recruitments of ERα, CoREST, and LSD1 on the proximal binding site of the promoters of ERα target genes, (a) *pS2* and (b) *PR*. Data is presented as ± SEM for triplicate wells. IgG was used as a negative control.
5.3.5.2 Decreased Transcription Levels on ERα-Dependent Genes

We next assessed the effects of the decreased recruitments of LSD1/CoREST/ERα by disruption of the binding interaction between LSD1 and CoREST on the ERα-dependent transcription of genes in order to confirm the role of LSD1 through its interaction with CoREST. mRNA levels of endogenous ERα-dependent genes was examined by treating MCF7 cells with the Linker–TAT peptide in dose-dependent manner for 24 hours. The disruption of the binding interaction between LSD1 and CoREST noticeably diminished the transcriptions of ERα-dependent genes in dose-dependent manner. The genes impacted include *pS2, PR, MCM2, CatD, GREB1, WISP2, AMyb, SDF1*, and *SIAH2* (Table 18). In contrast, the genes such as *TIF2* and *1DH3A*, which are known as non-ER target genes, were not affected by the disruption of the binding between LSD1 and CoREST as expected (Figure 53). In previous studies, treatments of MCF7 cells with LSD1 small molecule inhibitors and LSD1 knockdown by siRNA have shown the comparable results [133]. This study indicates LSD1 enzymatic activity through interaction with CoREST is important for the transcriptional activities of ERα.
Figure 53: RT-qPCR analysis after the treatment of ERα-positive MCF7 cells with the Linker-TAT peptides at various concentrations. Some genes were affected by the disruption of the binding interaction between LSD1 and CoREST. Data is presented as ± SEM for triplicate wells.
Table 18: Genes probed by RT-PCR. All information was obtained from NCBI Gene database.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Full name</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>pS2</td>
<td>Trefoil factor 1</td>
<td>Breast cancer marker</td>
</tr>
<tr>
<td>PR</td>
<td>Progesterone receptor</td>
<td>Steroid receptor mediating physiological effects of progesterone</td>
</tr>
<tr>
<td>MCM2</td>
<td>Minichromosome maintenance complex component 2</td>
<td>Involved in the initiation of DNA replication, highly expressed in breast cancer</td>
</tr>
<tr>
<td>CatD</td>
<td>Cathepsin D</td>
<td>Lysosomal aspartyl protease, implicated in breast cancer pathogenesis</td>
</tr>
<tr>
<td>GREB1</td>
<td>Growth regulation by estrogen in breast cancer 1</td>
<td>Early response gene in the ER-regulated pathway</td>
</tr>
<tr>
<td>WISP2</td>
<td>WNT1 inducible signaling pathway protein 2</td>
<td>A member of the connective tissue growth factor/cysteine-rich 61/neuroblastoma overexpressed (CCN) family of growth factor, overexpressed in preneioplastic and human breast cancer cells</td>
</tr>
<tr>
<td>AMyB</td>
<td>V-myb myeloblastosis viral oncogene homolog (avain)-like 1</td>
<td>Transcription activator, its expression dependent on estrogen stimulation</td>
</tr>
<tr>
<td>SDF1</td>
<td>Chemokine (C-X-C motif) ligand 12</td>
<td>Stromal cell-derived alpha chemokine member of the intercrine family, implicated in the metastasis of breast cancer</td>
</tr>
<tr>
<td>SIAH2</td>
<td>Seven in absentia homolog 2</td>
<td>Involved in the regulating the ubiquitination and degradation of synaptophysin</td>
</tr>
<tr>
<td>IDH3A</td>
<td>Isocitrate dehydrogenase 3 (NAD+) alpha</td>
<td>Important for metabolism in brown adipose tissue</td>
</tr>
<tr>
<td>TIF2</td>
<td>Translation initiation factor IF-2</td>
<td>Essential component for the initiation of protein synthesis</td>
</tr>
</tbody>
</table>
5.3.5.3 Decreased Proliferation of ERα-Positive Breast Cancer cells

Previous study showed that LSD1 inhibitors such as 2-PCPA and its derivatives slow down cell growth and proliferation of ERα-positive breast cancer cells, indicating that the inactivation of LSD1 activity impacts on the proliferation of breast cancer cells [133]. In the same manner, the inhibition of LSD1 activity by disrupting its interaction with CoREST should show the similar results on the proliferation of the cells. To this end, MCF7 cells were treated with various concentrations of the Linker-TAT peptide, and their growth pattern was observed over the course of 8 days. As shown in Figure 54, the disruption of the binding interaction with the Linker-TAT peptide showed the reduced rate of cellular proliferation of MCF7 cells in both dose- and time-dependent manners. The treatments of the MCF7 cells for more than 8 days result in cell death. Interestingly, the Linker-TAT peptide was stable for 8 days without any significant degradation in cells (Figure 55). Presumably, the formation of the helical complex with LSD1 might endow the Linker-TAT peptide with capability to resist proteolytic degradation. However, still we treated the cells with the Linker-TAT peptide every other day to keep its level constant during the experiment.

The cell viability assay, which was monitored by CellTiter Blue reagent, supported that the decreased cell proliferation is ascribed not to the cytotoxicity of the Linker-TAT peptide to cells but most likely to the capability of the Linker-TAT peptide to disrupt the binding interaction (Figure 56). The various concentrations of the Linker-
TAT peptide showed little effects on the cell viability; the cells treated with the highest concentration of the Linker-TAT peptide (1 μM) did not show any difference in viability as compared to the viability of the non-treated cells. This study suggests that the enzymatic activity of LSD1 through its interaction with CoREST contributes to the proliferation of breast cancer cells as well as transcriptional activities of ERα as shown above.
Figure 54: Proliferation of ERα-positive MCF7 cells after repeated treatment with indicated concentration of the Linker-TAT peptide over the course of 8 days. Disruption of the binding interaction leads to the reduced rate of cellular proliferation in both dose- and time-dependent manners.

Figure 55: Examination of the stability of the Linker-TAT peptide. The Linker-TAT peptide was stable for 8 days without any significant degradation in MCF7 cells. GAPDH was used as a loading control. D: day.
Figure 56: Viability of ERα-positive MCF7 cells after treatment for 24 h with indicated concentrations of the Linker-TAT peptide. The various concentrations of the Linker-TAT showed little effects on the cell viability, which indicates that cytostatic capability of the Linker-TAT peptide is not ascribed to its cytotoxicity.
5.4 *In vitro* Pull-Down Assay to Investigate Interaction between LSD1 and ERα

Clearly the disruption of the binding interaction between LSD1 and CoREST with the Linker-TAT peptide significantly affects ERα-mediated activities, indicating the importance of LSD1 through its interaction with CoREST. Despite the lack of current information on the implications of CoREST in LSD1/ERα-mediated activities in breast cancer, based on our ChIP experiments, we believe that CoREST coexists in the transcriptional complexes containing LSD1 and ERα by being recruited to the same target genes where LSD1 and ERα are recruited. In order to investigate their physical associations, initially we performed co-immunoprecipitation (IP) experiment by isolating endogenous LSD1, CoREST, and ERα from ERα-positive MCF7 cells and probing their presence by Western blotting. However, we were not able to observe clear signals showing their coexistence. This is presumably because antibodies used to pull-down the complexes inhibit their associations. Recently Garcia-Bassets, I.et al. successfully showed their associations by the same experiment with their own customized antibodies (Figure 57-(a) and (b)) [61], supporting our results of ChIP experiment. Then we employed an *in vitro* pull-down assay to examine their direct interactions. Full length LSD1, truncated LSD1 containing only the SWIRM domain (residues 1-279), CoREST^{286-482}, and control GFP protein were expressed as GST-fusion proteins. After immobilization of these proteins to glutathione resin, $^{35}$S[methionine-
labeled full length of human ERα was incubated with them and washed extensively. Any bound ERα was then eluted and visualized using autoradiography. Interestingly, neither full length of LSD1 nor the SWIRM domain of LSD1 was able to bind to ERα regardless of the presence of estrogen (Figure 57-(c) and (d)), however, intriguingly CoREST showed a binding interaction with ERα, suggesting that CoREST may serve as a bridge between LSD1 and ERα.
Figure 57: Western blot analyses showing (a) immunoprecipitation of LSD1 and coimmunoprecipitation of CoREST from ERα-positive MCF7 cell extracts by using a LSD1 antibody and (b) coimmunoprecipitation of LSD1 by ERα antibody from MCF7 cell extracts [61]. This suggests the potential coexistence of LSD1/CoREST/ERα in ERα-positive breast cancer cells. (c) A 12 % SDS-PAGE gel showing normalized GST fusion proteins used for a pull-down assay (performed by Dr. Julie A. Pollock in our group). Yellow bands indicate the location of the indicated proteins. (d) Radiometric blot showing the binding interaction between CoREST and ERα.
5.5 Discussion and Future Directions

In this chapter, our study shows an importance of LSD1 demethylation activity through its physical and functional interaction with CoREST in ERα-positive breast cancer. LSD1 does not belong to the list of co-regulatory proteins that have been known so far [147], but nonetheless previous studies and our experimental results have proved that its expression and activity play a significant role in ERα-positive breast cancer. The specific mechanisms underlying the implication of LSD1 through CoREST interaction in ERα-positive breast cancer still remain elusive. Originally CoREST is known to direct the demethylation activity of LSD1 toward H3K4 that is often associated with active genes; thus the resultant demethylation accompanies gene repression. In contrast, methylation of H3K9 is associated with silenced genes; hence removal of H3K9 methyl mark in the presence of AR/ER coincides with gene activation. Obviously our study reveals that CoREST is involved in the demethylation at H3K9 to some degree as well as H3K4. Here, we have determined that the disruption of the binding interaction between LSD1 and CoREST affects not only the demethylation activity of LSD1 on both H3K4 and H3K9 but also the ERα-transcriptional activities on target genes and processes required for proliferation of cancer cells. Particularly, since LSD1 is only the amine oxidase that contains a central protruding tower domain, our study suggests that the disruption of the binding interaction between LSD1 and CoREST most likely inhibits LSD1 among FAD-dependent amine oxidases. Based on our observations, we assume that
interrupting the binding interaction will in effect sequester LSD1 and physically occlude it from proper localization to the nucleosome, which makes the enzyme lose its capability to perform the demethylation activity on nucleosomal substrates. This may further lead to changes in the local levels of other post-translational modification marks such as acetylation and phosphorylation in chromatin as well as the methylation level, which presumably prevents the activity of ERα-transcriptional apparatus on its target genes.

Interestingly, unlike our expectation that LSD1 directly binds to ERα through its LKELL (LXXLL) motif in the tower domain, our in vitro pull-down assay presents that CoREST physically associates with ERα regardless of the presence of estrogen. This allows us to hypothesize the role of CoREST as a bridge connecting LSD1 to ERα. Consistently, examination of the primary sequence and crystal structure of CoREST reveals two potential ERα binding motifs, LPEVI sequences between CoREST Linker region and CoREST SANT2 domain and ISDVI sequences in CoREST SANT2 domain, whose sequences are similar to the LXXLL motif (Figure 58). Indeed, some known ERα binding proteins such as NCoR and SMRT utilize their L/I-XX-V/I-I sequences to bind to ERα [171]. These sequences may be the molecular basis for the observed binding interaction between CoREST and ERα, which implies the potential role of CoREST as a scaffolding module in this context. Thus, the likely scenario is that the binding of ERα to specific DNA sites at the promoter of target genes through its
DNA-binding domain recruits CoREST (or CoREST-containing complex) to facilitate the concomitant recruitment of downstream demethylase LSD1, resulting in the demethylation of proximal histone marks, mainly H3K9, to enable ERα-mediated transcriptions. Additionally the recognition of DNA sites may be reinforced through the SANT domains of CoREST. In order to confirm their direct interaction, a competitive assay can be employed as similarly we did in Chapter 4 and Chapter 5 where GST-tagged CoREST\textsuperscript{286-482} and potent competitive proteins incubate together with ERα, especially ER ligand binding domain (LBD), and any bound ERα is eluted with CoREST\textsuperscript{286-482} and examined. Furthermore, the studies of the binding interaction between each domain of CoREST\textsuperscript{286-482}, the Linker region and SANT2 domain, and ER LBD will allow us to predict their binding motif at a molecular level. Nonetheless, there may be other possibilities that the full length CoREST behaves differently from this truncated CoREST (CoREST\textsuperscript{286-482}) in binding to ERα, and that CoREST and LSD1 may associate with ERα through other co-regulatory proteins that primarily bind to ERα in the cellular environment.
Figure 58: Two potential ERα-binding motifs of CoREST286-482 are represented in green color. Figure was generated using PyMoL (PDB file: 2IW5).
The fact that LSD1 is highly expressed in ERα-negative breast cancer allows us to postulate that LSD1 as well as CoREST may serve as basal proteins that regulate transcription levels at various target genes in breast cancer. Previous study has demonstrated that LSD1 serves not only as a biomarker for ERα-negative breast tumors, but also as an alternative, novel therapeutic target in the treatment of ERα-negative breast cancer that is not amendable to hormone therapy [60]. Accordingly several studies were followed to show that small molecule LSD1 inhibitors and polyamine analogs inhibit LSD1 enzymatic activity with resultant enhanced levels of the global methylation level on H3K4 in ERα-negative breast cancer cells [172, 173]. Dr. Pollock in our group has utilized 2-PCPA derivatives to probe their effects on ERα-negative breast cancer cells such as MDA-MB-231 and showed increased methylation level and decreased cell proliferation [133]. In the same manner, we questioned whether the disruption of the binding interaction between LSD1 and CoREST inhibits LSD1 activity in ERα-negative breast cancer cells. To this end, ERα-negative breast cancer cells, MDA-MB-231 cells, were treated with the Linker-TAT peptide, and the global methylation levels on H3K4 and H3K9 and the cellular proliferation were examined. In these assays, we were able to observe the robust increases in the dimethylation levels on both sites and slower proliferation of cells in a dose-dependent manner (Figure 59), which suggests that the disruption of the binding of LSD1 with CoREST is not impacted by the
presence of ERα and can be a novel target for treatments of both ERα-positive and negative breast cancer cells.
Figure 59: (a) Global dimethylation levels of H3K4 and H3K9 in ERα-negative MDA-MB-231 breast cancer cells after treatment with Linker-TAT peptide in a dose-dependent manner. The methylation levels on H3K4 and H3K9 were significantly changed with the addition of Linker-TAT peptide to cells. (b) ERα-negative cell proliferation after repeated treatment with indicated concentration of the Linker-TAT peptide over the course of 8 days. Disruption of the binding interaction leads to the reduced rate of cellular proliferation in both dose- and time-dependent manners. Abbreviation: NT; non-treated, BF; PBS buffer treated.
Furthermore, as mentioned in Chapter 1, LSD1 has been proposed to perform its demethylation activity that requires the intimate collaboration with HDAC 1/2; HDAC 1/2 in the LSD1 complex are likely to function upstream of LSD1 by generating a hypoacetylated substrate that is better recognized by LSD1 [66, 174]. Consistently, the combined treatment with LSD1 and HDAC small molecule inhibitors has indeed shown the cooperation and synergy in regulating gene expression and growth inhibition in ERα-negative breast cancer as well as other solid tumors such as hematological malignancies where they are implicated [173, 174]; these studies showed an increased methylation on H3K4 concomitant with inhibition of nucleosomal deacetylation by HDAC inhibitors such as SAHA and TSA. A likely scenario for this synergistic effect entails the role of CoREST as a bridge between two enzymes and the nucleosome because each enzyme is associated with a different domain of CoREST. Taken together, this led us to speculate with some degree of confidence that the disruption of the binding interaction between LSD1 and CoREST with HDAC inhibitors will generate the similar or more higher synergistic effect in ERα-negative breast cancer cells because targeting the binding interface between LSD1 and CoREST is most likely specific to LSD1 as compared to other LSD1 small molecule inhibitors. At this moment, although we proved that CoREST is involved in ERα-directed demethylation activity of LSD1 on H3K9 and further implicated in ERα-positive breast cancer, we do not know the precise molecular complex in which LSD1 functions with CoREST. There may be other
Accessory proteins in the signaling cascades between LSD1/CoREST and ERα that amplify the effect of the disruption of the binding interaction between LSD1 and CoREST in ERα-mediated breast cancer. Future studies, hence, should fundamentally aim at the identification of other factors such as HDACs in the complex containing LSD1/CoREST/ERα and the nature of their timing in the ERα-mediated signaling, so that we can prove the precise molecular mechanisms underlying their interactions. This could be performed by chromatin immunoprecipitation (ChIP), co-immunoprecipitation (IP) assays, domain deletional study, phage display, or pull-down experiment. Other methodologies such as SPROX (Stability of Proteins from Rates of Oxidation) can be potentially used to map out protein-protein associations in this context.

Intriguingly, Garcia-Bassets, I. et al. suggested that both H3K4 and H3K9 demethylation occur upon estrogen stimulation of LSD1/ERα-dependent gene targets. Although high level of H3K4 methyl mark was still observed after LSD1 binding to ERα, a certain degree of H3K4 demethylase activity was observed upon activation of ERα target genes [61], which indicates that the biological consequences of disrupting the binding interaction in ERα-positive breast cancer may be ascribed to the demethylated H3K4 by LSD1/CoREST. Hence, it is imperative to study the precise molecular mechanisms by which LSD1, CoREST, ERα, and potentially other accessory factors such as HDACs functionally interplay each other in ERα-positive breast cancer in order to
define their roles in the complex. Also how LSD1 changes its demethylase specificity should be studied.

It is noteworthy that the peptide targeting the binding interface between LSD1 and CoREST is potentially not specific to LSD2 despite its structural and catalytic similarities with LSD1 because of the lack of the tower domain in LSD2 essential for the interaction with CoREST helical segment. With knockdown of LSD2, no significant changes in ERα-mediated transcriptional activities and other histone marks such as acetylation were observed [133, 173], which implies that LSD2 is likely to be part of chromatin-remodeling complexes different from the complex containing LSD1 and HDACs.

In summary, our study has verified that modulating LSD1 enzymatic activity by interrupting a critical protein-protein interaction offer a distinctively new avenue of therapeutic intervention in breast cancer via sequestering LSD1 and in turn disrupting ERα recruitment and following transcriptional activities. Moreover, our study does represent that there is a considerable possibility in regulating the ERα-mediated functions by indirectly targeting the binding between ERα and co-regulators using molecules that inhibit the activity and/or expression of co-regulators.
5.6 Experimental Section

Reagents and Materials. Media, antibiotics, and all other buffer reagents were purchased from Sigma, New England Biolabs, MP Biomedicals, BD Biosciences, Millipore, Santa Cruz, Roche, and Abcam. Antibodies (Ab) that recognize LSD1 (ab17721), CoREST (Millipore 07455), Hemagglutinin (HA) (Roche 11867423001), and different histone modifications, namely anti-diMeK4H3 (ab32356) and anti-diMeK9H3 (ab1220), were purchased from Abcam and Millipore. Chemically competent BL21 Star (DE3) *E. coli* cells were purchased from Invitrogen. Chromatographic protein purifications were carried out on an AKTA FPLC (GE Healthcare), and the protein concentration was determined by UV absorbance spectroscopy.

Cell Culture. MCF7 cells were maintained in DMEM/F12 (Gibco) medium supplemented with 8 % fetal bovine serum (FBS) (Sigma), 1 mM sodium pyruvate, and 0.1 mM non-essential amino acids. Cells were grown in a 37 °C incubator with 5 % CO₂.

Purification of the Linker Peptide. The vector encoding the Linker peptide was transformed into electrocompetent BL21 Star (DE3) *E. coli* cells, and the cells were grown overnight as described above for LSD1 purification. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, 0.1 % Triton X-100, and 0.4 mM PMSF (pH 7.8). The linker was then purified via nickel-affinity chromatography with a linear gradient from 50 to 500 mM imidazole in 50 mM sodium phosphate and 300 mM NaCl (pH 7.8) and cation-exchange
chromatography (SP-Sepharose Fast Flow, Sigma) with a linear gradient from 100 to 800 mM NaCl in 50 mM sodium phosphate (pH 7.8). The protein concentration was measured spectrophotometrically at 280 nm with its extinction coefficient of 1,490 cm$^{-1}$ M$^{-1}$.

**Expression and Purification of the Linker with TAT domain (Linker-TAT).** The gene encoding CoREST Linker$^{293-380}$ was cloned into the expression vector of pTAT-HA that was a kind gift from Dr. Steven Dowdy. The coding sequence of CoREST Linker$^{293-380}$ was extracted from CoREST$^{286-482}$ and amplified by using PCR with a forward primer (5’ GCG GGT ACC GTC AAA AAA GAA AAA CAT AGC ACA CAA -3’) and a reverse primer (5’- GCG CTC GAG TTA ATT ACA TTT CTG AAT GAC CTC TGG AGG-3’) under the conditions as previously described. The primers were designed to contain KpnI and XhoI restriction sites at the N- and C-termini respectively to allow for facile ligation into the pTAT-HA vector. The vector was transformed into electrocompetent C43 (DE3) *E. coli* cells, and the cells were grown overnight as described above for LSD1 purification. The protein was purified via nickel-affinity chromatography with a linear gradient from 100 to 700 mM imidazole in 50 mM sodium phosphate and 300 mM NaCl (pH 7.8) and followed by purification via ammonium sulfate precipitation. The desired protein was precipitated with 30 % ammonium sulfate and the precipitated protein was resuspended in PBS buffer. The protein concentration was determined at 280 nm with its extinction coefficient of 13,370 cm$^{-1}$ M$^{-1}$. This protein was further used for *in vivo* studies.
Expression and Purification of GFP protein with TAT domain (GFP-TAT). The gene encoding a GFP protein was extracted from pEGFP-C1 plasmid (Clontech) with a forward primer (5’ AAA GGT ACC ATG GTG AGC AAG GGC 3’) and a reverse primer (5’ AAA CTC GAG TTA CTT GTA CAG CTC GT) and cloned into the pTAT-HA expression vector by utilizing KpnI and XhoI restriction sites at the N- and C-termini respectively to allow for facile ligation. The vector was transformed into electrocompetent C43 (DE3) E. coli cells, and the cells were grown overnight as described above for LSD1 purification. The protein was purified via nickel-affinity chromatography with a linear gradient from 100 to 700 mM imidazole in 50 mM sodium phosphate and 300 mM NaCl (pH 7.4) and followed by purification via HiPrep 26/60 Sephacryl S200 gel filtration chromatography (GE Life Sciences) with a buffer containing 50 mM sodium phosphate, 100 mM NaCl, and 5 % glycerol (pH 7.4). The protein concentration was measured spectrophotometrically at 280 nm with its extinction coefficient of 31,890 cm⁻¹M⁻¹.

Expression and Purification of GST-tagged CoREST²⁸⁶-⁴⁸². The gene encoding a truncated form of CoREST (residues 286-482) was cloned into the pDEST15 vector that contains glutathione S-transferase (GST)-tag gene by Gateway cloning system, and the resulting plasmid was used for expression and purification. The BL21 (DE3) E. coli cells were transformed with the pDEST15 vector containing the CoREST²⁸⁶-⁴⁸² gene, which were grown on a LB agar plate supplemented with ampicillin overnight at 37 °C. Streaks
of the colonies on the plate were used to grow 4 L of cells in TB media with shaking (200 rpm) at 23 °C. Gene expression was induced with 0.5 mM of IPTG, and cell were allowed to grow overnight and collected by centrifugation at 4225 × g. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, and 0.4 mM PMSF (pH 7.8). GST-tagged CoREST286-482 was then purified via GSTrapTM 4B column (GE Life Sciences) with elution buffer containing 10 mM reduced glutathione and 50 mM Tris (pH 8.0) and HiPrep 26/60 Sephacryl S200 gel filtration chromatography (GE Life Sciences) with buffer containing 140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 1.8 mM KH₂PO₄ (pH 7.3). The final concentration of GST-tagged CoREST286-482 was determined by absorption spectroscopy at 280 nm with its extinction coefficient of 68,300 cm⁻¹M⁻¹ and the protein was stored at -20 °C in 80 % glycerol.

GST Pull-down In Vitro Inhibition Assay to Disrupt the Binding Interaction between LSD1 and CoREST286-482. Purified GST-CoREST286-482 was incubated with glutathione-sepharose resin (GE Healthcare) in binding buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.3) for 30 min at room temperature, and the resin was washed extensively with the binding buffer. Then the Linker peptide (or Linker-TAT, GFP-TAT) at the indicated concentrations was added, which was followed by addition of LSD1. The mixture was incubated for 1 h at room temperature. The LSD1 and Linker peptide (or Linker-TAT, GFP-TAT) complexes were washed off with the binding buffer
and any bound LSD1 to CoREST\textsuperscript{286-482} was eluted off from resin with the buffer containing 50 mM Tris and 10 mM reduced glutathione (pH 8.0) by incubating for 20 min at room temperature. The eluted LSD1 was resolved by SDS-PAGE, transferred to a PVDF membrane (Bio-Rad), and analyzed by Western blotting. The PVDF membrane was first blocked with 5 % milk in TBST (5 mM Tris, pH 7.5, 15 mM NaCl, 0.27 mM KCl, 0.1 % Tween 20) at room temperature for 2 h. The membrane was then incubated with 0.1 µg/mL of monoclonal rabbit antibody of LSD1 (Abcam ab 17721) at 4 °C overnight. After three washing in TBST, 0.02 µg/mL a goat anti-rabbit IgG (H+L)-horseradish conjugated antibody (Bio-Rad 170-6515) was applied to the membranes for 1h at room temperature. The binding of LSD1 was finally detected by an enhanced chemiluminescence reaction (Amersham Life Science).

Examination of Methylation Levels on H3K4 and H3K9 by Western blot. When MCF7 cells were confluent to 70 %, the cells were treated with different concentration of the Linker-TAT peptide and grown overnight at 37 °C in 6-well plates in DMEM/F12 medium supplemented as described above. The cells were washed with ice-cold PBS, harvested in PBS, and centrifuged for 1 min. Then the whole cell extracts were isolated using RIPA buffer (50 mM Tris (pH 8.0), 200 mM NaCl, 1.5 mM MgCl\textsubscript{2}, 1 % Nonidet P-40 (NP40), 1 mM EGTA, 10 % glycerol, 50 mM NaF, 2 mM Na\textsubscript{3}VO\textsubscript{4}, and 1x protease inhibitor mixture). Crude histones were extracted from the lysate by resuspending in water and precipitating in 25 % TCA. The precipitants were washed with acetone and
resuspended in PBS buffer. The concentration of resuspended histones was determined using Bio-Rad Bradford assay. The proteins were then resolved by SDS-PAGE and analyzed by Western blotting.

**Transfection Assay.** MCF7 cells were seeded in 6-well plate in DMEM/F12 medium (Gibco) supplemented and transfected with CoREST siRNA (Invitrogen) and DharaFECT1 (Invitrogen) in the absence and presence of the Linker-TAT peptide according to the supplier's protocol. After 2 days, the cells were harvested. The protein levels of the siRNA-treated cells with or without the Linker-TAT peptide were estimated by Western blotting.

**Chromatin Immunoprecipitation (ChIP) Assay.** MCF7 cells were grown in phenol red-free medium until the confluence reached to 90 % and the cells were serum starved for 24 h. After treatment with vehicle (ethanol) or estradiol (E2) at the final concentration of 100 nM for 45 min, the cells were fixed with 1 % formaldehyde, washed with ice-cold PBS, harvested in PBS, and centrifuged for 1 min. The cells were then lysed by sonication in 1 mL sonication buffer (50 mM HEPES (pH 7.8), 140 mM NaCl, 1 mM EDTA, 1 % Triton X-100, 0.1% sodium deoxycholate, 0.1 % SDS, 1X protease inhibitor). The lysate was clarified by centrifugation (15 min, 4 °C, 18000 x g), and precleared with 100 µL protein A/G agarose beads. The precleared lysate was incubated with appropriate primary antibodies for 4-6 h at 4 °C, followed by addition of protein A/G agarose beads overnight at 4 °C. The beads were washed and the precipitates were
eluted with elution buffer containing 50 mM Tris (pH 8.0), 1 mM EDTA, 1 % SDS at 65 °C. The cross-linking was reversed with addition of NaCl, and DNA was isolated with QIA-quick PCR purification kit (Qiagen). qRT-PCR was then performed. The data were normalized to the input for the immunoprecipitation.

**RNA Isolation and RT-PCR.** MCF7 cells were seeded in 12-well plates in phenol red-free medium. When the cells were 70 % confluent, the cells were treated with the Linker-TAT peptide in a dose-dependent manner. After 6 h, the cells were treated with vehicle (ethanol) and estradiol (E2, 100 nM) respectively, incubated for 18 h, and harvested. The total RNA was isolated using the Aurum Total RNA Mini Kit (Bio-Rad) according to the supplier’s protocol. 0.5 µg of total RNA was reverse-transcribed using iScript cDNA synthesis kit (Bio-Rad). cDNAs were then amplified by PCR and quantified by Bio-Rad iCycler iQ Realtime PCR system. The data was normalized to the housekeeping gene, 36B4 and presented as fold induction over control. All data points were collected in triplicate, and the overall assay was run in triplicate.

**Cell Proliferation Assay.** MCF7 cells were seeded at 3000 cells per well in 96-well plates. After 2 days, the cells were treated with Linker-TAT peptide at the indicated concentrations. Every 2 days, over the course of 8 days, the cells were similarly treated. The total DNA contents were measured by fluorescence using Hoechst 33258 dye using a SpectraMax Gemini EM microplate reader (Molecular Devices).
Knockdown of CoREST and Peptide Treatment. MCF7 cells were seeded in 6-well plate in DMEM/F12 medium supplemented as described above and transfected with CoREST siRNA in DharmaFECT 1 (Invitrogen) according to the supplier’s protocol. After 2 days, the cells were harvested. The protein levels of the siRNA-treated cells with or without Linker-TAT peptide were estimated by Western blotting and fluorescence microscope.

Cell Viability Assay. MCF7 cells were seeded in 8000 cells per well in 96-well plates. After 4 days, the cells were treated with the Linker-TAT peptide at various conditions (0.5 μM – 1 mM) for 24 h. Cells were the incubated for 2 h with the addition of CellTiter-Blue reagent (Promega), and the fluorescence was measured (excitation 535 nm, emission 630 nm) using a SpectraMax Gemini EM microplate reader (Molecular Devices). The data were normalized to background with no cells.

Microscopy Assay. MCF7 cells were plated into a 12-well plate that contains sterile circular glass coverslips and incubated until they were 70-80 % confluent. The cells in the indicated wells were transfected with CoREST siRNA according to the supplier’s protocol. The addition of the Linker-TAT peptide was followed. After 48 h, the cells were fixed with MeOH, incubated at -20 °C for 5 min, and washed with PBS buffer. The cells were then blocked with 5 % normal goat serum in PBST (PBS buffer containing 0.1 % Triton X-100). Primary antibodies of CoREST (Millipore 07455) and LSD1 (Abcam 17721) were added to appropriate wells at a dilution of 1:200 and incubated for 1 h at
room temperature. After washing with PBS buffer, the cells were incubated with secondary antibodies (Alexa Fluor 568 goat anti-rabbit IgG (H+L) (Invitrogen, A-11011) at a 1:1000 dilution for 1 h at room temperature. The cells in the indicated wells were counterstained with DAPI (Sigma, D-9542) to locate the nuclei. The appropriately stained cells on the glass coverslips were visualized under the fluorescence microscope (Zeiss Axio Imager).
6. Inducible Expression of the Linker peptide in Xenograft Animal Model

6.1 Background

Given the importance of LSD1 through its interaction with CoREST in mediating ERα-transcriptional activities and processes required for cell proliferation, we decided to conduct xenograft experiments in nude mice in order to shed light on the significance of the modulation of LSD1 activity by disrupting the critical protein-protein interaction and subsequent biological consequences in animal models. To this end, we utilize a lentiviral system to deliver the gene encoding the Linker peptide into MCF7 cells tightly regulated by a tetracycline-inducible system, which will subsequently be xenotransplanted into mice. Expression of the Linker peptide will be conducted in a controlled manner; it should inhibit LSD1 activity in vivo and concomitant ERα-mediated function. We expect this will provide support for LSD1 as a target for ERα-breast cancer.

The lentiviral system is a promising tool currently under preclinical development for gene therapy [175]. Lentiviral infections have some advantages over other gene therapy methods including high-efficiency infection of both dividing and non-dividing cells, long-term stable expression of a transgene, and low immunogenicity [175, 176]. Lentiviruses are typically produced by transfecting 293FT cells with three plasmids including an expression vector carrying the gene to be expressed in target cells, an
envelope vector carrying the gene encoding the proteins to enable the entry of expression construct into target cells, and a packaging plasmid providing all of the proteins essential for transcription and packaging of an RNA copy of the expression construct into the viral particles [176, 177]. Then lentiviruses function as a typical virus does by interacting with the surface of target cells and transferring the genetic information – often referred to as therapeutic genes in gene therapy - into target cells where the viral genes are integrated into the DNA of the target cells in order to be expressed. Here MCF7 cells are infected with lentivirus carrying the genes encoding the Linker peptide, so that MCF7 cells will produce the Linker peptide in vivo (Scheme 6). Expression of the Linker peptide will provide LSD1 antagonist that is specific for events when CoREST is required. We expect that this will inhibit ERα-transcriptional activation and processes required for cell proliferation, which should subsequently suppress tumor growth in mice. Also it is noteworthy that expression of LSD1 antagonist in cells may circumvent the problems of potential toxicity associated with current small molecule inhibitors of LSD1.
Scheme 6: Schematic outline of the xenograft experiment. The 293FT cell line is transfected with three plasmids (an expression plasmid, a packaging plasmid, and an envelope plasmid) for production of lentiviral particles. The lentiviruses are then used to transduce the Tet-controlled MCF7 cells where the viral gene is integrated into the DNA of MCF7 cells to be expressed. This is followed by selection of stably transduced MCF7 cells. The cells should be tested to validate whether they express the Linker peptide in a controlled manner and affect ERα-mediated functions before being xenotransplanted into nude mice. Once transplanted, the volume of tumor will be observed after 7-8 weeks of the treatment with the Linker peptide expressed. Abbreviation: Tet: tetracycline.
For successful treatment of ERα-positive breast cancer with efficient delivery, expression of the Linker peptide needs to be highly regulated in terms of administration, induction, and termination. To achieve this set of conditions, a regulatory system that offers tight control over gene expression in response to inducible agents is valuable. The tetracycline regulatory system, which was developed by Bujard et al. in 1992, offers such substantial regulation of target gene expression in response to tetracycline (Tc) or doxycycline (tetracycline derivative, Dox) that can be repetitively administrated in a time- and dose-dependent manner [178, 179]. This system comprises two complementary control circuits, initially described as the tTA (tetracycline-controlled transactivator protein) dependent and rtTA (reverse tetracycline-controlled transactivator protein) dependent expression system, commonly referred to as the Tet-Off system and Tet-On system, respectively. In the Tet-Off expression system, tTA regulates expression of a target gene that is under control of a tetracycline-responsive promoter element (TRE). In the absence of Tc or Dox, tTA binds to the TRE and activates transcription of the target gene. In the presence of Tc or Dox, which diffuses in their uncharged forms through lipid bilayers of membranes, tTA loses its ability to bind TRE and the expression from the target gene remains inactive [178, 179, 180]. Conversely, in the Tet-On system, transcription of the TRE-regulated target gene is stimulated by rtTA only in the presence of Tc or Dox. rtTA is unbound to the TRE in the absence of Tc or Dox, resulting in an inactive gene, but in the presence of Tc or Dox, rtTA binds to the
TRE and activates the expression of a target gene (Figure 60) [178, 179, 180]. In this manner, MCF7 cells will be engineered by two basic constructs with the first one to stably express Tc or Dox-regulated transactivators, either tTA or rtTA depending on the system of choice, and the second one containing the gene encoding the Linker peptide under the control of TRE. Using the tetracycline regulatory system, we expect the precise control of the timing and expression level of the Linker peptide to be possible. In this chapter, only a general overview of the experiment and initial cloning results will be described.
Figure 60: A tetracycline (Tet)-inducible system. In the Tet-Off system, tTA (tetracycline-controlled transactivator protein) binds to the tetracycline-responsive promoter elements (TRE) and activates target gene expression. In the presence of tetracycline (Tc) or doxycycline (Dox), tTA is unbound and gene expression is turned off. In the Tet-On system, reverse tetracycline-controlled transactivator protein (rtTA) is unbound in the absence of Tc or Dox, resulting in an inactive target gene. In the presence of Tc or Dox, rtTA binds to the TRE and activates the expression of the target gene.
6.2 Generation of Stable Human Doxycycline-Inducible Breast Cancer MCF7 cells Overexpressing the Linker Gene

Initial cloning for production of the lentivirus started with extraction of the genes encoding the Linker peptide from the vector containing genes encoding CoREST\textsuperscript{286-482} by PCR. The gene was then ligated into a pM vector with appropriate restriction enzyme sites in order to utilize the first 147 N-terminal amino acids of a GAL4 DNA-binding domain (DBD) contained in the pM vector. Importantly with respect to the gene delivery approach, the first 74 amino acids of the GAL4 DBD are known to sufficiently confer a nuclear localization signal on peptides or proteins, which makes it a good candidate for use as part of a gene transfer vehicle [181]. The genes encoding the Linker peptide ligated to the genes encoding the GAL4 DBD were then extracted and cloned into the entry vector (pENTR2B) of the Gateway cloning system that allows easy transfer of DNA between different cloning vectors with characteristic Gateway \textit{att} sites. The genes encoding only the GAL4 DBD were isolated separately to be used as a control and cloned into the Gateway system as well. The genes were then transferred to a destination vector, pLenti CMV/TRE3G PuroDEST, from the entry vector by a Gateway LR reaction (Figure 61).
Figure 61: Cloning steps for preparation of the lentiviral plasmid. A picture of 2 % agarose gel shows the PCR product of the genes encoding the Linker peptide extracted from the pET28b vector containing genes for CoREST<sub>286-482</sub> at two different annealing temperatures, 58 °C and 60 °C. The PCR product was then cloned into a pM vector containing a GAL4 DNA-binding domain (DBD). Genes of the Linker and GAL4 DBD were extracted from the pM vector and cloned into the Gateway entry vector (pENTR2B), which were finally transferred to a pLenti CMV/TRE3G PuroDEST destination vector.
The lentiviral plasmid was used to transfec 293FT cells with two accessory plasmids to aid in an efficient delivery of the gene encoding the Linker-GAL4 peptide into the Tet-controlled MCF7 cells. The 293FT cell line derived from human embryonal kidney cells is fast-growing and highly transfecetable, so it is commonly used for generating high-titer lentivirus. Two days after transfection of the 293FT cells, the cell supernatant containing recombinant lentiviral particles was used to transduce the target MCF7 cells. Here we used MCF7 Tet-On cells kindly provided from the McDonnell group in the Department of Pharmacology and Cancer Biology at Duke University. As described above, in this system the Linker-GAL4 gene will be expressed only in the presence of Tc or Dox. We chose to use Dox instead of Tc due to its stability in cells and efficiency at lower concentrations. The infected MCF7 Tet-On cells were then selected for puromycin-resistance to establish a stable cell line for 2 weeks. At the moment, the experiment is still ongoing, and the MCF7 Tet-On cells are being tested to validate whether the Linker-GAL4 peptide is properly expressed and affects ERα-mediated activities. As described in Chapter 5, the changes in global methylation levels on H3K4 and H3K9, ERα-transcriptional activities, and proliferation of cells will be examined as well as expression level of the Linker-GAL4 peptide before the cells are introduced into nude mice. Once validated, the cells will be xenotransplanted into nude mice by Dr. Suzanne Wardell in the Department of Pharmacology and Cancer Biology at Duke
University, and the volume of tumors will be observed 7-8 weeks after treatment with the Linker-GAL4 peptide expressed in the presence of Dox.

6.3 Conclusion and Future Directions

This chapter describes a general overview of the xenograft study to examine how expression of the Linker peptide affects tumor growth in mice. Specifically, we employed a tetracycline-inducible system, which is widely used for controlled expression of target genes (therapeutic genes) in gene therapy [182]. As described earlier, there are two variants: the Tet-Off system and the Tet-On system. The Tet-Off system is negatively controlled and relies on Tc or Dox to deactivate expression, whereas the Tet-On system relies on Tc or Dox to active gene expression. Here we generated MCF7 Tet-On cells infected by lentiviruses carrying the gene encoding the Linker-GAL4 peptide; thus the expression of the Linker-GAL4 peptide will be tightly regulated in both a time-and dose dependent manner with Dox. We expect that the controlled expression of the Linker-GAL4 peptide in this system plays an important role in facilitating the precise determination of the relationship between the expression of the Linke-GAL4 and its pathophysiolocial consequences such as the change in tumor growth. As we observed in Chapter 5, the controlled expression of the Linker-GAL4 peptide is expected to inhibit LSD1 activity and thereby affect the processes required for cell proliferation. To this end, the preliminary study to validate the modified MCF7 cells
with the Tet-On system needs to be completed first. Once validated, the MCF7 Tet-On cells will be xenotransplanted into nude mice in hope of reducing tumor growth in ERα-positive breast cancer. Additionally, since the size of the GAL4 DNA-binding domain (DBD) fused with the Linker peptide is twice as large as the Linker peptide itself, it is important to examine the effect of the GAL4 domain on the inhibitory activity of the Linker peptide. Addition of the GAL4 domain may change the conformation of the Linker peptide, which may change the ability of the Linker peptide to disrupt the binding interaction between LSD1 and CoREST. This experiment was achieved by employing the GST pull-down assay as described in Chapter 4 and Chapter 5. By utilizing the Gateway system, the genes encoding the GAL4 domain and Linker-GAL4 domain in the entry vector (pENTR2B) could be readily transferred to a destination vector, pDEST17, individually for expression in E.Coli cells. The result of Western blotting analysis indicates that the additional GAL4 domain affects the ability of the Linker peptide to disrupt the binding interaction to some degree as it inhibits the binding interaction at a higher concentration than the Linker-TAT peptide (Figure 62). As a positive control, the GAL4-domain alone was used in the pull-down assay. Within the concentration range examined, it did not affect their binding interaction, supporting the idea that the Linker-GAL4 peptide is specific to the LSD1-CoREST286-482 binding.
Figure 62: 15 % SDS-PAGE gels showing (a) the purified GAL4 protein (23 kDa) and (b) the purified Linker-GAL4 protein (33 kDa). Western blot analyses probing for LSD1 binding to GST-CoREST286-482 with addition of (c) the GAL4 alone and (d) Linker-GAL4. The additional GAL4 domain affects the capability of the Linker peptide to disrupt the binding interaction to some degree by inhibiting the binding interaction at a higher concentration than the Linker-TAT peptide (see Figure 47-(a) for comparison). Nonetheless, the Linker-GAL4 peptide specifically inhibits the binding interaction between LSD1 and CoREST286-482. GAL4 was used as a control and did not affect the binding interaction as expected.
6.4 Experimental Section

Reagents and Materials. Media, antibiotics, enzymes, and all other buffer reagents were purchased from Invitrogen, Millipore, New England Biolabs, Cellgro, Santa Cruz, Roche, and Abcam. Antibodies (Ab) that recognize LSD1 (ab17721), CoREST (Millipore 07455), Hemagglutinin (HA) (Roche 11867423001), GAL4 DBD (sc-577), and different histone modifications, namely anti-diMeK4H3 (ab32356) and anti-diMeK9H3 (ab1220), were purchased from Abcam and Millipore. Chemically competent One Shot stbl3 cells were purchased from Invitrogen.

Cell Culture. MCF7 cells were maintained in DMEM/F12 medium (Gibco) supplemented with 8 % fetal bovine serum (FBS) (Sigma), 1 mM sodium pyruvate (Gibco), and 0.1 mM non-essential amino acids (Gibco). 293 FT cells were maintained in DMEM medium (Cellgro) supplemented with 8 % FBS, 1 mM sodium pyruvate, and 0.1 mM non-essential amino acids. Cells were grown in a 37 °C incubator with 5 % CO₂.

Construction of Linker-GAL4-TRE Plasmid. The gene encoding a CoREST Linker293-380 was cloned into the pM vector that was a kind gift from Dr. Donald McDonnell in the Department of Pharmacology and Cancer Biology at Duke University. The pM vector contains the first 147 amino acids of the GAL4 domain. Initially the coding sequence of the Linker peptide was extracted and amplified by using PCR with a forward primer (5’ GCG AGATCT ATG GTC AAA AAA GAA AAA CAT AGC ACA CAA-3’) and a reverse primer (5’- AAA GGA TCC ATT ACA TTT CTG AAT GAC CTC TGG AGG-3’).
under the following conditions: an initial denaturation step for 5 min at 95 °C, 30 cycles of denaturation for 1 min at 95 °C, annealing for 1 min at a gradient of 54-65 °C, elongation for 1:40 min at 70 °C, and followed by a final elongation step for 10 min at 70 °C. The primers were designed to contain BglII and BamHI restriction sites at the N- and C-termini respectively to allow for facile ligation into the pM vector. The sequence-verified genes encoding the Linker-GAL4 peptide were then extracted from the pM vector and cloned into BamHI and XbaI sites of the pENTR2B entry vector. The genes were then inserted into the pLenti CMV/TRE3G PuroDEST destination vector using an LR reaction in the Gateway cloning system. As a control, the genes encoding the GAL4 DBD itself was cut from the pM vector with BglII and XbaI restriction enzyme sites and ligated into BamHI and XbaI sites of the pENTR2B entry vector, which was then followed by the transfer to the pLenti CMV/TRE3G PuroDEST destination vector.

*Generation of Lentiviruses Carrying the Genes Encoding the Linker-GAL4 and GAL4.*

293FT cells were seeded at the density of 2 ×10⁶ cells into a 10 cm culture dish in DMEM medium supplemented with 8 % FBS, 1 mM sodium pyruvate, and 0.1 mM non-essential amino acids and incubated overnight until cells reach ~70 % confluence. The lentiviral vector carrying the Linker-GAL4 gene and packaging vectors were mixed in 400 µL OPTIMEM and incubated for 5 min at room temperature, and an extra 30 min incubation was followed by addition of TransIT-LT1 transfection reagent (Mirus Bio). The transfection mixture was added dropwise to cells. Two days after transfection, the
virus-containing medium was collected, filtered through a 0.45 µM low-protein binding filter, and used to infect MCF7 Tet-On cells (below). The same procedures were applied to generate the lentiviral vector carrying the gene encoding the GAL4.

*Generation of Stable Human Doxycycline-Inducible Breast Cancer MCF7 cells*

*Overexpressing the Linker-GAL4 and GAL4.* The stably transfected MCF7 cells with the pTet-On regulator plasmid were kindly provided from the McDonnell lab in the Department of Pharmacology and Cancer Biology at Duke University. The cells were selected in the presence of 5 µg/mL blasticidin (Invitrogen), which was then infected with lentiviral particles collected above. The transfected cells were then selected with 1 µg/mL puromycin.

*Validation of MCF7 Tet-On Cells.* After the selection, the cells were seeded into appropriate plates as previously described in Chapter 5 to validate whether the Linker–GAL4 and GAL4 peptides are properly expressed and affect ERα-mediated functions. Doxycycline was added to cells every other day to induce the expression of the Linker-GAL4 and GAL4 peptides.

*Expression and Purification of the Linker-GAL4 and GAL4 domain.* The coding sequences of the Linker-GAL4 and GAL4 alone on the pENTR2B entry vector were transferred to a pDEST17 destination vector individually using the LR reaction of Gateway cloning system for expression in *E. coli* strains. Once the sequences were verified, the pDEST17 vectors were transformed into electrocompetent BL21 Star (DE3)
*E. coli* cells, which were grown on a LB agar plate supplemented with ampicillin overnight at 37 °C. Streaks of the colonies on the plate were used to grow 4 L of cells in LB media with shaking (200 rpm) at 23 °C. When the cell density reached an OD_{600} of 0.6, 0.5 mM IPTG was added to the flasks to induce the expression of each protein. The cells were allowed to grow overnight and collected by centrifugation at 4225 × g. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM sodium phosphate, 300 mM NaCl, 5 % glycerol, 0.1 % Triton X-100, 0.4 mM PMSF, and 1X protease inhibitor cocktail solution (pH 7.4). The Linker-GAL4 was then purified via nickel-affinity chromatography with a linear gradient from 50 to 500 mM imidazole in 50 mM sodium phosphate and 300 mM NaCl (pH 7.4), which was followed by purification via cation-exchange chromatography (SP-sepharose Fast Flow, Sigma-aldrich) with a linear gradient from 150 to 1000 mM NaCl in 50 mM sodium phosphate and 5 % glycerol (pH 7.4). The protein was finally polished by HiPrep 26/60 Sephacryl S100 gel filtration chromatography (GE Life Sciences) with buffer containing 50 mM sodium phosphate, 150 mM NaCl, and 5 % glycerol (pH 7.4). Similarly, GAL4 protein was purified via nickel-affinity chromatography and HiPrep 26/60 Sephacryl S100 gel filtration chromatography with the same buffer as described above for the Linker-GAL4 purification. The final concentration of the purified proteins was determined by absorption spectroscopy at 280 nm with the extinction coefficient of 16,500 cm\(^{-1}\)M\(^{-1}\)for
GAL4 and 17, 780 cm$^{-1}$M$^{-1}$ for the Linker-GAL4, respectively. The proteins were flash-frozen in liquid nitrogen and stored at -20 °C until use.

*GST Pull-down In Vitro Inhibition Assay to Disrupt the Binding Interaction between LSD1 and CoREST$^{286-482}$. Purified GST-CoREST$^{286-482}$ was incubated with glutathione-sepharose resin (GE Healthcare) in binding buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, pH 7.3) for 30 min at room temperature, and the resin was washed extensively with the binding buffer. Then the GAL4 and GAL4-Linker peptides were added at the indicated concentrations, which was followed by addition of LSD1. The mixture was incubated for 1 h at room temperature. The LSD1 and peptide complex was washed off with binding buffer, and any bound LSD1 to CoREST$^{286-482}$ was eluted with the buffer containing 50 mM Tris and 10 mM reduced glutathione (pH 8.0) by incubating for 20 min at room temperature. The eluted LSD1 was resolved by SDS-PAGE, transferred to a PVDF membrane (Bio-Rad), and analyzed by Western blotting as previously described in Chapter 4 and Chapter 5.
Appendix A. Expression and Purification of CoREST\(^{53-482}\)

**A.1 Background**

This appendix chapter describes our attempts to express and purify CoREST\(^{53-482}\) that contains all of the functional domains of CoREST. As mentioned throughout previous chapters, CoREST is essential for both the deacetylation activity of HDAC1/2 and demethylation activity of LSD1 towards nucleosomal substrates, indicating that CoREST entails a role as a bridge between two enzymes and nucleosomes. Although HDACs and LSD1 require CoREST for nucleosomal targeting, they seem to associate with CoREST using different strategies. For example, the ELM2 domain of CoREST has shown to associate with HDAC1 in various transcriptional complexes and confer nucleosomal deacetylation activity to HDAC1 [70, 174]. Similar to the role of SANT2 of CoREST in demethylation activity of LSD1, the SANT1 domain serves as a bridge sequence between HDAC1 and its nucleosomal substrates. Particularly, LSD1 requires distinctive regions of CoREST including the Linker region and SANT2 domain for its demethylation activity toward nucleosomes (Figure 63). As such, given the significance of CoREST in mediating the activities of the two enzymes, the mechanistic and enzymatic study of full length of CoREST in associations with LSD1 and HDAC1/2 will provide insights into the nature of their interactions. Towards this end, we initially we set out to express and purify CoREST\(^{53-482}\).
Figure 63: Domain structure of CoREST. CoREST entails its role as a bridge between two enzymes, LSD1 and HDAC1, and nucleosomes. The ELM2 and SANT1 domains are essential for deacetylation activity of HDAC1/2, whereas the Linker region and SANT2 domain are required for LSD1 demethylation activity toward nucleosomes.
A.2 Expression and Purification of CoREST$^{53-482}$

CoREST$^{53-482}$ was overexpressed in several *E. coli* strains in order to determine optimal yield. The cDNA encoding CoREST$^{53-482}$ was cloned into the pET151/D-TOPO and pDEST15 vectors respectively by Dr. Dawn M. Z. Schmidt, a former postdoctoral researcher in our group. The resulting plasmid was transformed into BL21 (DE3), BL21 Star (DE3), BL21 (DE3) codonplus-RIL, Rosetta2 (DE3) *E. coli* strains, and co-transformed with a plasmid expressing the chaperone GroESEL into BL21 (DE3) *E. coli* strain. The cells were grown under various growth conditions, such as differing the concentrations of IPTG added (0.5 mM and 1 mM), varying the temperatures (23 °C and 37 °C) and incubation time after IPTG-induction. Of the strains used, BL21 Star (DE3) *E. coli* cells, which carry an *rne* gene mutation encoding a truncated RNaseE gene that loses the ability to degrade mRNA, exhibited the best expression of CoREST$^{53-482}$ (Figure 64).

Growing cells at higher temperatures (37 °C) led to better expression but yielded poor solubility; CoREST$^{53-482}$ was mostly expressed as insoluble aggregates called inclusion bodies. In contrast, growing cells at lower temperature such as 23 °C and 4 °C seems to increase the solubility of protein, but the overall yield of CoREST$^{53-482}$ was reduced. The expression level and solubility of CoREST in various growth conditions are summarized in Table 19 and Table 20. Interestingly, while running SDS-PAGE gels to confirm the expression of CoREST$^{53-482}$, it was found that CoREST$^{53-482}$ tends to run higher on gels than its molecular weight (53kDa). This is possibly due to the high contents of proline in
the primary sequence of CoREST$_{53-482}$. In fact, CoREST$_{53-482}$ has 23 prolines in its sequence; thus, the rigid CoREST$_{53-482}$ in structure may not migrate through polyacrylamide gel pores as well as fully linearized molecules. To clarify the exact location of CoREST$_{53-482}$ on gels, it may need to be run on a SDS-PAGE gel, which contains low percentage of polyacrylamide.
Figure 64: Pictures of 10 % SDS-PAGE gels showing representative expression levels of CoREST$^{53-482}$ in (a) BL21 Star (DE3) *E.coli* strain, (b) BL21 (DE3) *E.coli* strain, (c) BL21 (DE3)-codonplus RIL *E.coli* strain, (d) Rosetta (DE3) *E.coli* strain, (e) BL21 (DE3) *E.coli* strain with co-transformation of a plasmid expressing the chaperone GroESL, and (f) BL21 (DE3) Star *E.coli* strain. The cells were grown at 37 °C for 4 h after 0.5 mM IPTG induction except (f). (f) Cells were cultured at 4 °C for 7 days with 0.5 mM IPTG induction. Pre: before IPTG induction, 1h-4h: time points after IPTG induction, d1-d7: day points after IPTG induction S: soluble, IS: insoluble. Expressed CoREST$^{53-482}$ is indicated by red arrows.
Table 19: Expression levels of CoREST<sup>53-482</sup> under various culture conditions. The coding sequence of CoREST<sup>53-482</sup> was cloned into pET151/D-TOPO.

<table>
<thead>
<tr>
<th>E.coli strain</th>
<th>Growth Temp.</th>
<th>IPTG Conc.</th>
<th>Growth time</th>
<th>Expression</th>
<th>Solubility</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21 (DE3)</td>
<td>23 °C</td>
<td>0.5 mM</td>
<td>overnight</td>
<td>Y (M)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>overnight</td>
<td>Y (M)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td>37 °C</td>
<td>0.5 mM</td>
<td>4 h</td>
<td>Y (M)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>4 h</td>
<td>Y (M)</td>
<td>IS</td>
</tr>
<tr>
<td>Rosetta2 (DE3)</td>
<td>23 °C</td>
<td>0.5 mM</td>
<td>overnight</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>overnight</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td>37 °C</td>
<td>0.5 mM</td>
<td>4 h</td>
<td>Y (M)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>4 h</td>
<td>Y (M)</td>
<td>IS</td>
</tr>
<tr>
<td>BL21 Star (DE3)</td>
<td>4 °C</td>
<td>0.5 mM</td>
<td>7 days</td>
<td>Y (M)</td>
<td>S?</td>
</tr>
<tr>
<td>BL21 (DE3)</td>
<td>23 °C</td>
<td>0.5 mM</td>
<td>overnight</td>
<td>Y (M)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>overnight</td>
<td>Y (M)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td>37 °C</td>
<td>0.5 mM</td>
<td>4 h</td>
<td>Y (S)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>4 h</td>
<td>Y (S)</td>
<td>IS</td>
</tr>
<tr>
<td>BL21 (DE3) with GroESEL</td>
<td>23 °C</td>
<td>0.5 mM</td>
<td>overnight</td>
<td>na</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>overnight</td>
<td>na</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td>37 °C</td>
<td>0.5 mM</td>
<td>4 h</td>
<td>na</td>
<td>S?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>4 h</td>
<td>na</td>
<td>S?</td>
</tr>
<tr>
<td>BL21 (DE3) Codonplus-RIL</td>
<td>23 °C</td>
<td>0.5 mM</td>
<td>overnight</td>
<td>na</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>overnight</td>
<td>na</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td>37 °C</td>
<td>0.5 mM</td>
<td>4 h</td>
<td>na</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>4 h</td>
<td>na</td>
<td>IS</td>
</tr>
</tbody>
</table>

W: weak; M: moderate; S: strong
S: soluble; IS: insoluble
na: Expression level could not be determined
Table 20: Expression levels of CoREST\textsuperscript{53-482} under various culture conditions. The coding sequence of CoREST\textsuperscript{53-482} was cloned into pDEST15.

<table>
<thead>
<tr>
<th>E.coli strain</th>
<th>Growth Temp.</th>
<th>IPTG conc.</th>
<th>Growth time</th>
<th>Expression</th>
<th>Solubility</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21 (DE3)</td>
<td>23 °C</td>
<td>0.5 mM</td>
<td>overnight</td>
<td>N</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>overnight</td>
<td>N</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td>37 °C</td>
<td>0.5 mM</td>
<td>4 h</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>4 h</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td>Rosetta2</td>
<td>23 °C</td>
<td>0.5 mM</td>
<td>overnight</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td>(DE3)</td>
<td></td>
<td>1.0 mM</td>
<td>overnight</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td>37 °C</td>
<td>0.5 mM</td>
<td>4 h</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>4 h</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td>BL21 Star</td>
<td>23 °C</td>
<td>0.5 mM</td>
<td>overnight</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td>(DE3)</td>
<td></td>
<td>1.0 mM</td>
<td>overnight</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td>37 °C</td>
<td>0.5 mM</td>
<td>4 h</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0 mM</td>
<td>4 h</td>
<td>Y (W)</td>
<td>IS</td>
</tr>
</tbody>
</table>

W: weak; M: moderate; S: strong
S: soluble; IS: insoluble
na: Expression level could not be determined
As illustrated in the expression and solubility tests above, since the majority of CoREST\textsuperscript{53-482} was an insoluble form, the refolding of CoREST\textsuperscript{53-482} from inclusion bodies seemed to be a necessary step to obtain soluble CoREST\textsuperscript{53-482}. A typical procedure for the refolding of proteins from inclusion bodies involves a separation of the insoluble protein by centrifugation, resuspension of the insoluble protein in denaturants such as guanidine or urea, and subsequent refolding by gradually reducing the denaturants (Scheme 7). With high concentration of denaturants, proteins lose their ordered structures that result from intramolecular non-covalent interactions; that is, the denaturants are capable of making direct interactions with side chain residues and backbone peptide bonds of proteins through hydrogen bonding. This mutual influence weakens the intermolecular bonds and interactions of proteins, diminishing the overall secondary and tertiary structure. Then proteins are refolded by gradually removing the denaturants through dialysis. Various additives such as L-arginine and reducing reagents can be used to increase efficiency of the refolding step. After refolding, proteins are expected to adopt appropriate native conformations. To this end, initially we started lysing cell pellets and separating the supernatant from the insoluble CoREST\textsuperscript{53-482} by centrifugation. We then attempted to solubilize and unfold CoREST\textsuperscript{53-482} by addition of 8 M urea in the presence of 1 mM dithiothreitol (DTT) to reduce any disulfide bonds. The refolding of CoREST\textsuperscript{53-482} was followed by dialysis against refolding buffers containing sequentially lower concentrations of urea. Aliquots (1 mL) were removed from each step
of process and analyzed by electrophoresis on 10 % SDS-PAGE gel to show the progress of the preparation of the refolded CoREST\textsuperscript{53-482} (Figure 65).
Scheme 7: A schematic flow chart for refolding process of insoluble proteins.
Figure 65: A picture of 10 % SDS-PAGE gel showing the progress for the preparation of the refolded CoREST\textsuperscript{53-482}. The insoluble CoREST\textsuperscript{53-482} was washed multiple times to remove unwanted proteins associated with or entrapped in the insoluble protein and unfolded by addition of 8 M urea. CoREST\textsuperscript{53-482} was then refolded by gradually removing the denaturant through step dialysis. IS: insoluble CoREST\textsuperscript{53-482}, w1-w3: washing the insoluble protein, UF: unfolded CoREST\textsuperscript{53-482}, RF: refolding CoREST\textsuperscript{53-482}, CRF: concentrated refolded CoREST\textsuperscript{53-482}. A red arrow indicates CoREST\textsuperscript{53-482}. 
The refolded CoREST^{53-482} was further purified via Ni (II)-affinity chromatography on an AKTA FPLC. Since CoREST^{53-482} contains a His<sub>6</sub> tag, it is presumed to bind to the column and be eluted at a certain concentration of imidazole. Approximately 10 mL of the concentrated, refolded CoREST^{53-482} was applied to the column. To track the presence of CoREST^{53-482}, fractions from three major peaks, indicated by colored arrows in Figure 66, were analyzed by 10 % SDS-PAGE. However, fractions from all peaks showed bands that correspond to the molecular weight of CoREST^{53-482}, indicating that CoREST^{53-482} eluted off from the column throughout the entire run (Figure 66). This might be due to the high concentration of either the denaturant or other buffer additives such as L-arginine still remaining in solutions, which may interfere with purification of CoREST^{53-482}. Nonetheless, CoREST^{53-482} was purified to near homogeneity.
Figure 66: A representative AKTA FPLC profile and pictures of 10 % SDS-PAGE gels showing CoREST\(^{53-482}\) purification on chelating sepharose Ni (II) affinity chromatography. A blue line represents UV absorbance at 280 nm as a function of retention volume (mL), and a green line represents a gradient of elution buffer. Arrows in the chromatography indicate the putative peaks where CoREST\(^{53-482}\) might be eluted off from the column. Consistently analysis of 10 % SDS-PAGE gels shows that CoREST\(^{53-482}\) was eluted off from the column throughout the entire run. The protein in the 2\(^{nd}\) lane indicates the concentrated, refolded CoREST\(^{53-482}\) loaded onto the column for purification.
In order to observe whether the refolded and purified CoREST$^{53-482}$ is functionally active and subsequently binds to the purified LSD1, the samples were analyzed via an 8\% native PAGE. CoREST$^{53-482}$ was incubated with LSD1 in a 1:1 molar ratio and subsequently electrophoresed. However, it was hard to tell whether or not there was the formation of complex due to ambiguity of bands in the 4$^{th}$ lane of the gel (Figure 67). The gel bands indicated by orange arrows in Figure 67 presumably suggest the formation of LSD1 and CoREST$^{53-482}$ complex, but they might be the traces of LSD1 impurities seen in the 2$^{nd}$ lane of the gel as well. The amount of the complex formed may be too small to be detected and visualized via this method. Thus, the native PAGE could not clearly prove as to whether the refolded CoREST$^{53-482}$ is functionally active.
Figure 67: A picture of 8% native PAGE gel showing the formation of the LSD1-CoREST$^{53-482}$ complex. lane 1: protein maker; lane 2: LSD1 with molecular weight of 80 kDa; lane 3: CoREST$^{53-482}$ with molecular weight of 53 kDa; lane 4: a LSD1-CoREST complex with molecular weight of approximately 130 kDa. The native PAGE could not clearly prove as to whether the refolded CoREST$^{53-482}$ is functionally active and forms a complex with LSD1.
A.3 Discussion and Future Directions

In this chapter, we briefly present our attempts to express and purify CoREST\textsuperscript{53-482} containing all the functional domains of CoREST. Currently, most of the studies in LSD1 activity with respect to CoREST association focus on the C-terminal region of CoREST (residues 286-482). Previous studies by Shiekhattar and coworkers have examined the effect of full length CoREST in regulating the activity of LSD1 and showed that addition of full length CoREST increased the demethylase activity of LSD1 \cite{71}. This may be ascribed to the role of the SANT1 domain as a bridge of LSD1 to nucleosomal substrates even though the stimulatory activity of SANT1 was weak as compared to that of SANT2. Also, it is possible that the association of LSD1 with CoREST\textsuperscript{53-482} causes conformational changes, which are favorable for substrate binding. Despite the role of the full length CoREST, its functional characterization such as kinetic, thermodynamic, and structural studies has not been fully elucidated. To this end, initial preliminary studies including optimized expression and purification should be completed. Here, we showed that CoREST\textsuperscript{53-482} was overexpressed in BL21 Star (DE3) \textit{E.coli} cells, but under most of culture conditions, it was expressed as inclusion bodies, suggesting a pressing need to refold CoREST\textsuperscript{53-482}. Although the identity of the refolded protein was verified to be CoREST\textsuperscript{53-482} by its sequence analysis, we were not sure whether it was folded correctly. Consistently, the function of refolded CoREST\textsuperscript{53-482} in binding to LSD1 and thus mediating its activity could not be confirmed. We believe the problem with solubility of
CoREST\textsuperscript{53-482} we encountered may account for the reason why significant studies with the full length CoREST have not been performed. Thus, it is necessary to obtain soluble CoREST\textsuperscript{53-482} to circumvent the refolding process. This may be overcome by growing cells at low temperature with more rich TB media. Once the conditions for expression and purification of CoREST\textsuperscript{53-482} are optimized, we believe that the study of the role of CoREST\textsuperscript{53-482} in mediating LSD1 activity will provide insights into molecular mechanisms of regulating the binding interaction between two proteins and their association with HDAC1/2.

\textbf{A.4 Experimental Section}

\textit{Reagents and Materials.} Media, antibiotics, and all other buffer reagents were purchased from Sigma, New England Biolabs, MP Biomedicals, and BD Biosciences. Chemically competent BL21 Star (DE3) \textit{E. coli} cells were purchased from Invitrogen. Chromatographic protein purifications were carried out on an AKTA FPLC (GE Healthcare), and the protein concentration was determined by UV absorbance spectroscopy.

\textit{Expression of CoREST\textsuperscript{53-482}.} The pET151/D-TOPO and pDEST15 plasmids containing the cDNAs coding for CoREST\textsuperscript{53-482} were transformed into several \textit{E. coli} strains and were used to grow cells for protein expression. The colonies grown on LB agar plates supplemented with ampicillin were used to grow 6 mL small-scale cultures
in LB media supplemented with ampicillin. These cultures were used to inoculate with 100 mL of LB media next morning. Cultures were grown with shaking (200 rpm) at various temperatures. When the cell density reached an OD$_{600}$ of 0.6, 100 µL of either 500 mM IPTG or 1M IPTG was added to induce protein expression. The cells were allowed to grow for 3-4 h at 37 °C and overnight at 23 °C, respectively, before harvesting by centrifugation at 4225 x g for 10 min.

*Determinination of Solubility of CoREST$^{53-482}$. The protocol provided by the Bug-Buster Protein Extraction Kit (Novagen) was used to perform the solubility test. Aliquots (1 mL) of cell culture were removed to a 1.5 mL microtube and harvested by centrifugation at 18000 x g for 5 min in the benchtop microcentrifuge. The cell pellets were resuspended in 300 µL of Bug-Buster reagent, and the sample was centrifuged for 15 min at 18000 x g. The supernatant was saved as a soluble fraction and the pellets (inclusion bodies) were washed with 500 µL of 1:10 dilution of Bug-Buster reagent three times. The centrifugation was followed and the final pellet was resuspended in SDS-PAGE loading buffer. The soluble fraction and inclusion bodies in buffer were analyzed by SDS-PAGE.

*Refolding of CoREST$^{53-482}$. The pET151/D-TOPO plasmid containing the cDNA coding for CoREST$^{53-482}$ was transformed into chemically competent BL21 Star (DE3) *E. coli* cells, which were grown on a LB agar plate supplemented with ampicillin overnight at 37 °C. Streaks of the colonies on the plate were used to grow 2 L of cells in LB media
with shaking (200 rpm) overnight at 23 °C. When the cell density reached an OD _{600} of 0.6, 1 mL of 0.5 mM IPTG was added to the flasks to induce protein expression. The cells were allowed to grow overnight and collected by centrifugation at 4225 × g next morning. The cell pellets were lysed with an Emulsiflex C-5 cell cracker in buffer containing 50 mM Tris, 300 mM NaCl, 5 % glycerol, 0.1 mM PMSF, 0.1 % NaN₃, 0.5% Triton X-100, and 1 mM DTT (pH 8.0). 10 mM MgSO₄, 100 mg of lysozyme, and DNaseI were then added to the crude lysate and incubated for 20 min at room temperature, and the centrifugation at 4225 × g for 15 min was followed. Afterwards the supernatant was discarded. The pellets were resuspended in fresh lysis buffer, and the inclusion bodies were recovered by centrifugation. The washing step was repeated three times. After washing, the pure inclusion bodies were dissolved in 20 mL of 50 mM Tris and 50 mM glycine (pH 8.0), which was gradually added to 200 mL of unfolding buffer (50 mM Tris, 50 mM glycine, and 8 M urea, pH 8.0). Unfolding was allowed overnight at 4 °C with constant rapid stirring. Refolding was performed via dialysis against buffers containing lower concentrations of urea; the unfolded protein was transferred to dialysis tubing with 10 kDa MWCO and sequentially placed in buffer with 6 M urea, 4 M urea, 2 M urea, and no urea in 50 mM Tris, 300 mM NaCl, and 1 mM DTT (pH 8.0). The dialysis was allowed for 4 h each buffer. The dialyzed sample was further concentrated to 10 mL by using an Amicon concentrator with a 10 kDa MWCO membrane.
Purification of CoREST$^{53-482}$. The concentrated, refolded CoREST$^{53-482}$ was purified on an AKTA FPLC via Ni (II)-affinity chromatographic column pre-equilibrated with 150 mL of equilibration buffer (50 mM Tris, 300 mM NaCl, and 5 mM imidazole, pH 8.0). The column was washed with 200 mL of buffer containing 50 mM Tris, 300 mM NaCl, and 25 mM imidazole (pH 8.0), which was then followed by elution with a linear gradient of 25-500 mM imidazole. The fractions collected were analyzed by electrophoresis on 10 % SDS-PAGE gel.

Native Gel Electrophoresis. The purified LSD1 and the refolded CoREST$^{53-482}$ were mixed in 1:1 molar ratio and incubated for overnight at 4 °C. The LSD1, CoREST$^{53-482}$, and LSD1-CoREST$^{53-482}$ complex were analyzed via an 8% native PAGE gel at 100 V at 4 °C.
References


272


Biography

Sunhee Hwang was born June 25, 1979 in Seoul, South Korea. She completed two years of coursework in the Department of Chemistry at Dongguk University in Seoul and transferred to the Department of Chemistry and Biochemistry in Arizona State University in 2003. She graduated *magna cum laude* with a Bachelor of Science in Chemistry in December 2005 and was honored on the Dean’s list. While there, she participated in research studying the surface modifications of chip-scale chemical sensors under Dr. Junseok Chae.

She began her graduate career in August 2007 in the Department of Chemistry at Duke University and joined the chemical biology lab of Dr. Dewey McCafferty. For duration of her studies, she was awarded Kathleen Zielek Fellowship and conference travel fellowship to attend the ACS national meeting. She won the first place award for her poster presentation in the 125th North Carolina ACS meeting. She published the journal: Thermodynamic Characterization of the Binding Interaction between Histone Demethylase LSD1/KDM1 and CoREST in *Biochemistry* and will be publishing another paper on the regulation of LSD1 activity by targeting its interaction with CoREST soon. After graduation from Duke University, She will begin a position as a postdoctoral researcher at Stanford University in the lab of Dr. Daria Mochly-Rosen to study the mechanism of protein kinase C-mediated signal transduction.