Rockefeller University
Digital Commons @ RU

Student Theses and Dissertations

2008

Crystal Structure and Mechanism of Lysine Specific Demethylase-1

Pete G. Stravropoulos

Follow this and additional works at: http://digitalcommons.rockefeller.edu/student_theses_and_dissertations

Part of the Life Sciences Commons

Recommended Citation
Stravropoulos, Pete G., "Crystal Structure and Mechanism of Lysine Specific Demethylase-1" (2008). *Student Theses and Dissertations*. 376.
http://digitalcommons.rockefeller.edu/student_theses_and_dissertations/376

This Thesis is brought to you for free and open access by Digital Commons @ RU. It has been accepted for inclusion in Student Theses and Dissertations by an authorized administrator of Digital Commons @ RU. For more information, please contact mcsweej@mail.rockefeller.edu.

CRYSTAL STRUCTURE AND MECHANISM OF LYSINE SPECIFIC DEMETHYLASE-1

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Pete G. Stavropoulos

June 2008

© Copyright by Pete G. Stavropoulos 2008

CRYSTAL STRUCTURE AND MECHANISM OF LYSINE SPECIFIC DEMETHYLASE-1

Pete G. Stavropoulos, Ph.D.
The Rockefeller University 2008

The reversible methylation of specific lysine residues in histone tails is crucial in epigenetic gene regulation. LSD1, the first known lysine-specific demethylase, selectively removes monomethyl and dimethyl, but not trimethyl modifications of Lys4 or Lys9 of histone-3. Here, we present the crystal structure of LSD1 at 2.9-A resolution. LSD1 forms a highly asymmetric, closely packed domain structure from which a long helical 'tower' domain protrudes. The active site cavity is spacious enough to accommodate several residues of the histone tail substrate, but does not appear capable of recognizing the different methylation states of the substrate lysine. This supports the hypothesis that trimethylated lysine is chemically rather than sterically discriminated. We present a biochemical analysis of LSD1 mutants that identifies crucial residues in the active site cavity and shows the importance of the SWIRM and tower domains for catalysis.

ACKNOWLEDGMENTS

This work, described in my thesis, was conducted between March of 2003 and June of 2008 in the Laboratory of Cell Biology at The Rockefeller University. I would like to express my greatest appreciation and gratitude to my thesis advisor, Günter Blobel, for supporting and mentoring me throughout my tenure in his laboratory as a graduate fellow. Gunter is one of the most exceptional scientists, an extremely enthusiastic individual, and unquestionably supportive. Günter possesses many admirable qualities that have left a permanent imprint on how I think about and conduct my research, however it is his ability to draw attention to the global picture that I admire most. His candid ability to identify important questions, formulate ideas and to conceptualize them into a higher ordered and a global perspective has permanently influenced my view and approach to science. I am also appreciative of the trust Gunter displayed to me by allowing me to seek my own path, choose my own direction, and ask/answer the biological questions I had an interest in. It is this freedom with unquestionable support that allowed for my development and cultivation into an independent scientist. I could not ask more of a mentor.

As a student in Günter's lab I was fortunate enough to become acquainted with a postdoctoral fellow André Hoelz. What originated as a casual work/laboratory relationship where we enjoyed conversing about science, politics and daily life, quickly escalated to a friendship and fruitful collaboration. I cannot dilute André s contribution to my success and training and I must also define him as a scientific mentor. André has also influenced my approach to science. Very few possess the eye and taste for quality experiments as André does. His grand approach to science allows one to funnel down to the answer in the quickest possible time and leaves one with little doubt, if any, about the conclusions. His creativity, insightful views and extreme enthusiasm are qualities that I deeply admire.

There are additional individuals I would like to acknowledge. Within the lab, I would like to thank Ivo Melcak with whom I shared many nights in the lab with; Tobias Stuwe, an extremely motivated and bright individual who was a Masters Diploma Student in Gunter's lab; Elias Coutavas who took the time to edit my thesis as well as for his kind gestures which made my transition into the Blobel lab easy; Erik Debler for all the great conversations (and drinks) we have shared as well as the time he allocated to edit my thesis; Alina Patke for editing my thesis; Bernardo Reina San Martin for the great conversations and times we shared before he left for France; and Lourdes Quirolgico and Tita Isberto for administrative support. I also would like to thank Carlos Gomez and Chris Karampahtsis for years of support, friendship and great times.

During the course of ones education one comes across individuals who leave permanent marks and impressions. My first mentor, whom I hold the highest esteem for, Dr. Nouria Hernandez, did just that when I joined her laboratory at Cold Spring Laboratories during my undergraduate studies. I would also like to acknowledge another individual who I hold great admiration for, Dr. Serge Fuchs with whom I worked with at Mount Sinai before beginning my graduate studies.

I would like to express my gratitude to all members of my thesis committee, which includes Dr. Seth Darst, Dr. Nina Papavasiliou and my external committee member, Dr. Danny Reinberg. Dr. Darst, who served as the chairperson of my committee, has been a pleasure to have on my committee and I thank him for all that he has done. Dr. Papavasiliou has been a pleasure to know and extremely supportive since the start of my graduate work at The Rockefeller University. It is always enjoyable and insightful to discuss experiments, ideas and papers with Dr. Papavasiliou. I have had a long standing admiration for Dr. Reinberg's work and I thank him greatly for taking the time out to participate in my thesis defense.

I would like to thank everyone in the Dean's office for all of their help and support throughout the years. A great appreciation also exists for the late Jean Devlin, with whom I became well acquainted with at SUNY Stony Brook University where she was director of the undergraduate pharmacology program that I participated in. I would also like to thank Dr. Sidney Strickland, dean of the graduate program at The Rockefeller University, for all his help and support during my graduate work. I would also like to acknowledge and thank the past and present members of the Dean's office including Kristen Cullen, Marta Delgado, Emily Harms and Cristian Rosario.

Last, but most importantly, I would like to thank all of my family. I am grateful for all the years of support my parents, Fr. George and Angeliki, have given me. Their unwavering efforts to make sure I had all the resources I needed to succeed will always be appreciated. I thank my brother, Terry, for all that he has done including removing my mind from work by listening to his stock picks and economic forecasts. I would like to thank my in-laws, Nick, Kitsa and Lou, for all of their support.

I would like to express my deepest gratitude to my wife, Vasiliki, the mother of my beautiful boy Angelo, who has given me unwavering support but most importantly, has been extremely patient. She has always been there to listen to all the exciting developments, good and bad, that I experienced during my graduate studies.

Thank you…..

TABLE OF CONTENTS

Introduction..... 1

1.1 HISTONE BINDING PROTEINS 4

1.2 HISTONE ARGININE AND LYSINE METHYLATION 11

1.3 HISTONE METHYLTRANSFERASES 13

1.4 DEMETHYLATION OF HISTONE RESIDUES 15

1.5 JMJC DOMAIN HISTONE LYSINE DEMETHYLASES 18

1.6 AMINE OXIDASE DEMETHYLASES 20

1.6.1 BIOCHEMICAL ACTIVITY OF LSD1 20

1.6.2 *In Vitro* ANALYSIS OF HUMAN LSD1 23

1.7 HORMONE DEPENDENT ACTIVATION OF LSD1 27

1.7.1 ANDROGEN RECEPTOR ACTIVATION OF LSD1 27

1.7.2 ESTROGEN DEPENDENT ACTIVATION OF LSD1 29

1.8 LSD1 AND MAMMALIAN DEVELOPMENT 34

1.8.1 ROLE OF LSD1 IN ORGANOGENESIS 34

1.8.2 LSD1 AND LYMPHOID DEVELOPMENT 37

1.9 *S. pombe* LSD1 LIKE AMINE OXIDASES 38

1.10 *D. melanogaster* LSD1 LIKE AMINE OXIDASE 41

1.11 LSD1 AND NON-HISTONE SUBSTRATES 43

1.12 FUNCTIONAL COUPLING BETWEEN 44

DEMETHYLASE

AND DEACETYLASE ACTIVITIES

1.13 KNOWN AMINE OXIDASE INHIBITORS 45

MATERIAL AND **48**

METHODS.....

2.1 PROTEIN EXPRESSION 48

2.2 PROTEIN PURIFICATION 50

2.3 EXPRESSION OF SELENOMETHIONINE-LABLED 56

PROTEINS

2.4 CRYSTALLIZATION AND STRUCTURE 56

DETERMINATION

2.5 ILLUSTRATIONS AND FIGURES 64

2.6 PHOTOMETRIC ACTIVITY ASSAY 64

RESULTS AND **67**

DISCUSSION.....

3.1 OVERALL ARCHITECTURE OF HUMAN LSD1 67

3.1.1 DOMAIN ORGANIZATION OF HUMAN LSD1 67

3.1.2 ARCHITECTURAL OVERVIEW OF THE HUMAN LSD1 CRYSTAL STRUCTURE 67

3.2 THE LSD1 SWIRM DOMAIN 70

3.3 BIOCHEMICAL ANALYSIS OF MUTANTS LOCATED IN THE SWIRM-OXIDASE INTERFACE 78

3.4 THE ACTIVE SITE CAVITY OF LSD1 78

3.5 ENZYMATIC DISCRIMINATION OF TRI-METHYLATED LYSINE 86

3.6 THE TOWER DOMAIN 89

3.7 BIOCHEMICAL CHARACTERIZATION 93

3.8 SUMMARY AND FUTURE DIRECTION 96

REFERENCES..... **99**

LIST OF FIGURES

Figure 1.1	Post-Translational Modifications of Histones	3
Figure 1.2	Structure of Human Double Bromodomain of TAFII250	6
Figure 1.3	Structure of *Drosophila* HP1 Chromodomain	7
Figure 1.4	Structure of the Human 53BP1 Tandem Tudor Domains	9
Figure 1.5	Structure of the Human NURF PHD Finger	10
Figure 1.6	Structure of the *Drosophila* ISWI SANT Domain	12
Figure 1.7	Different Methylation States of Arginine and Lysine Residues	14
Figure 1.8	Reaction Mechanism of the PAD4 Deiminase	16
Figure 1.9	Structure of the Human JMJD2A Catalytic Core	19
Figure 1.10	Proposed Reaction Mechanism of JmjC Domain Containing Demethylases	21
Figure 1.11	Model of LSD1 Activity on Various Substrates	22
Figure 1.12	Potential Reaction Mechanisms of Amine Oxidase Demethylases	24
Figure 1.13	Model of REST Transcriptional Corepressor Complex	25
Figure 1.14	Androgen Receptor Stimulated Activity of LSD1	28
Figure 1.15	Estrogen Receptor Alpha Activation of Transcription	30

Figure 1.16	Model of LSD1-Containing Complexes Involved in Organogenesis	36
Figure 1.17	Model of Functional Activity of *S. pombe* and *D. melanogaster* LSD1-Like Demethylases	40
Figure 2.1	Flowchart of Protein Purification of Human LSD1	49
Figure 2.2	Immobilized Nickel Affinity Purification	51
Figure 2.3	Gel Filtration Chromatography on Superdex 200 16/60	52
Figure 2.4	Gel Filtration Chromatography on Superdex 75 16/60	54
Figure 2.5	Ion Exchange Chromatography on Mono Q 5/5	55
Figure 2.6	Representative Crystal of Human LSD1 and its Diffraction Pattern	57
Figure 2.7	Electron density maps	59
Figure 2.8	Representative 2|Fo|-|Fc| Electron Density	62
Figure 2.9	Ramachandran Plot of the Human LSD1 Model	63
Figure 2.10	Western Blot Analysis of Purified LSD1 Mutants	65
Figure 3.1	Domain Structure of Human LSD1	68
Figure 3.2	Crystal Structure of Human LSD1	69
Figure 3.3	The SWIRM domain of LSD1	71
Figure 3.4	Structure of the SWIRM Domain of LSD1	74
Figure 3.5	Multi-Species Sequence Alignment of LSD1 Homologs	76

Figure 3.6	Biochemical Analysis of LSD1 Mutants Located in the SWIRM-Oxidase Interface	79
Figure 3.7	Structure of the Oxidase Domain and Active Site Cavity of LSD1	80
Figure 3.8	Complex Crystal Structure of LSD1 Bound To Substrate H3 Peptide	85
Figure 3.9	Trimethylated H3-K4 Peptide Acts as a Competitive Inhibitor of LSD1 Activity	87
Figure 3.10	Proposed Chemical Mechanisms of LSD1 Demethylation Activity	88
Figure 3.11	Allosteric Regulation of LSD1	91
Figure 3.12	Structure of the Hetero-Dimeric Human LSD1-CoREST Complex	92
Figure 3.13	Biochemical Analysis of LSD1 Mutants	94

LIST OF TABLES

Table 1	Crystallographic Analysis	60
Table 2	Crystallographic Analysis	61

Introduction

In eukaryotes, genomic DNA exists as a dynamic complex of nucleic acid and protein, which forms a compact structure termed chromatin. Chromatin formation is established by the efforts of numerous proteins, but largely by a set of proteins termed histones. Histones are small globular proteins that are highly basic in charge with flexible, ~40-residue-long N-terminal tails. Although there are a number of different histones, four core histones seem to dominate the landscape of eukaryotic genomic DNA. An octamer of histones, namely a H3/H4 tetramer which assembles with two H2A/H2B dimers, interacts with 147 base pairs of helical DNA to form the fundamental unit of chromatin, the nucleosomal particle [1]. In this nucleosomal assembly, the DNA is wrapped around the histone octamer, constituting a total of 1.75 turns of the superhelical structure, with the flexible N-terminal histone tails protruding from the complex [1]. The general belief is that the nucleosomal units are organized to form higher ordered chromatin structures [2].

Biochemical analysis has revealed that histones undergo an array of post-translational modifications which are highly concentrated on the flexible N-terminal tails, although a number of residues within the histone fold have also been found to be modified. The dynamic post-translational modifications of histones include acetylation, adeno-ribosylation, phosphorylation, ubiquitination, sumoylation, and methylation, resulting

from the interplay of opposing enzymes (Figure 1.1). Examples of such antagonistic sets of enzymes are acetylases:deacetylases for acetylation, kinases: phosphatases for phosphorylation, and methyltransferases: demethylases for methylation.

The packaging of genomic DNA into chromatin restricts access to the DNA for a variety of biological events, such as replication, recombination, DNA repair, and transcription. The fundamental question of how the nucleic acid of the complex is accessed arises when one examines the restrictive nature of chromatin. To begin to address this question, the "Histone Code Hypothesis" was set forth [3-5]. The central paradigm of the histone code hypothesis is that post-translational modifications of histones act as recognition motifs by creating or removing affinities for proteins, also refered to as effectors, involved in particular chromatin-associated processes [3-5]. Various combinations of modifications within a given nucleosome may exist, and these modifications may be combinatorial [3-5]. The presence and concentration of particular modifications within distinct locations in chromatin may be responsible for the epigenetic states present within a cell [4, 6].

The structure and mechanism of the first lysine demethylase discovered, Lysine Specific Demethylase -1 (LSD-1) is the subject of this thesis.

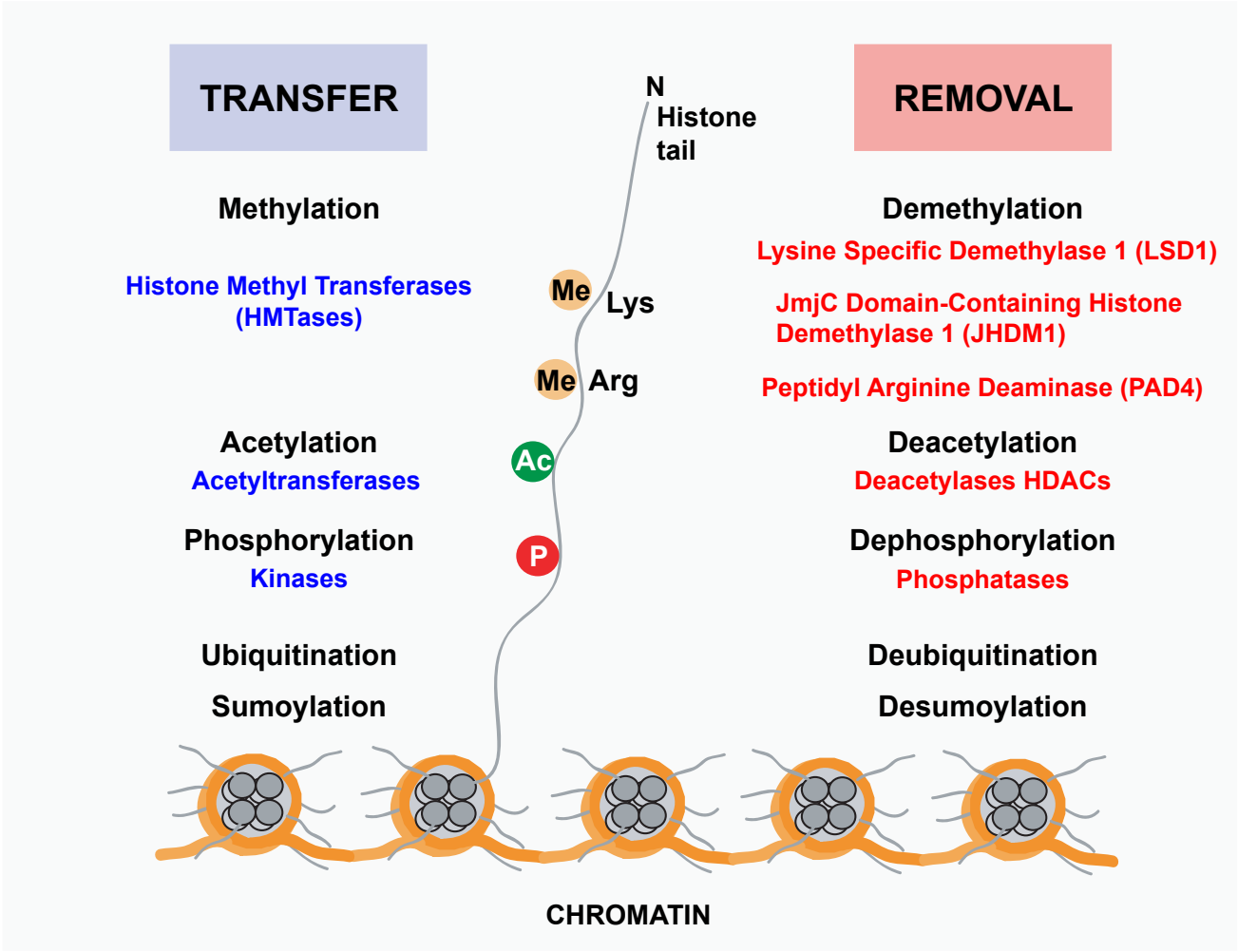


Figure 1.1 Post-Translational Modifications of Histones

A number of histone modifications have been described. These modifications are dynamic; antagonizing sets of enzymes catalyze the addition and removal of these modifications.

1.1 Histone Binding Proteins

Covalent modification of histones creates binding sites for various effector proteins [3-5]. The specific modification of different histone residues has profound consequences for the state and function of chromatin. Biochemical analysis of a variety of transcriptional repressor and activator complexes has revealed a number of protein motifs and domains that are shared between various complexes. Some of these domains contain enzymatic activity, whereas others do not. The most common domains found in effector proteins include the bromodomain, chromodomain, tudor, plant homeodomain (PHD) finger, and the SWI3, ADA2, N-CoR, and TFIIIB (SANT) domain. A brief description of these most commonly found domains known to bind modified histones is given below. Within a given co-activator or repressor complex, these domains will be found in various combinations, which allows one to envision how a sequential or combinatorial readout could occur.

The bromodomain is an evolutionarily conserved domain found in a number of proteins involved in chromatin-associated processes and was first identified in a *D. melanogaster* protein called brahma [7, 8]. Furthermore, the bromodomain is also found in many histone acetyltransferases [9], where it mediates protein-protein interactions by specifically binding to acetyl-lysine [10, 11]. Bromodomains adopt a conserved structural fold of a left-handed bundle of four helices αZ, αA,

αB, αC (Figure 1.2) [12,13]. The ZA and BC inter-helical loops are of variable length and sequence. The BC loops constitute a hydrophobic pocket that serves to stabilize the structure, and are responsible for the binding to the acetyl-lysine [13].

The chromodomain was first identified in the two *D. melanogaster* proteins, HP1 and Polycomb, which are involved in regulating chromatin structure [14]. The chromodomain is present in a large number of proteins known to be involved in chromatin biology, including helicases, histone acetyltransferases, and histone methyltransferases [15]. Several chromodomains have been shown to directly interact with methylated histones. For example, the chromodomain of HP1 interacts with methylated H3K9, a marker for heterochromatin which is involved in maintaining the chromatin in a silent state and, possibly, in the formation of higher ordered chromatin structures [16, 17]. Another example is CHD1, a helicase involved in nucleosomal assembly and mobilization, which contains two chromodomains that recognize methylated H3K4 through their concerted effort [18]. Structure determination of HP1 has revealed a small, three-stranded antiparallel β-sheet with an α-helix that runs across the sheet (Figure 1.3) [19-22]. This structure is held together by a hydrophobic core formed by conserved hydrophobic amino acid residues. Another domain present in chromatin regulatory proteins is the tudor domain. The double tudor domain of histone demethylase JMJD2A binds

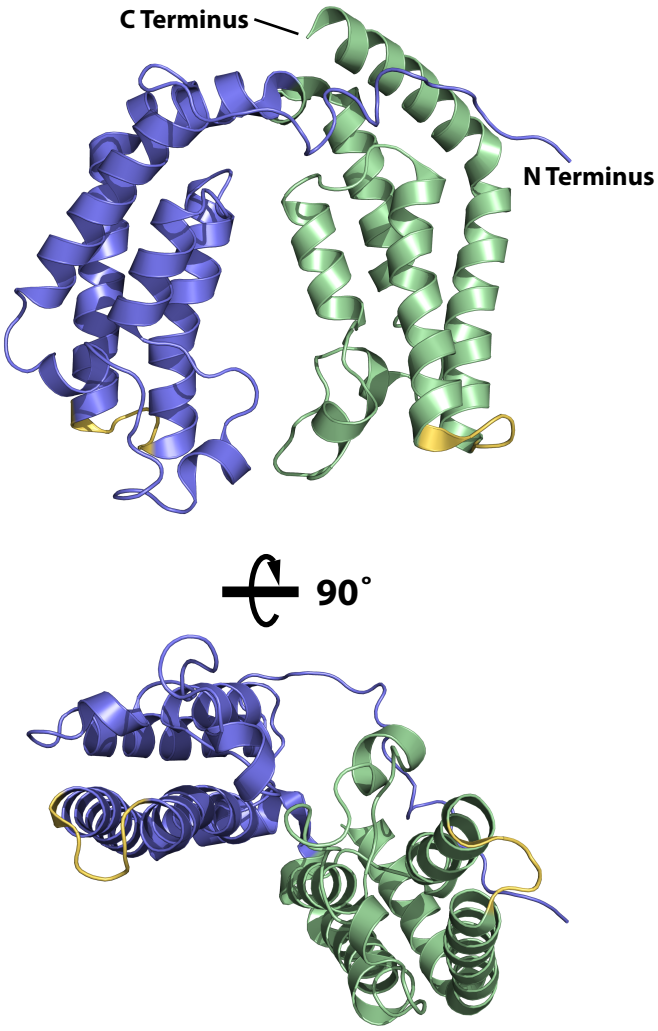


Figure 1.2 Structure of Human Double Bromodomain of TAFII250

Ribbon representation of the crystal structure of the human double bromodomain module of TAFII250 (PDB accession code 1EQF). Each bromodomain consists of a conserved left-handed bundle of four a-helices. Each bromodomain is differentially colored. One lobe of the double bromodomain is colored blue whereas the other is colored green. The BC loops, which are commonly used for protein-protein interactions, are colored yellow. TAFII250 uses the BC loops to bind to histone H4 N-terminal tails that are acetylated at multiple lysine residues (K5, K8, K12 and K16). Two different orientations of TAFII250 are displayed.

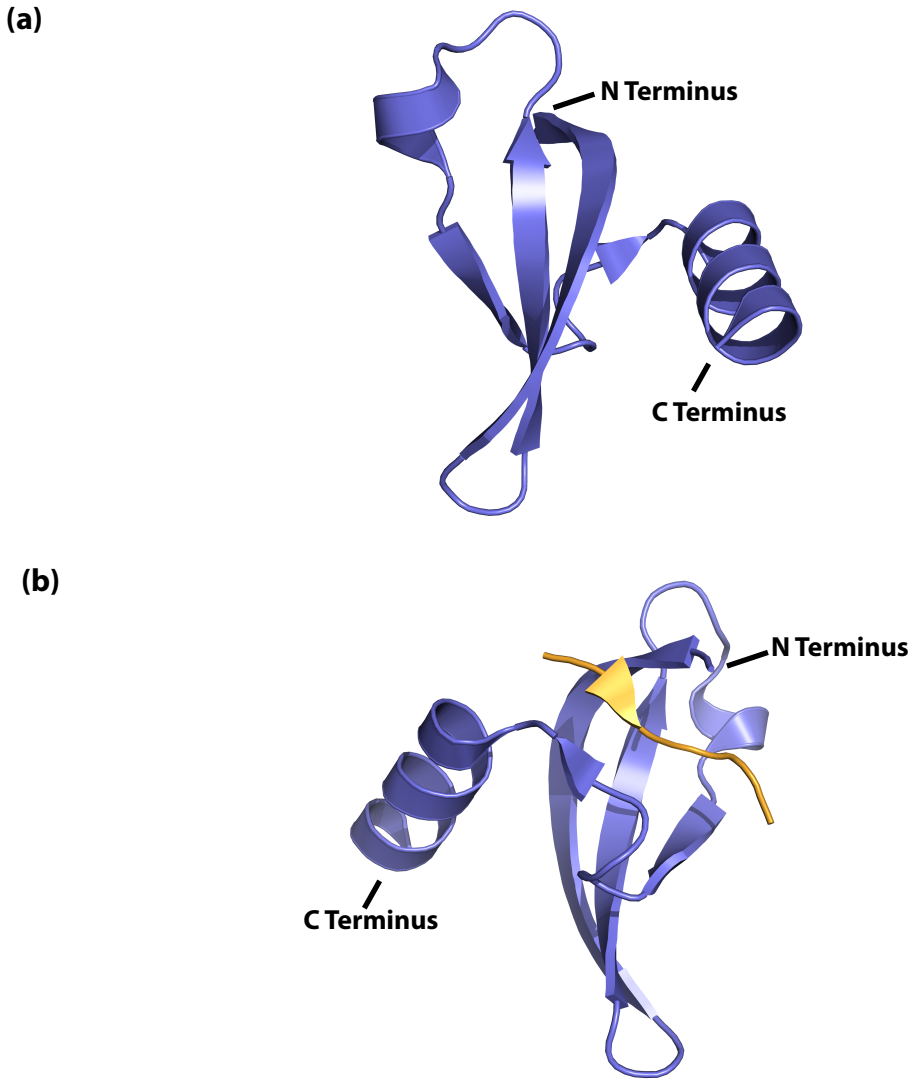


Figure 1.3 Structure of D. melanogaster HP1 Chromodomain

(a) A ribbon representation of the crystal structure of D. melanoganster HP1 chromodomain is shown (PDB accession code 1KNA). The three stranded antiparallel ß-sheet which interacts with an α-helical segment constitutes the conserved core of a chromodomain. (b) A ribbon representation of the crystal structure of D. melanoganster HP1 chromodomain is shown in blue, in complex with the first 15 residues of histone H3 di-methylated at lysine 9 shown in yellow.

preferentially to tri-methylated H3K4 [23, 24]. Interestingly, tudor domains of the spinal muscular atrophy gene product SMN, the splicing factor 30 kDa (SPF30), and the Tudor domain-containing 3 (TDRD3), have been shown to bind to arginine-glycine-rich motifs in a methyl-arginine-dependent manner [25]. 53BP1 is a DNA damage checkpoint protein which contains a double tudor domain. *In vitro* studies have shown that the double tudor domain of 53BP1 binds to di-methyl lysine 79 of Histone H3 (2meH3K79). Structural studies of the double tudor domain of 53BP1 revealed that each tudor domain consists of a five stranded β-sheet that forms a β-barrel (Figure 1.4) [26].

The role of PHD fingers in histone binding has been elucidated recently. The PHD finger from the nucleosome remodeling factor (NURF) has been shown to bind methylated H3K4 with a preference for the tri-methylated state [27]. Biochemical analysis of ING2, a PHD containing protein which is a member of the mSin3a-HDAC1 histone deacetylase complex [28], revealed its ability to bind tri-methyl H3K4 [29]. The PHD finger forms a small compact globular structure with a ββα motif as the only common structural element of the core. Each PHD finger binds two zinc atoms that are coordinated by loops that bear a Cys_4-$HisCys_3$ zinc binding motif. The two zinc binding sites are separated by the two stranded β-sheet. The β-sheet is followed by a small α-helix which provides the histidine and cysteine to coordinate the first zinc ion (Figure 1.5) [30].

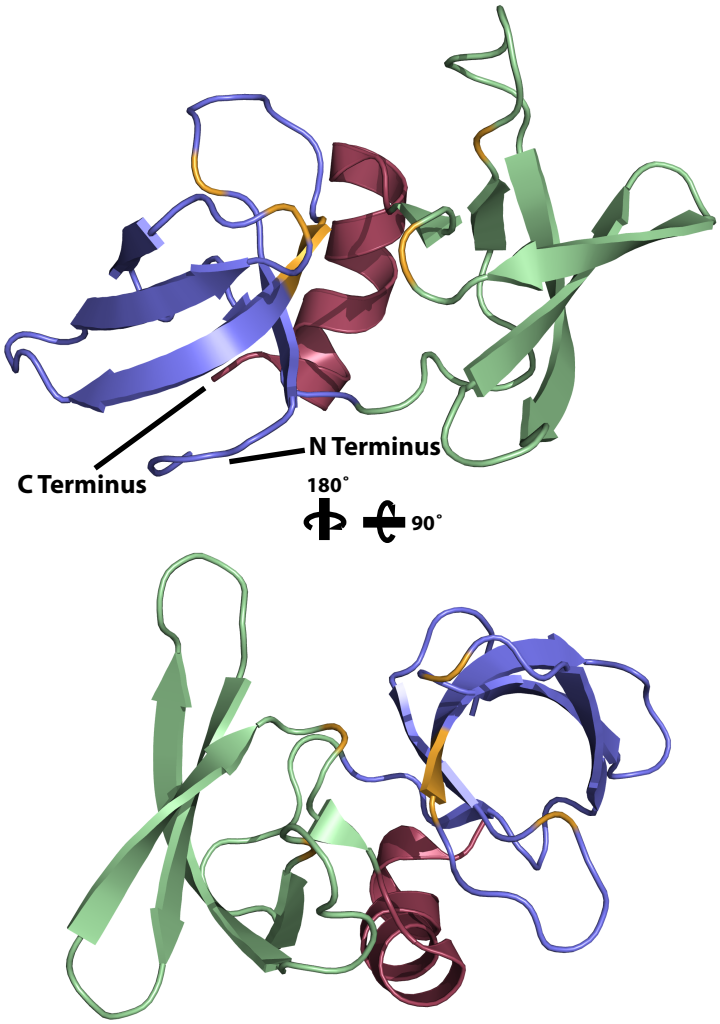


Figure 1.4 Structure of the Human 53BP1 Tandem Tudor Domains

The ribbon representation of the crystal structure of the human 53BP1 tandem tudor domains is shown in two different orientations (PDB accession code 1XNI). The tudor domain contains a five stranded ß-sheet which adopts a ß-barrel structure. The two tudor domains of 53BP1 are differentially colored. One lobe of the double tudor domain is colored in blue, the second in green, and an extra α-helical segment in red. The portion of the double tudor domain of 53BP1 which is believed to bind di-methyl lysine 79 of histone H3 is colored in orange.

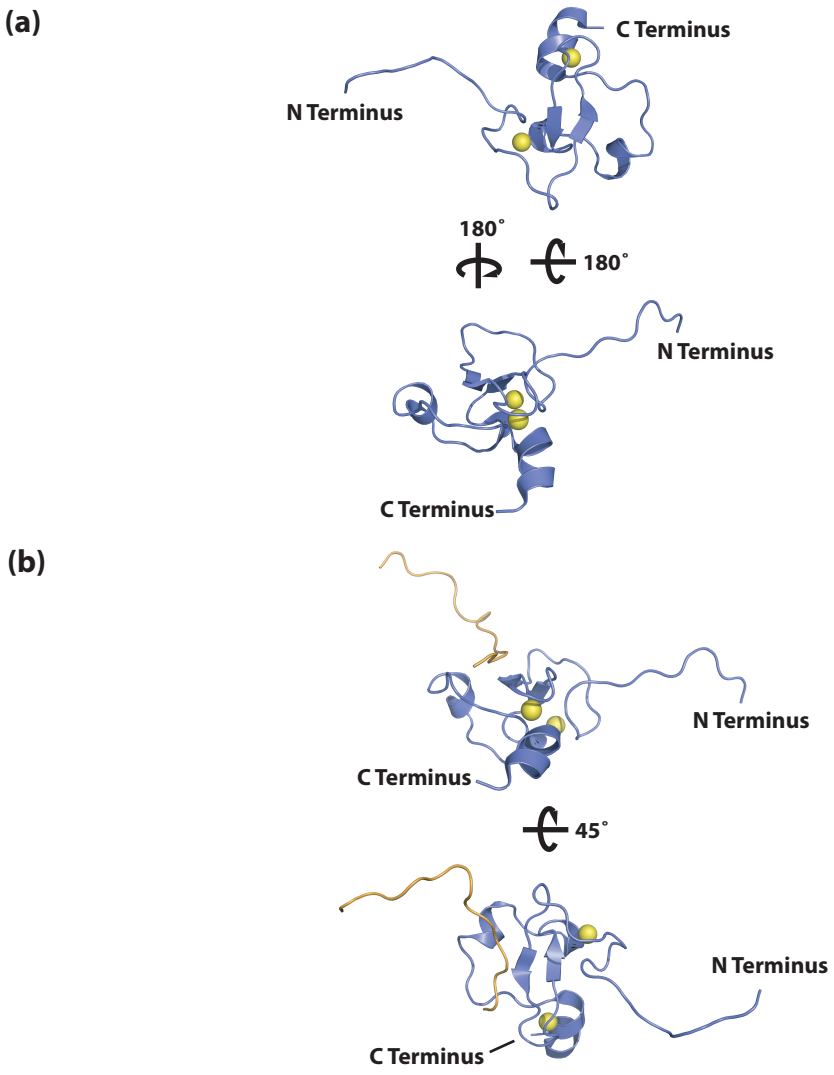


Figure 1.5 Structure of the Human NURF PHD Finger

(a) The ribbon representation of the NMR structure of the PHD finger from human NURF is shown in two different orientations (PDB accession code 2FUU). The core element of a PHD finger consists of two zinc ions, a small two stranded ß-sheet followed by an a-helix. The two zinc ions, coordinated by a Cys_4-$HisCys_3$ zinc binding motif, are seen as yellow spheres (b) The ribbon representation of the NMR structure of the PHD finger from human NURF is shown in complex with the N-terminal tail of histone H3, which is tri-methylated on lysine 4. Two different orientations are shown where the PHD finger is colored in blue, the histone peptide is colored orange, and the two zinc ions are seen as yellow spheres.

The SANT domain was originally identified as a motif present in nuclear receptor co-repressors [31]. The SANT domain has a high similarity to the DNA binding domain (DBD) of Myb related proteins, although the residues required for DNA binding are not well conserved [32]. Two nuclear receptor co-repressors, SMRT and NCoR, are involved in recruiting histone deacetylase activity to chromatin. Both SMRT and NCoR contain two closely spaced SANT domains which have been shown to bind histones H3 and H4 [33]. SANT domains consist of three α-helical segments that are arranged in a helix-turn-helix motif (Figure 1.6) [34]. A conserved large bulky aromatic residue is present in each of the α-helices which form a hydrophobic core [34]. Structural studies have shown that DBD domains of Myb proteins and SANT domains share overall structural similarities however, key residues required for DNA binding are not conserved in the SANT domain [32].

1.2 Histone Arginine and Lysine Methylation

Patterns of histone methylation and their biological consequences are not as simple to correlate as are other modifications, such as acetylation and transcriptional activation. However, a general trend seems to associate areas of transcriptionally active chromatin with methylated lysine residues at H3K4, H3K36, and H3K79, whereas methylation of histones in

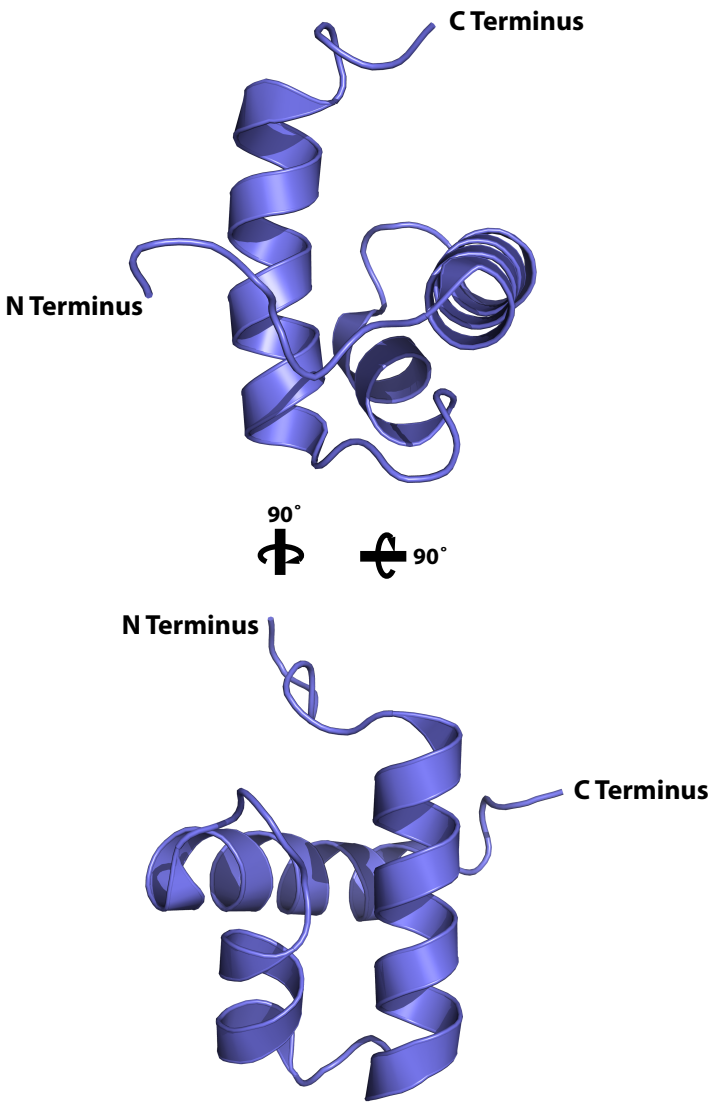


Figure 1.6 Structure of the D. melanoganster ISWI SANT Domain

The ribbon presentation of the crystal structure of the SANT domain from Drosophila ISWI shown in two different orientations (PDB accession code 1OFC). The helix-turn-helix motif is the conserved structural element of SANT domains.

transcriptionally silenced chromatin seems to occur on H3K9, H3K27, and H4K20 [35].

Arginine residues have been shown to be either mono-methylated or di-methylated on the terminal amino groups of the guanidinium moiety (Figure 1.7a) [36]. Di-methylated arginine modifications are further classified into symmetric and asymmetric modifications, depending on whether both methyl groups are attached to the same or different terminal amino groups. Methylated lysine residues exist in the mono-, di-, or tri-methylated state (Figure 1.7b). The majority of the methylated residues exist in the N-terminal tails of histones, although a number of residues that are located within the histone fold are also subject to methylation [37].

1.3 Histone Methyltransferases

Arginine methylation is carried out by arginine methyltransferases, such as PRMT1 and CARM1 [38-41]. S-adenosyl-L-methionine (AdoMet) is utilized as the methyl group donor, which is subsequently transferred to the guanidinium moiety of arginine. Lysine methylation is carried out by two types of lysine methyltransferases (HKMTs). The predominant type of HKMT contains an ~130- residue domain termed SET, which was originally discovered in three *D. melanogaster* genes: Su(var)3-9, En(zeste), and Trithorax [42]. The second type of HKMT does not contain a SET domain, but rather belongs to a family of methyltransferases termed Class-I

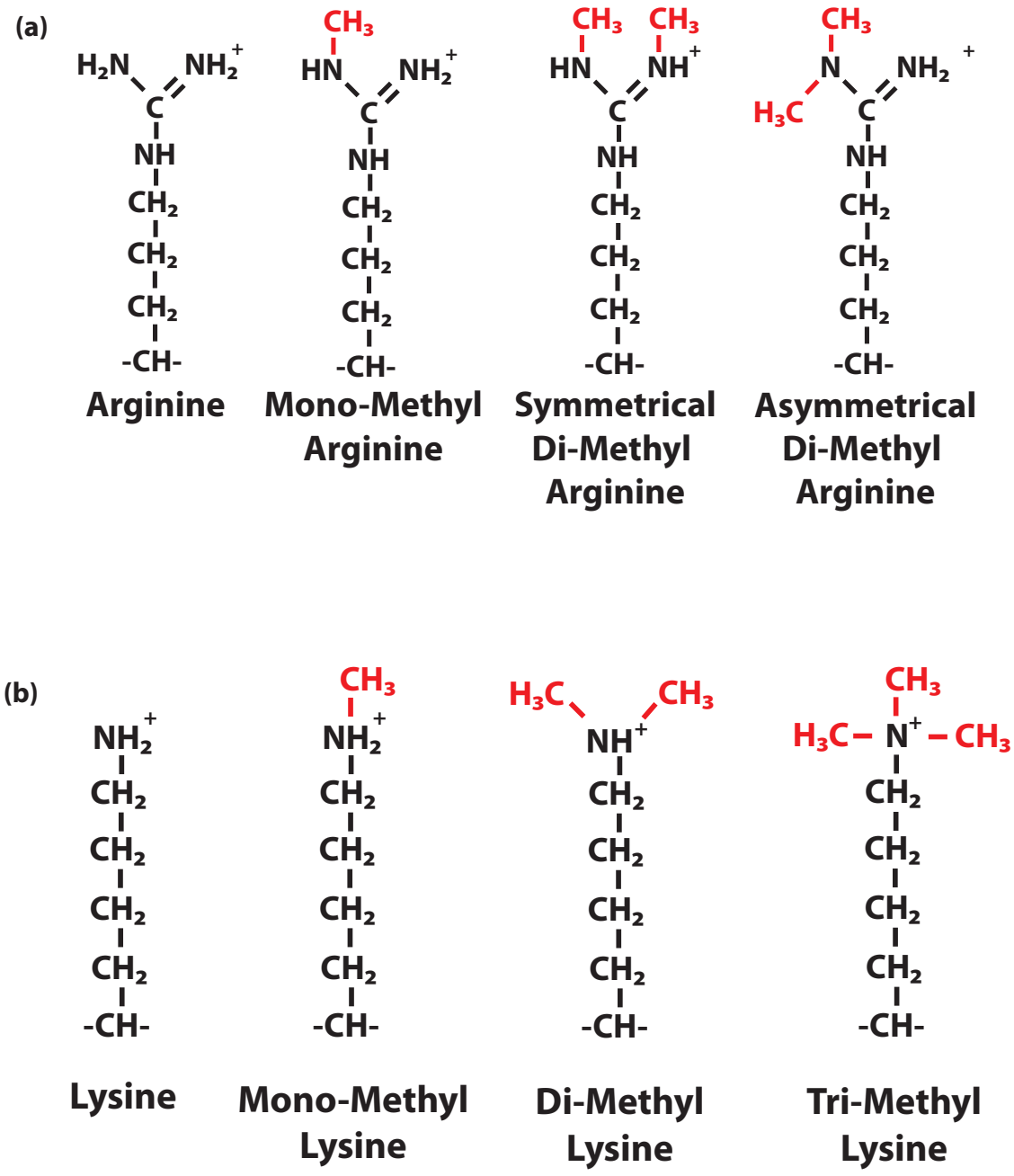


Figure 1.7 Different Methylation States of Arginine and Lysine Residues

(a) Arginine residues can be methylated to form mono-methyl arginine, symmetrical di-methyl arginine and asymmetric di-methyl arginine. (b) Lysine residues can be methylated to form mono-methyl lysine, di-methyl lysine and tri-methyl lysine.

MTases. To date, Dot1p is the only HKMT to be discovered which does not contain a SET domain. Regardless of which family of methyltransferases the HKMTs belong to, they all utilize AdoMet as a methyl group donor. HKMTs display a high degree of substrate specificity and are structurally gauged to selectively modify the substrate lysine to a discrete methylation state [for review see 42].

1.4 Demethylation of Histone Residues

The general belief was that methylation was an irreversible modification, the most stable of all modifications found in histones. This view was established by several groups which showed that global methyl group turnover had similar kinetics to histone turnover [43]. This view started to change with the discovery of an enzymatic activity that selectively removes methyl groups from arginines [43]. To date, no arginine demethylase has been described that converts methylated arginine residues back to arginine. However, the removal of methyl groups from arginines in histones has been shown to occur through a demethyliminate reaction, where methyl arginine is converted to citrulline (Figure 1.8) and thus is not a bona fide demethylation reaction. This reaction is catalyzed by PAD4, a peptidylarginine deiminase in a calcium dependent reaction [43].

A truly dynamic view of methylation came with the discovery of the first lysine demethylase, LSD1, also known as BHC110 and AOF2 [44]. LSD1

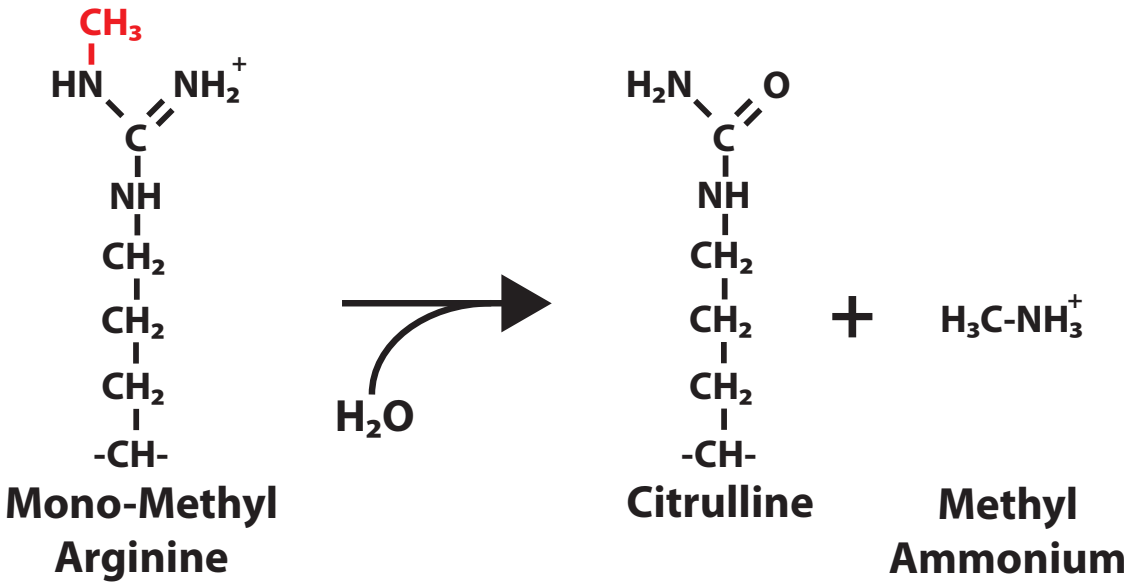


Figure 1.8 Reaction Mechanism of the PAD4 Deiminase

PAD4, the arginine deiminase also catalyzes the methyl deimination of mono-methylated arginine residues. A representation of the demethyl imination reaction is shown.

resides in a number of transcriptional co-repressor complexes [45-49], and belongs to the amine oxidase family of flavoenzymes [50]. Database searches for mammalian orthologues have revealed ten possible candidates, six of which contain a putative nuclear localization signal (NLS) [51]. In humans, LSD1 appears to have only one homologue, LSD2, also known as AOF1 [44].

Shortly after the discovery of LSD1, a second class of demethylase enzymes, JmjC domain-containing demethylase-1 (JHDM1), was discovered [52]. JHDM1 belongs to a class of enzymes that contain a Jumonji C (JmjC) domain. JmjC domains show a similarity to the cupin superfamily, which predominantly consists of metalloenzymes and comprises a subfamily of the oxygenase family [53, 54]. An analysis of the JmjC domain-containing proteins has revealed that a set of 98 proteins from *Homo Sapiens*, *Mus Musculus*, *Drosophila melanogaster*, *Caenorhaditis elegans*, *S. pombe* , and *S. cerevisiae* exists [55]. A phylogenic analysis of the 98 JmjC domain-containing proteins was used to define seven groups to which each of these proteins belong. These families were defined according to overall protein architecture, as well as the presence of additional domains [for review see 55]. About 30 JmjC domain-containing proteins are encoded in the human genome.

1.5 JmjC Domain Histone Lysine Demethylases

To date, several JmjC domain-containing proteins have been shown to be capable of demethylating histone substrates. These include JHDM1A, which affects mono- and di-methyl H3K36 [52], JHDM2A, which affects mono- and di-methyl H3K9 [56], JMJD2A and JMJD2D, which affect tri-methyl H3K9 and tri-methyl H3K36 [57], JMJD2B, which affects tri-methyl H3K9 [57, 58], and JMJD2C, which affects tri-methyl H3K9 and tri-methyl H3K36 [57]. Furthermore, an *in vitro* analysis has demonstrated that recombinant JHDM1A retains the ability to demethylate nucleosomal substrates without the addition of other proteins or factors [52]. Most recently, the JmjC domain-containing JARID protein family, which includes RBP2, PLU1, SMCX, and Yjr119Cp has been shown to possess catalytic demethylation activity towards di- and tri-methylated H3K4 [59, 60].

The structure of the catalytic core of JMJ2A has recently been determined and consists of several domains (Figure 1.9) [53]. The amino terminal portion contains a JmjN domain followed by a JmjC domain and a C-terminal α-helical domain [53]. A zinc finger motif is formed by residues from both the JmjC domain and the C-terminal α-helical domain [53].

JmjC domain proteins catalyze oxidative reactions utilizing Fe(II) and α-ketoglutarate as cofactors. The chemical mechanism proposed for the enzymatic reactions carried out by this family of oxygenases includes the

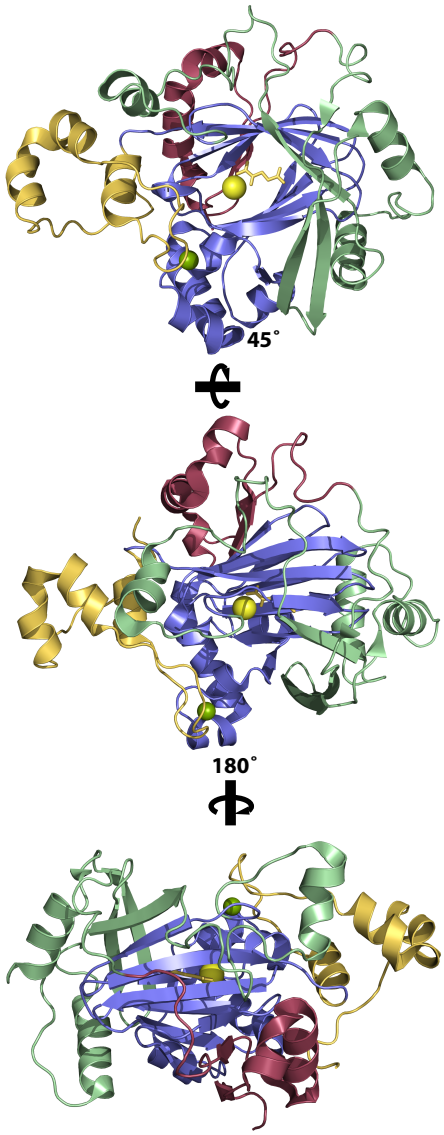


Figure 1.9 Structure of the Human JMJD2A Catalytic Core

The ribbon representation of the crystal structure of the human JMJD2A catalytic core is shown in several orientations (PDB accession code 2GP5). Various colors illustrate the various domains within the catalytic core. The N-terminal JmjN domain is colored dark red. The mixed domain following the JmjN domain is colored green. The JmjC domain is colored blue. The C-terminal domain is colored yellow. The a-ketoglutarate co-factor is represented as a stick figure and is colored yellow. The coordinated zinc and iron atoms are represented as spheres and are colored green and yellow, respectively.

coupling of an oxidative breakdown of α-ketoglutarate to the hydroxylation of the substrate [54]. The hydroxyl group transferred to the substrate is an activated oxygen species, such as an Fe(IV)-oxo intermediate [54]. By-products of the enzymatic reaction are carbon dioxide, succinate from the breakdown of α-ketoglutarate, and formaldehyde from the oxidation of the demethylated lysine residue (Figure 1.10) [52].

1.6 Amine Oxidase Demethylases

1.6.1 Biochemical Activity of LSD1

LSD1 was first shown to contain enzymatic demethylation activity towards lysine 4 of histone 3 (H3K4) [44]. Interestingly, recombinant purified LSD1 can demethylate a mono- or di-methyl peptide substrate and bulk histones; however, it cannot demethylate a nucleosomal substrate by itself and requires an additional protein cofactor named CoREST (Figure 1.11a) [61, 62]. FAD is a common electron acceptor found in many enzymes that catalyze various biological reactions. LSD1 demethylates lysine residues in a flavin dependent oxidation [44]. Two flavin-dependent reaction mechanisms are possible [63]. However, the inability of LSD1 to demethylate tri-methyl lysine as well as structural studies indicate that one of the two mechanisms is more plausible. LSD1's ability to catalyze the oxidation reaction depends on the availability of a free electron pair at the

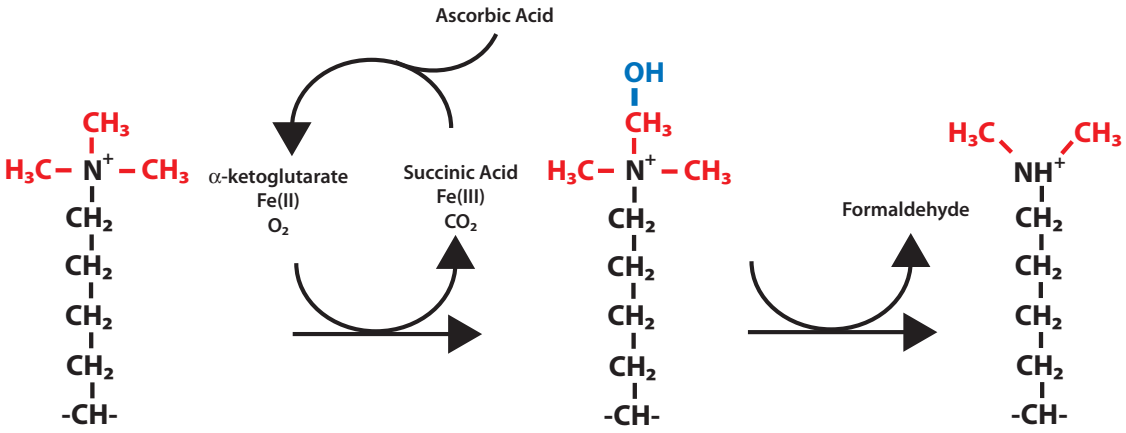


Figure 1.10 Proposed Reaction Mechanism of JmjC Domain Containing Demethylases

The proposed reaction mechanism for oxidative reactions of Jumonji domain proteins is depicted. The oxidative breakdown of the cofactor a-ketoglutarate to succinic acid is coupled to the hydroxylation of a methyl group on the substrate lysine. Following hydroxylation of the methyl group, a spontaneous breakdown results in the formation of the demethylated lysine residue and formaldehyde. Reoxidation and regeneration of a-ketoglutarate from succinic acid is accomplished by ascorbic acid.

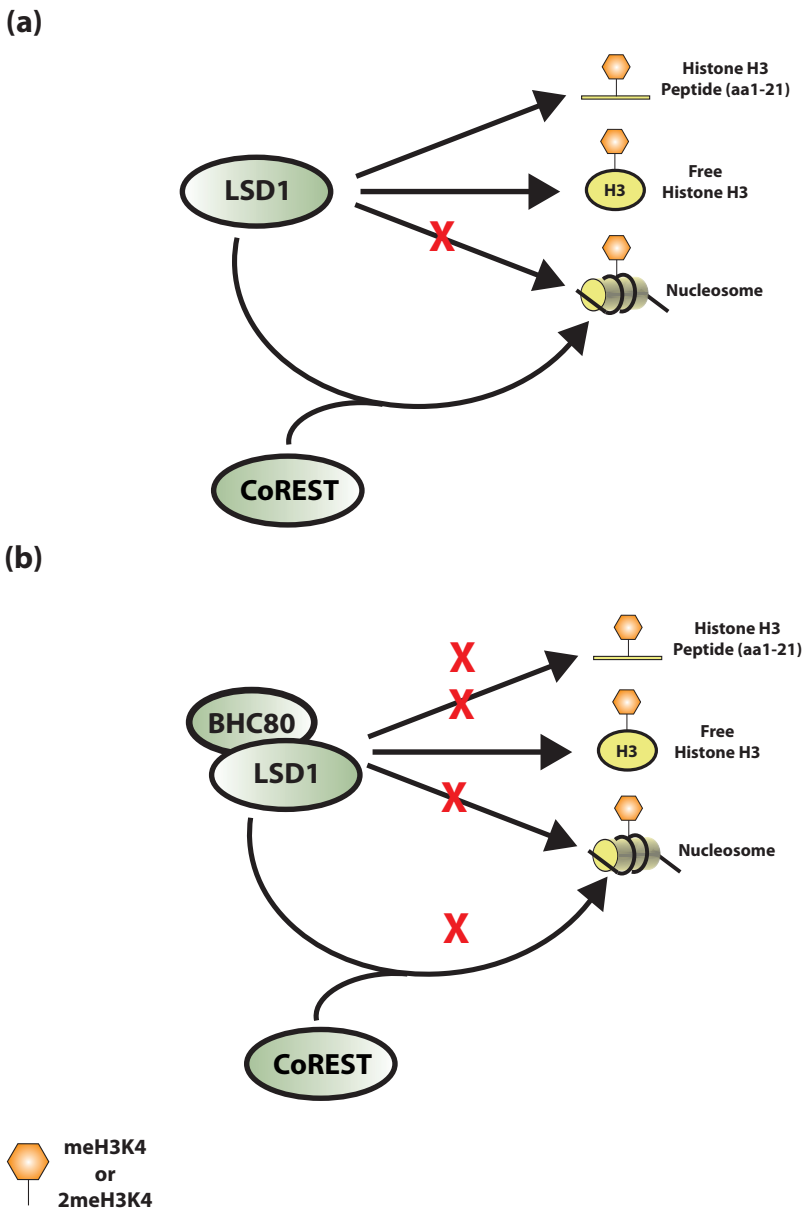


Figure 1.11 Model Of LSD1 Activity on Various Substrates

(a) Purified recombinant LSD1 is capable of demethylating a peptide substrate consisting of the first 21 amino acid residues of histone H3, purified recombinant histone H3 and bulk histone H3 which are mono- or di- methylated on lysine residue 4 (meH3K4 and 2meH3K4). LSD1, however, cannot confer activity against a nucleosomal substrate without the addition of CoREST. (b) LSD1 activity is inhibited against all substrates in the presence of BHC80.

methylated lysine residue. The FAD moiety accepts the electron pair from the nitrogen residue, which allows for the formation of an aminium cation intermediate. Then, the aminium cation intermediate is hydrolyzed to a carbinolamine intermediate, which spontaneously breaks down to formaldehyde and lysine (Figure 1.12) [63]. The flavin cofactor is readily oxidized with molecular oxygen to regenerate FAD, forming hydrogen peroxide as a by-product [63].

1.6.2 *In vitro* Analysis of Human LSD1

Demethylation of a nucleosomal substrate requires additional components of the REST co-repressor complex (CoREST complex), namely CoREST (Figure 1.11a) [61, 62]. Briefly, targeted to the REST-responsive repressor element 1 (RE1), the CoREST complex is required for the repression of neuronal genes in non-neuronal tissues [47]. The complex consists of a number of proteins, including LSD1, CoREST, and Histone deacetylases 1 and 2 (HDAC1,2) (Figure 1.13). CoREST contains several distinct domains. The N-terminal domain consists of an EGL-27 and MTA1 (ELM2) domain followed by two SANT domains that are separated by a linker region. The ELM2 domain is believed to interact with histone deacetylases with stimulatory effects on their deacetylase activity [64, 65]. SANT domains are known to behave as either DNA binding or protein interaction motifs [32, 66]. A domain deletion analysis of CoREST has demonstrated

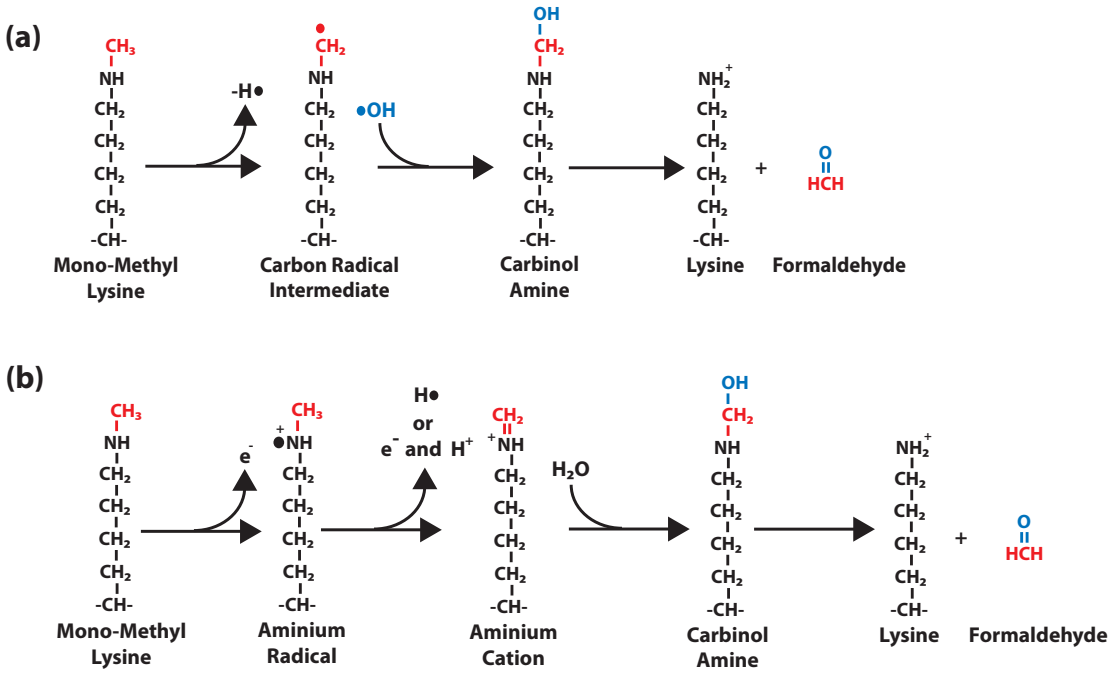


Figure 1.12 Potential Reaction Mechanisms of Amine Oxidase Demethylases

Two possible reaction mechanisms for LSD1 dependent demethylation are depicted. (a) The first pathway requires the removal of a hydrogen radical from the target methyl group, which results in the formation of a carbon radical intermediate. A hydroxyl radical would then need to be transferred to the carbon radical of the methyl group during a hydroxylation reaction. The resulting carbinol amine could then spontaneously breakdown to the demethylated product and formaldehyde. (b) The second pathway requires the removal of an electron from the e-nitrogen group of the lysine residue undergoing demethylation, resulting in the formation of an aminium radical. A hydride ion or an electron plus a proton would then be removed from the target methyl group which would result in a aminiumcation intermediate. Via hydrolysis, a carbinol amine is formed which spontaneously breaks down to the demethylated lysine residue and formaldehyde.

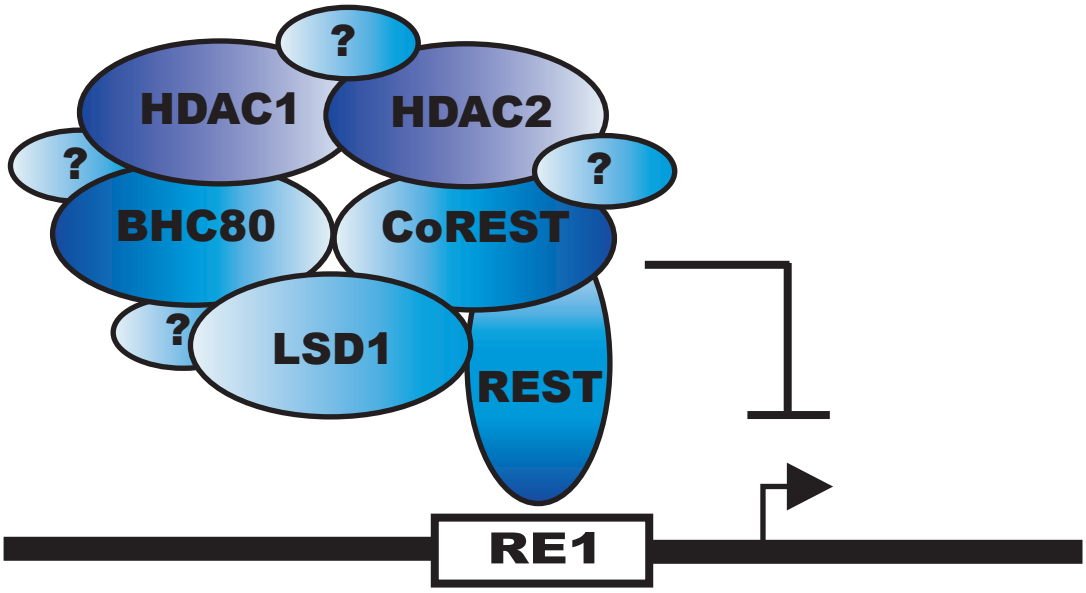


Figure 1.13 Model of REST Transcriptional Corepressor Complex

REST, a zinc finger protein, binds to the REST Response Elements (RE). REST recruits the CoREST corepressor complex to the RE elements where it suppresses expression of neuronal specific genes in non-neuronal cells. The core CoREST complex is comprised of CoREST, LSD1, BHC80, HDAC1 and HDAC2.

that the linker region between the two SANT domains is required in order to stimulate LSD1 nucleosomal demethylation [61, 62] and, furthermore, has been shown to directly interact with LSD1 [67]. However, conflicting results as to the requirement of the SANT domains have been reported. One analysis has shown that SANT2 is required for nucleosomal demethylation by LSD1 [61], whereas another analysis has shown that either of the two SANT domains is sufficient for stimulating activity [62]. The function of SANT1 is not known; however, SANT2 is believed to contain DNA binding activity [67].

BHC80 is a PHD finger-containing protein that co-purifies with the LSD1/CoREST repressor complex. Initial biochemical studies showed that BHC80 inhibits the demethylase activity of LSD1 (Figure 1.11b), which structural studies showed was a result of an affinity of BHC80's PHD finger for unmodified H3K4 [61, 62, 68]. Methylation of H3K4 results in the disruption of the interaction between H3K4 and the PHD finger of BHC80 [68]. BHC80's affinity for demethylated H3K4 suggests that it functions downstream of LSD1 mediated demethylation of H3K4. Many repressor complexes contain both demethylases and methyltransferases and it has been suggested that BHC80 may function to tether LSD1 to the promoters it has demethylated in order to help prevent re-methylation of H3K4 and/or may also promote demethylation of neighboring nucleosomes by LSD1 [68].

1.7 Hormone Dependent Activation of LSD1

1.7.1 Androgen Receptor Activation of LSD1

Surprisingly, LSD1 has been shown to switch its target specificity towards mono- and di-methylated H3K9 in the presence of the Androgen Receptor (AR) [69]. LSD1 was shown to reside on Androgen Response Elements (AREs) in the absence of the AR [60]. A hormone-dependent co-localization of the AR and LSD1 on the AREs stimulates LSD1-mediated H3K9 mono- and di-methyl activity (Figure 1.14) where as the H3K4 methyl levels remained unaltered [69]. The knockdown of LSD1 results in the decreased activation of androgen receptor responsive promoters [56, 69]. TAP-tagged and purified LSD1-AR showed *in vitro* demethylation of di-methyl H3K9 on native nucleosomal substrate [69].

It is quite surprising that LSD1 is capable of changing its substrate specificity in a promoter-dependent manner. How the demethylation specificity is altered from methyl H3K4 to methyl H3K9 remains to be seen. The change in specificity could result from a conformational change induced by a protein-protein interaction or, alternatively, a post-translational modification. Indeed, it will be interesting to see how the specificity is altered, especially since this activity is regulating two opposing processes: gene repression by H3K4 demethylation and gene activation by H3K9 demethylation [35].

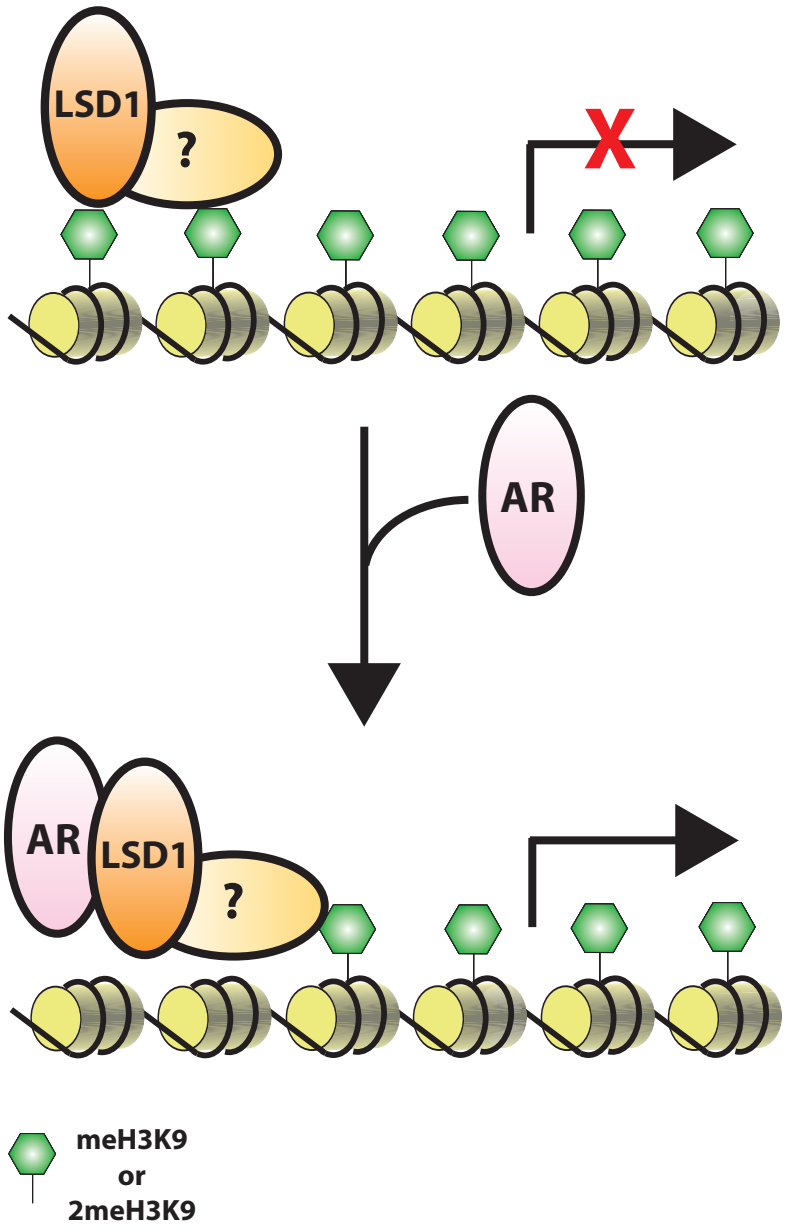


Figure 1.14 Androgen Receptor Stimulated Activity of LSD1

LSD1 is recruited to the Androgen Response Elements (ARE) by an as of yet unidentified mechanism, labeled with a question mark in the model. Ligand induced binding of the Androgen Receptor (AR) to the AREs results in activation of LSD1 demethylase activity towards histone H3 mono- and di-methylated lysine residue 9 (meH3K9 and 2meH3K9) which results in transcriptional activation.

1.7.2 Estrogen Dependent Activation of LSD1

Upon activation by an agonist such as 17β-estradiol (E2), the Estrogen Receptor-alpha (ERα) binds to the bcl-2 promoter and to the Estrogen Responsive DNA Elements (ERE) located 1.5kb downstream from the promoter in MCF7 cells, an estrogen receptor (ER) positive control breast cancer cell line [70]. Bridging between the promoter and enhancer was found to be mediated by ERα [71]. This bridging function has led to the idea that ERα causes looping of the DNA upon activation and is required for transcriptional regulation. The assembly of an active transcriptional unit on the promoter of ERα-regulated genes is also marked with an increase in the levels of di-methyl H3K4 on the promoter and enhancer regions of the loci studied [71]. Upon loss of ERα from chromatin and thus transcriptional down regulation of receptor responsive sites, an increase in the methylation state of H3K9 is seen with a concomitant decrease in the methylation state of H3K4 [71]. Demethylation of mono- and di-methyl H3K9 was found to be dependent on the activity of LSD1 (Figure 1.15) [71]. Treatment of cells with parglyine, a monoamine oxidase inhibitor that has been shown to inhibit LSD1, or knockdown of LSD1, resulted in unaltered mono- and di-methyl H3K9 levels upon activation of ERα [71]. Demethylation of mono- and di-methyl H3K4 at ERα regulated sites is also believed to be mediated by LSD1, which probably resides in a repressor complex that has yet to be described [71].

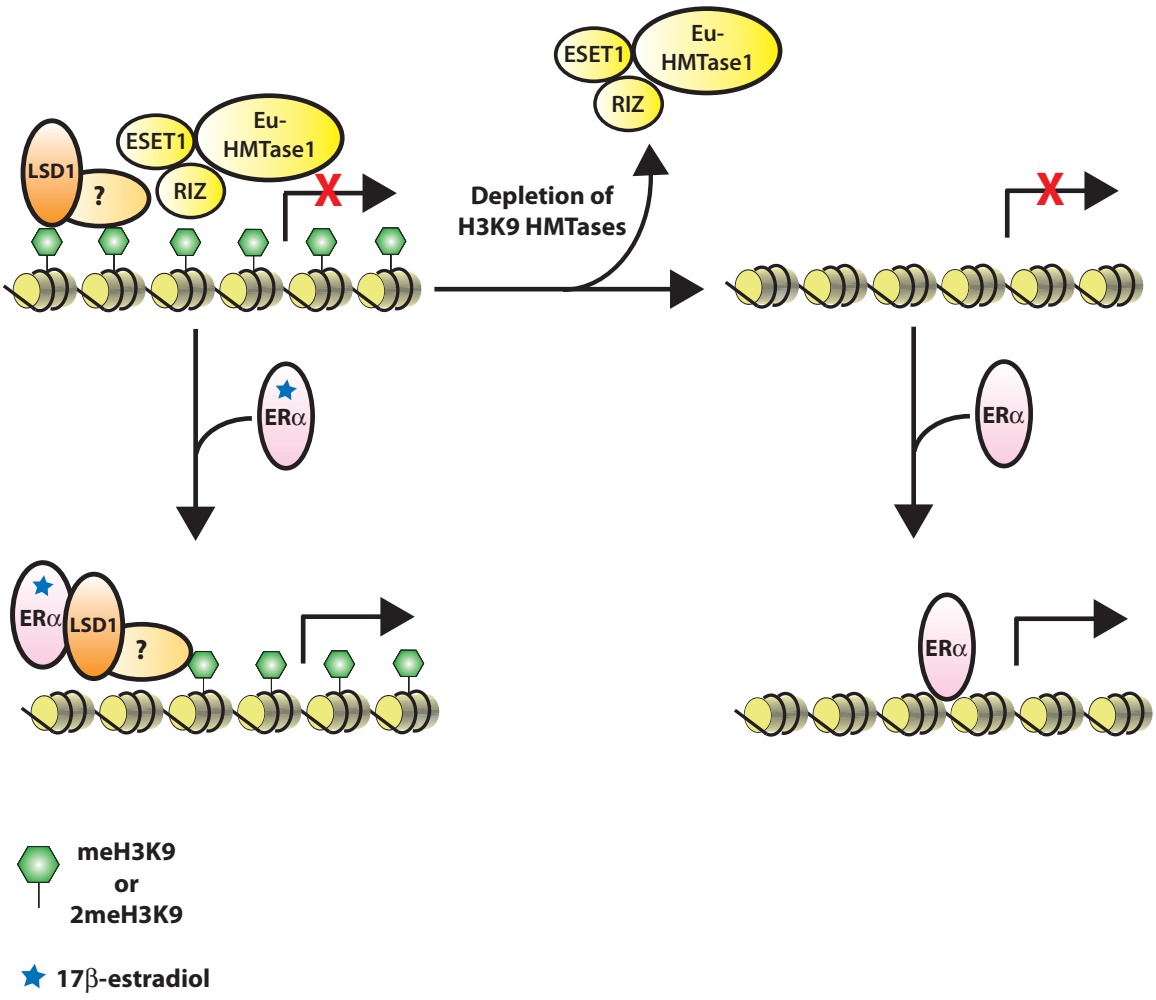


Figure 1.15 Estrogen Receptor Alpha Activation of Transcription

LSD1 is recruited to the Estrogen Response Elements (EREs). Activation of transcription from ERE containing loci normally requires the activation of the Estrogen Receptor-a (ERa) which is achieved by binding a ligand such as 17ß-estradiol, which is shown as a blue star. In the absence of several Histone Methyltransferases which have activity towards lysine 9 of histone H3 (H3K9), unliganded ERa can stimulate transcription from LSD1/ERa dependent transcriptional sites in a LSD1 independent manner.

ERα-dependent looping of chromatin is prevented when LSD1 inhibition occurs, even if the receptor is bound to the EREs [71]. A model has been proposed where one of the byproducts of LSD1-mediated demethylation, hydrogen peroxide (H_2O_2), is used as a signaling molecule in the estrogen dependent transcriptional activation pathway. H_2O_2 reacts with DNA to form adducts. The main product of peroxide-DNA adduct formation is 8-oxo-Guanine (8-oxo-G), and nuclear accumulation of 8-oxo-G at foci is seen upon treatment of cells with estrogen [72]. Estrogen induced formation of 8-oxo-G foci is dependent upon LSD1 activity [71]. 8-oxo-G is normally removed from chromatin by the base excision repair (BER) pathway, specifically by 8-oxo-Guanine DNA glycosylase I (OGGI), which was found to localize to the ERE regions of several loci upon ERα recruitment to those regions of the chromatin [73]. Topoisomerase IIβ is also recruited to the promoters of several ERα regulated genes upon E2 addition and is dependent on LSD1 activity. Topoisomerase IIβ recognizes nicked DNA [74]. The proposed model states that demethylation of H3K9 by LSD1 creates H_2O_2 which reacts with DNA to form 8-oxo-G. 8-oxo-G is removed using OGGI from the BER pathway. BER activities result in the formation of nicked DNA which is recognized by Topoisomerase IIβ which relaxes the chromatin, allowing for chromatin bending and transcriptional activation to occur [71].

A genome wide chromatin immunoprecipitation (ChIP) assay using LSD1 was performed on MCF7 cells after treatment with ERα agonist E_2 [75]. LSD1 was found to reside on approximately 20% of the promoters represented on the array used [75]. Approximately 84% of the promoters that were LSD1 positive were also Pol II positive, histone di-methyl H3K4 positive and histone acetylated H3K9 positive which indicates that LSD1 is present at the promoters of actively transcribed genes [75]. This was corroborated by the mRNA levels of the genes whose promoters where enriched for LSD1 occupancy. When comparing the promoter occupancy of LSD1 to the ligand induced promoter occupancy of the ERα, 58% of the ER promoters also showed LSD1 promoter recruitment [75]. Coimmunoprecipitation experiments also revealed a ligand dependent physical interaction between LSD1, ERα and CBP, an ERα coactivator [75].

To help elucidate the function of LSD1 in the regulation of ERα-dependent transcriptional activation, a series of H3K9 methyltransferases were knocked down using siRNA. Depletion of three methyltransferases, RIZ1, ESET1 and Eu-HMTase1, resulted in the derepression of pS2, a ligand activated ERα-dependent promoter, in the absence of ligand [75]. In fact, depletion of the methyltransferases made the pS2 promoter independent of LSD1 activity. RIZ1/ESET1 depleted cells displayed an increase of ERα recruitment to ligand specific promoters in a ligand

independent manner (Figure 1.15) [75]. This observation supports the idea that one function of H3K9 methyltransferases is to inhibit unliganded ERα recruitment to its target promoters. However, depletion of these three methyltransferases had no effect on the activation, nor on the unliganded ERα binding to LSD1-independent ERα target promoters.

Similar experiments to those for the ERα where carried out for the Androgen Receptor (AR) and the androgen responsive genes [75]. As stated above, LSD1 seems to play a similar role in activation of the AR [56, 69]. Furthermore, RIZ1/ESET1 methyltransferases seem to function in an analogous way with the AR-dependent PSA gene. Using transcriptional reporter assays where the lacZ coding region was preceded by a promoter specific for NFκB, AP-1, βRAR or CREB regulated genes, the effects of LSD1 on promoter activation was assessed. LSD1 had a role in activating NFκB, AP-1 and βRAR associated promoters but did not seem to be involved in activating CREB Response Elements (CRE) which are partially regulated by CREB [75]. In conclusion, LSD1 is involved in activation of a subset of promoters and is not a universally employed demethylase. It has been proposed that LSD1 is the opposing force to the RIZ1/ESET1/Eu-HMTase1 methyltransferases and works to communicate the presence of a liganded response or signaled activation for transcriptional activation of a specified pathway [75].

1.8 LSD1 and Mammalian Development

1.8.1 Role of LSD1 in Organogenesis

The anterior pituitary gland contains five specific hormone secreting cell types which arise from a common precursor cell type [76]. A conditional knockout mouse of LSD1 was created so that LSD1 is removed from the genome in the oral ectoderm primordium of the pituitary [77]. Conditional knockouts at this stage of development by use of the Cre-lox system under the control of Pitx1 Cre^{+} mice leads to deletion of desired loci in nearly all pituitary cells [78]. Morphogenesis of the pituitary gland was undisturbed in $LSD1^{-/-}$ embryos. However, growth hormone (*Gh*) and thyroid stimulating hormone β, the products of differentiated hormone secreting cells of Pit1/Pou1f1 lineage, were not detectable in pituitaries of $LSD1^{-/-}$ embryos [77]. Markers for gonadotrope and corticotrope cell types were significantly decreased in $LSD1^{-/-}$ embryos. Expression of cell-lineage-determining transcription factors and markers for terminally differentiated pituitary cells were diminished whereas the early cell lineage determining factors were expressed at normal levels [77]. Therefore, LSD1 seems essential for proper production of pituitary hormones, late cell-lineage determination and terminal differentiation events but not early cell-lineage commitment during pituitary development [77].

Pit1 is a transcription factor required for *Gh* expression [76]. By developmental stage E13.5-E14.5, Pit1 can be detected on *Gh* promoter

regulatory region by ChIP analysis of pituitary tissue. However, activation of *Gh* transcription occurs at E16.5-E17.5, a time coincident with LSD1 recruitment to the promoter regulatory region of the *Gh* locus [77]. LSD1 is also recruited to several other Pit1 targets and has been found to physically interact with Pit1 [77]. These observations suggest that LSD1 is spatially and temporally regulated for targeted gene activation for terminal differentiation of specified pituitary cell types.

Hey1 is a direct, downstream, target of the Notch signaling pathway [79]. At E17.5 time of development, Hey1 is normally not expressed in the pituitary; however, in the LSD1 deleted pituitary, Hey1 mRNA can be detected which, indicates that LSD1 may have a role in CtBP mediated transcriptional repression of the Notch signaling pathway [77]. LSD1 was also shown to be recruited to several other promoters that are repressed during pituitary development.

Through a series of double chip experiments on the *Gh* promoters, it was deduced that LSD1 is part of at least two complexes that get recruited to the promoter. One complex consists of the ZEB1-CoREST-CtBP-LSD1 corepressor complex and the second consists of the MLL1 coactivator complex (Figure 1.16) [77]. Through pairing of LSD1's enzymatic activity with various complexes, specific targeting and substrate presentation occurs, which results in a functional dichotomy.

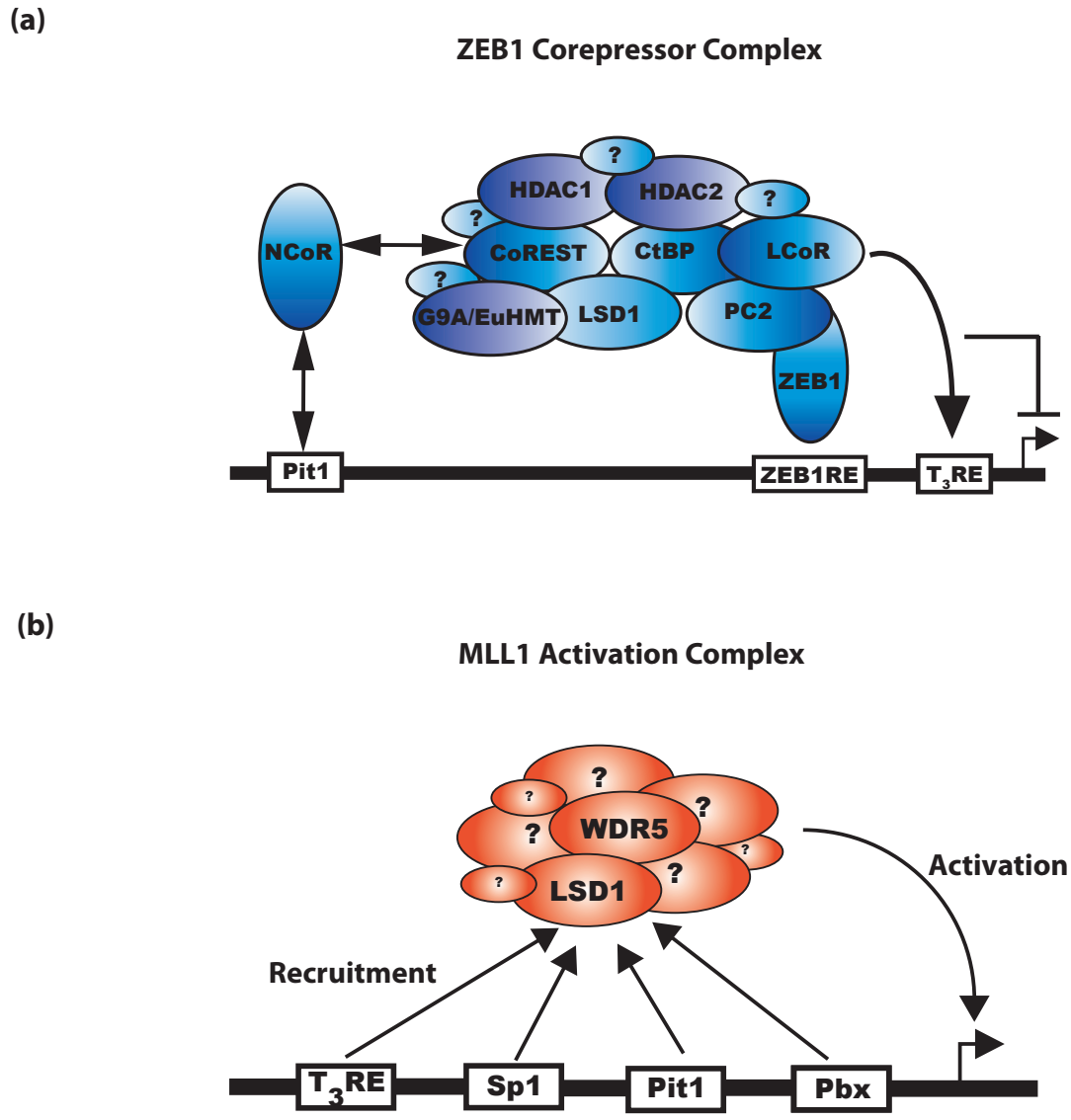


Figure 1.16 Model of LSD1-Containing Complexes Involved in Organogenesis

LSD1 has been found to reside in transcriptional repressor as well as transcriptional activation complexes. (a) The Zeb1 corepressor complex and its interchangeable subunits are shown. The Zeb1 corepressor complex has been shown to negatively regulate transcriptional activation from sites that are dependent on various transcription factors. (b) The MLL1 transcriptional activation complex is composed of LSD1, WDR5 and a number of other components. This complex is known to positively regulate transcriptional activation at sites which are dependent on various transcription factors.

1.8.2 LSD1 and Lymphoid Development

Growth factor independence (Gfi)1 and 1b are homologous proteins required for various developmental processes [80-83]. Both proteins contain a 20 residue N-terminal transcriptional repression SNAG domain and six zinc fingers at the C terminal end [84-86]. Gfi1 has been implicated in lymphoid development, maintenance of hematopoietic stem cells and in generation of neutrophils [87]. Gfi-1b is involved in the development of erythroid and megakaryocytic lineages [87].

Gfi-1b has been found to associate with CoREST, its three homologs Rcors 1-3, HDACs 1 and 2, and LSD1 [88]. Binding to LSD1 and CoREST is mediated via the SNAG domain [88]. CHiP-on-chip analysis for all three proteins revealed that LSD1 and CoREST are recruited to the majority of the Gfi-1 targets [88]. Knockdown of LSD1 or CoREST disrupted hematopoietic differentiation and leads to derepression of Gfi-1b target genes [88]. Analysis of the histone methylation state at the Gfi-1b target sites after LSD1 knockdown revealed an increase in H3K4 methylation status [88]. In conclusion, the SNAG domain recruits LSD1 and the repressor complex to sites of Gfi-1 regulated targets. At these target sites, enzymatic activity of the repressor complex modifies their nucleosomal targets to achieve repression of target genes and to allow proper developmental programs to proceed.

1.9 *S. pombe* LSD1-Like Amine Oxidases

In *S. pombe,* there are two LSD1-like proteins, spLsd1 and spLsd2. spLsd2 knockouts are not viable whereas spLsd1 are, albeit at a slower growth rate is observed as compared to wild type [89]. Biochemical analysis of spLsd1 and spLsd2 revealed a stoichiometric, stable protein complex between spLsd1, spLsd2 and two PHD domain proteins, Phf1 and Phf2, which are of unknown function [89]. Recombinant spLsd1 was shown to have activity towards mono- and di-methylated H3K9, but not mono- or di-methylated H3K4 on bulk histones [89]. Consistent with the observation that spLsd1 demethylates H3K9 *in vitro*, analysis of spLsd1 knockout strains show an increase in global di-methyl H3K9 levels. Deletion of clr4, the H3K9 methyltransferase in *S. pombe*, suppressed the slow growth rate of spLsd1 deletion strains [89].

A genome wide ChIP-chip experiment was conducted to determine whether the global di-methylation status of H3K4 and H3K9 was affected in splsd1 deletion strains. The resultant di-methylation levels of H3K4 and H3K9 were correlated with occupancy of spLsd1 and/or spLsd2 at these particular sites within the genome. A total of 560 euchromatic promoter sites where considered positive for spLsd1 and/or spLsd2 occupancy. spLsd1 and spLsd2 occupied 277 and 190 unique promoters respectively, and 93 promoters where co-occupied [89]. Analysis of the H3K9 methylation state of chromatin in spLsd1 deletion strains revealed an

increase in tri-methyl H3K9 at a significant percentage of the promoters examined but remained unchanged within the coding regions [89]. 60% of spLsd1/2 commonly occupied genomic locations are down regulated in the spLsd1 deletion strain which indicates that spLsd1 may have a direct role in activating transcription [89]. In splsd1 clr4 deletion strains, a portion of the loci regulated by spLsd1 where upregulated, suggesting that spLsd1 functions to prevent H3K9 tri-methylation and/or is involved in the recruitment of a demethylase capable of demethylating a tri-methyl lysine [89].

spLsd1/2 were also found to reside at heterochromatin boundary regions flanking the centromeres, but not within the heterochromatin region [89]. An increase in di- and tri-methyl H3K9 was seen at regions adjacent to the centromeres in the spLsd1 deletion strain [89]. This increase in H3K9 resulted in increased recruitment of Swi6, the *S. pombe* homolog of HP1 [89, 90]. It has been shown that Swi6 is involved in heterochromatin propagation thus the boundaries of the heterochromatin regions are expanded in the spLsd1 deletion strain [89, 90]. Expansion of the heterochromatin boundaries in the spLsd1 deletion strain suggests that spLsd1 may function as a boundary factor by regulating the methylation status of H3K9 within those regions (Figure 1.17a).

The functional effects of spLsd1/2 was also examined for heterochromatin expansion at the *MAT* locus. A reporter strain for the

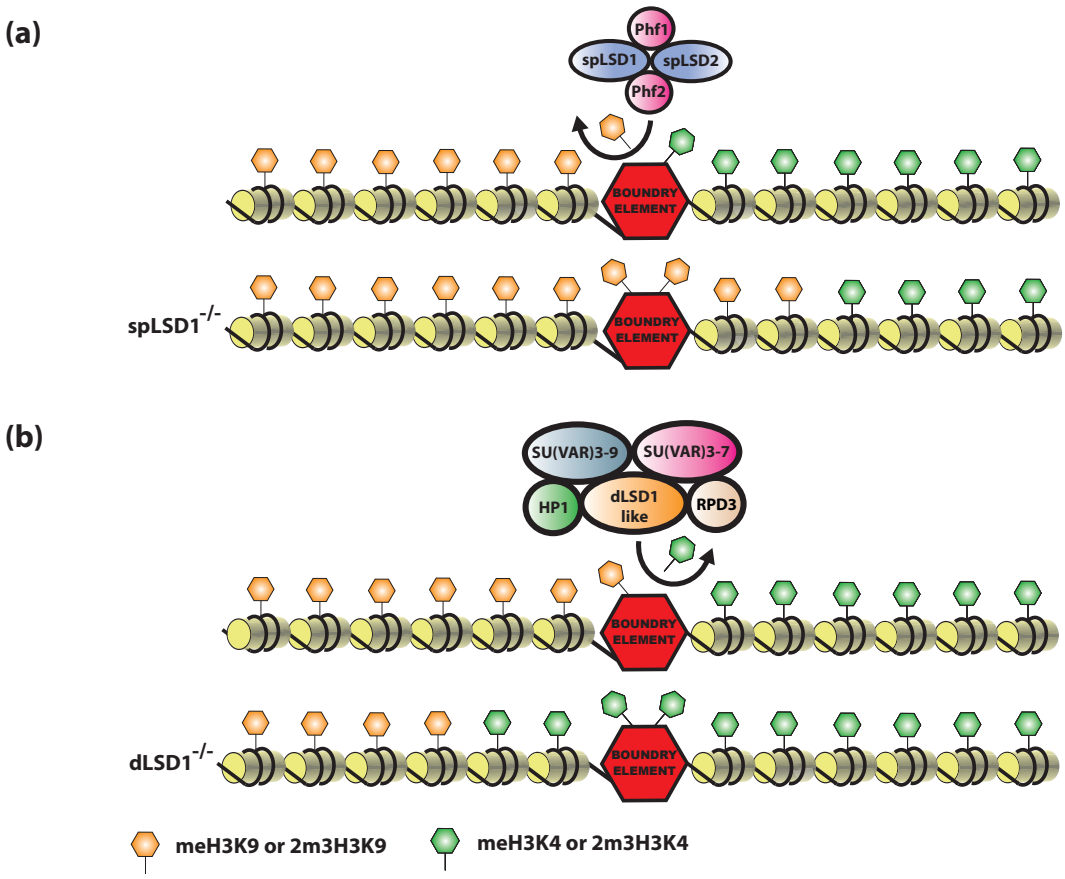


Figure 1.17 Model of Functional Activity of *S. pombe* and *D. melanogaster* LSD1-Like Demethylases

(a) Model of spLsd1/Lsd2, *S. pombe* LSD1-like demethylases, function for the maintenance of heterochromatin boundaries. The two LSD1-like demethylases, spLsd1/2, have demethylase activity towards mono- and di-methyl H3K9. The spLSD1 null mutation in yeast results in the propagation of heterochromatin from the boundary element outwards towards euchromatic regions as a result of increased H3K9 methylation status of those regions due to the lack of spLSD1 activity.

(b) Model of dLSD1, the *Drosphila* LSD1 like demethylase (dLSD1) termed SU(VAR)3-3, function in the maintenance of heterochromatin. dLSD1-like demethylase has activity towards methylated H3K4. dLSD1 null cells have decreased heterochromatin. The heterochromatic region moves away from the euchromatic region, as methylated H3K4 spreads into the heterochromatic regions.

MAT locus was utilized to evaluate the effects of low expression of spLsd1/2 on heterochromatin formation [91]. Decreased spLsd1/2 expression led to decreased expression of the reporter from the *MAT* locus insinuating increased heterochromatin formation [89]. Analysis of the methylation state at the locus showed elevated tri-methyl H3K9 and Swi6 binding [89].

1.10 *D. melanogaster* LSD1-Like Amine Oxidase

The *D. melanogaster* homolog to LSD1, SU(VAR)3-3, plays a functional role in heterochromatin formation. However, unlike the recombinant spLsd1/2 complex, recombinant SU(VAR)3-3 displayed activity towards mono- and di-methyl H3K4 rather than methylated H3K9 [92].

The function of SU(VAR)3-3 was examined during early embryonic development. During the first 13 cycles of development, SU(VAR)3-3 displays a dynamic redistribution throughout the various stages [92]. At cycle 14, when the cellular blastoderm is formed, SU(VAR)3-3 becomes concentrated at pericentric heterochromatin locations [92]. As development proceeds, SU(VAR)3-3 is found only at euchromatic sites [92].

Embryos produced by crossing males and females heterozygous for SU(VAR)3-3 null mutations were examined for heterochromatin formation at cycle 14 of development. SU(VAR)3-3 expression was greatly reduced

in the resulting embryos [92]. The levels of di-methyl H3K4 at heterochromatin locales increased in the null embryos, whereas HP1 and di-methyl H3K9 were reduced [92]. Coincident with these effects was the increase in global mono- and di-methyl H3K4 [92]. A reduction in tri-methyl H3K9 was also observed, without any significant changes in mono- and di-methyl H3K9 [92]. An increase in acetylated H3K9 in the heterochromatin regions was also observed [92].

In conclusion, H3K4 demethylation by SU(VAR)3-3 and deacetylation of H3K9 by RPD3 deacetylase, are coordinated during the early stages of embryogenesis, and a SU(VAR)3-3 null mutations leads to loss of heterochromatic chromatin. Further interpretation of the results leads to speculation that SU(VAR)3-3 may control the activity of SU(VAR)3-9, a methyltransferase with activity against H3K9, either directly or by creating a proper substrate, or indirectly by affecting its targeting.

The interplay of various combinations of modifications on histones has been seen and the modification of histones in a sequential manner has been observed in several instances [93]. The activity of the *D. melanogaster* methyltransferase SU(VAR)3-9 is inhibited when H3K4 is mono- or di-methylated; however, inhibition is relieved upon addition of SU(VAR)3-3, a demethylase [92]. Immunoprecipitation experiments revealed that SU(VAR)3-3 is associated with SU(VAR)3-9, RPD3 and HP1, which has been defined as a silencing complex [92]. One can

envision a plausible sequence of events that leads to heterochromatin formation: SU(VAR)3-3 demethylates H3K4, allowing RPD3 to deacetylate H3K9 in order to create a substrate for SU(VAR)3-9, which leads to H3K9 methylation, a high affinity binding site for HP1. Thus, during development, the H3K4 demethylase activity of SU(VAR)3-3 may help define the boundary elements between euchromatin and heterochromatin by preventing the methylation of H3K4 from entering the heterochromatic regions and by recruiting the deacetylase activity of RPD3 to create the SU(VAR)3-9 H3K9 substrate (Figure 1.17b) [92].

A comparison of heterochromatin formation and regulation between *S. pombe* and *D. melanogaster* reveals a dichotomy at the mechanistic level leading to the same final outcome. In fission yeast, spLsd1/2 work to establish and maintain heterochromatin boundaries by demethylating H3K9 in the boundary/euchromatin region, whereas in *D. melanogaster,* Lsd1 establishes and maintains heterochromatin boundaries by demethylating H3K4 in the boundary/heterochromatin regions.

1.11 LSD1 and Non-Histone Substrates

A number of non-histone proteins have been shown to undergo methylation on lysine residues. Among the growing list is the tumor suppressor and transcriptional activator p53 [94]. Di-methylation of p53 on lysine residue 370 (K370) promotes association of p53 with the coactivator

53BP1 through its Tudor domain [95-98]. LSD1 is capable of demethylating mono- and di- methyl K370 *in vitro* [94]. Knockdown of LSD1 in U2OS and H1229 cells results in an increase of p53 di-methylation at residue K370 [94]. Knockdown of LSD1 in U2OS cells followed by treatment of the cells with adriamycin, a DNA intercalating chemotherapeutic, results in an increased population of cells that undergo apoptosis [94]. These results indicate that LSD1 functions to inhibit the p53 mediated transcriptional activation that leads to apoptosis. The proposed mechanism of action for LSD1 inhibition is that LSD1 works by demethylating K370 of p53, which prevents 53BP1 binding and thus co-activation of p53 target genes [94].

1.12 Functional Coup ling Between Demethylase and Deacetylase Activities

In vitro analysis of the LSD1-containing complexes revealed a functional interplay between HDAC1-mediated histone deacetylation and LSD1-mediated histone demethylation [64]. As with LSD1, HDAC1 activity was detected on bulk histones; however, the deacetylation of a nucleosomal substrate required additional components of the co-repressor complex [100]. The addition of CoREST to the reaction conferred on HDAC1 the ability to deacetylate a nucleosomal substrate. Deletion analysis indicated that the ELM2 domain of CoREST is sufficient to

mediate nucleosomal deacetylation by HDAC1 [100]. However, the addition of LSD1 further enhanced the deacetylase activity, whereas a catalytically inactive LSD1 mutant did not [100]. In addition to the *in vitro* results, over-expression of LSD1 in HEK293 cells decreased di-methyl H3K4 and acetyl H3 levels at the synapsin promoter, an RE regulated site, whereas over-expression of the catalytically inactive mutant of LSD1 increased the levels of these modifications relative to the control [100]. Treatment of HEK 293 cells with Trichostatin A (TSA), a histone deacetylase inhibitor, resulted in an increase of acetyl H3 levels at the synapsin promoter as well as in the levels of di-methyl H3K4 [100]. These results suggest a functional coupling between the two enzymatic activities.

1.13 Known Amine Oxidase Inhibitors

Monoamine oxidase A and B (MOA A and MOA B) are flavin-dependent enzymes that are therapeutic targets for the treatment of neurological disorders [101]. MOAs catalyze the oxidative de-amination of neurotransmitters [101]. A number of MOA inhibitors have been synthesized and their specificity for MOA A or B determined [102]. Some of the inhibitors are specific to a single oxidase, whereas other inhibitors are non-selective [102]. MOAs share homology with LSD1 and potentially catalyze their respective oxidation reactions with a mechanism similar to that of LSD1. Due to these similarities, known MOA inhibitors were tested

for their ability to inhibit LSD1 activity. Six MOA inhibitors, three of which are specific for either MOA A or MOA B (Clorgyline, Deprenyl, and Pargyline) and three of which are non-selective (Phenelzine, Tranylcypromine, and Nialamide), were tested for their ability to inhibit LSD1 activity [103]. On bulk histones, all inhibitors at high concentrations had an inhibitory effect on recombinant LSD1 [103]. Although the inhibitory effect of Clorgyline, Deprenyl, Pargyline, and Nialamide was minimal, Phenelzine and Tranylcypromine exhibited a potent inhibitory effect [103]. In the presence of CoREST, Clorgyline, Deprenyl, and Pargyline did not have an inhibitory effect on a nucleosomal substrate, whereas Nialamide retained its weak inhibitory characteristics. However, Phenelzine and Tranylcypromine displayed, once again, a potent inhibitory effect [103]. Although Pargyline did not inhibit the LSD1/CoREST complex [103], it was shown to inhibit LSD1 H3K9 demethylation activity when associated with the androgen receptor complex [69].

The first attempt to synthesize an inhibitory molecule for LSD1 was based on the known mechanisms of action for MOA inhibitors [104]. A 21-residue, propargyl-Lys-derivatized histone 3 tail peptide was synthesized, and a kinetic analysis of this inhibitor peptide indicated that it acts as an irreversible inhibitor [104]. The mass spectroscopic analysis of the LSD1/inhibitor mixture showed a peak corresponding to the molecular weight of a covalent adduct between the FAD and inhibitor [104]. This

suggests that irreversible inhibition results from the covalent attachment of the inhibitor to the FAD molecule.

MATERIAL AND METHODS

2.1 Protein Expression

DNA fragments comprising the entire region or truncated versions of human LSD1 were amplified by PCR and cloned into a modified pET28a vector (Novagen) that contained a PreScission protease site directly after the N-terminal hexahistidine tag. Bacterial protein expression was carried out in triple baffled Fernbach flasks (Bellco) at a shaker speed of 200 rpm. The protein was expressed in *E. Coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene) and induced with 200 μM IPTG at 18° C for 16-18 hours in rich media (10 g bactotryptone, 2 g casamino acids, 2 g yeast extract, 5 g NaCl, 0.4% (w/v) glucose, and 10 μM riboflavin per liter). Typically, 40 liters of culture where grown per protein preparation. Cells were harvested by centrifugation and resuspended in 150 ml of a buffer containing 25 mM HEPES pH 8.0, 250 mM NaCl, 5% (v/v) glycerol, 8 mM β-mercaptoethanol, 8 mM imidazole, 100 μM phenyl-methane-sulphonyl-fluoride and Complete EDTA-free protease inhibitor cocktail tablets (Roche). The cells were lysed with a cell disrupter (Emulsiflex-C5, Avestin), and the lysate was centrifuged for 90 minutes at 40,000g. Flowchart of protein purification scheme can be seen in figure 2.1.

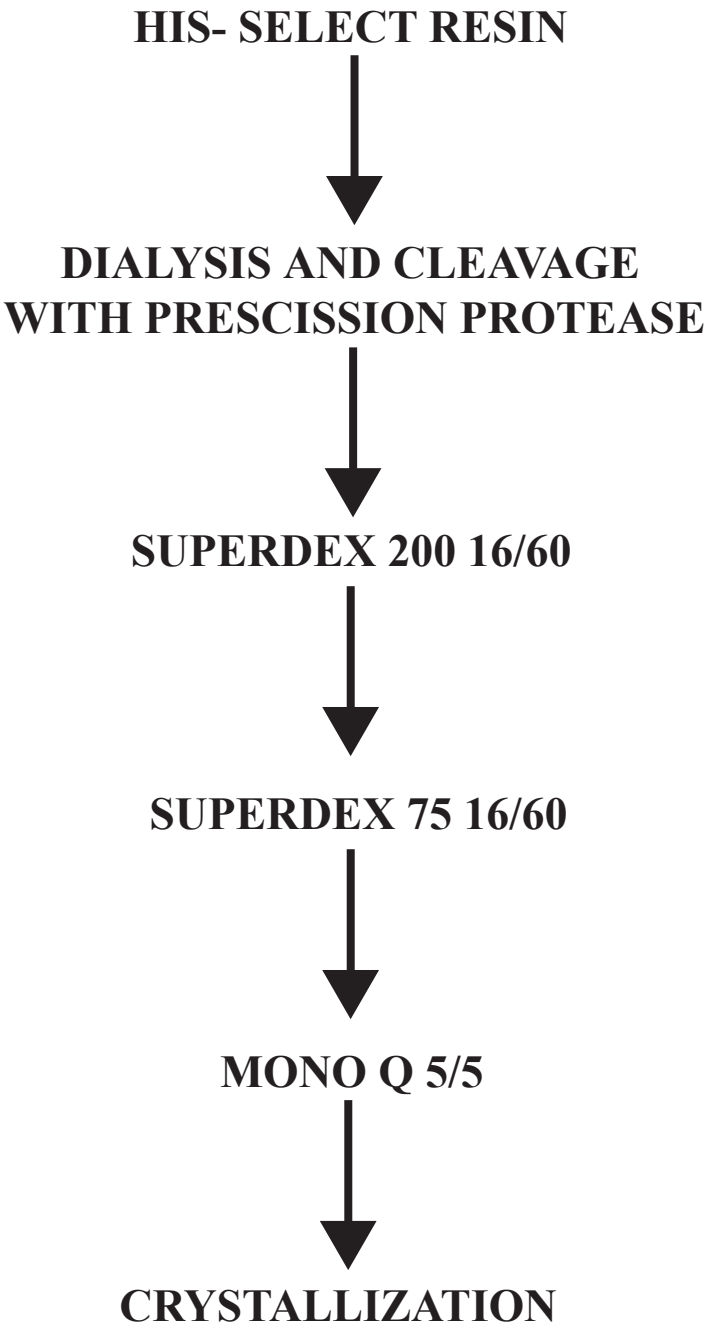


Figure 2.1 Flowchart of Protein Purification of Human LSD1

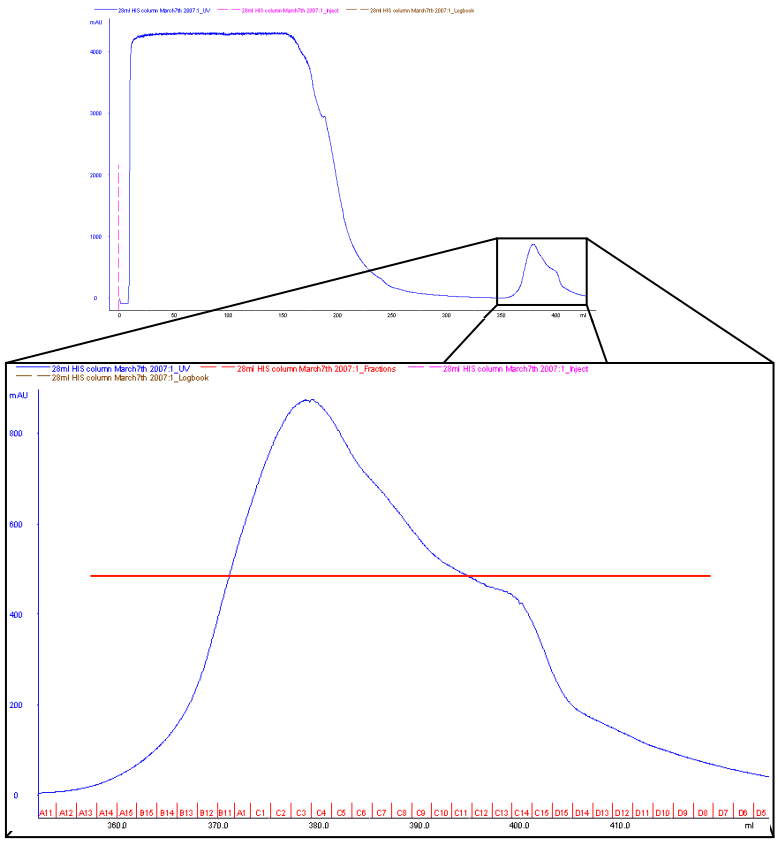
2.2 Protein Purification

The clarified lysate was applied to a 30 ml His-Select nickel column (Sigma) pre-equilibrated in Buffer A, which consists of 20 mM HEPES pH 8.0, 250 mM NaCl, 5% (v/v) glycerol, 8 mM imidizole and 20 mM 2-mercaptoethanol at a rate of 0.5 ml/min. Injection of lysate was followed by a 10 column volume wash with Buffer A. The protein was eluted via a 20 column volume linear gradient using Buffer A as the start buffer and Buffer B, which consists of 20 mM HEPES pH 8.0, 250 mM NaCl, 5% (v/v) glycerol, 500 mM imidizole, and 20 mM 2-mercaptoethanol as the final buffer (Figure 2.2).

Fractions containing LSD1 were pooled, dialyzed against a buffer containing 25 mM HEPES pH 8.0, 250 mM NaCl, 5% (v/v) glycerol, 5 mM dithiothreitol (DTT), and cleaved with PreScission protease (GE Healthcare) for 36 hours. 5 mM of fresh DTT was added every 12 hours during the duration of fusion protein cleavage with PreScission protease. The dialyzed and cleaved LSD1 was concentrated using a Millipore 50,000 Da MWCO 15 ml Ultra15 concentrator at 2,000g to a final volume of 4ml.

The concentrated protein was then purified over a 16/60 Superdex 200 column that was pre-equilibrated in a buffer containing 20 mM HEPES pH 8.0, 250 mM NaCl, 5% (v/v) glycerol, and 5 mM DTT using two consecutive runs at a flow rate of 1.0 ml/min (Figure 2.3).

(a)



(b)

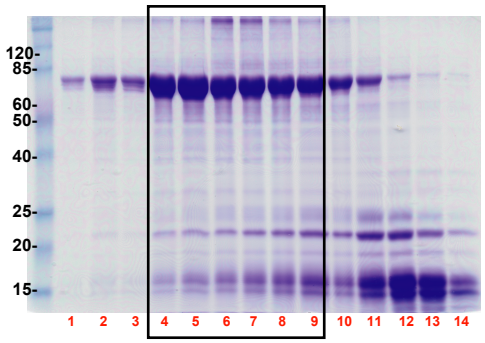


Figure 2.2 Immobilized Nickel Affinity Purification

A 30 ml His-Select column was used for the initial purification step of human LSD1. (a) A chromatographic representation of the His-Select purification UV trace at 280nm. Offset to the rear is the profile of the complete purification. An enlarged view of the elution profile is shown below. (b) A 12% denaturing SDS polyacrylamide gel stained with coomassie brilliant blue R-250 illustrating the peak fractions that are indicated by the red line. The black inset on the chromatograph and gel defines the fractions that were pooled and further purified.

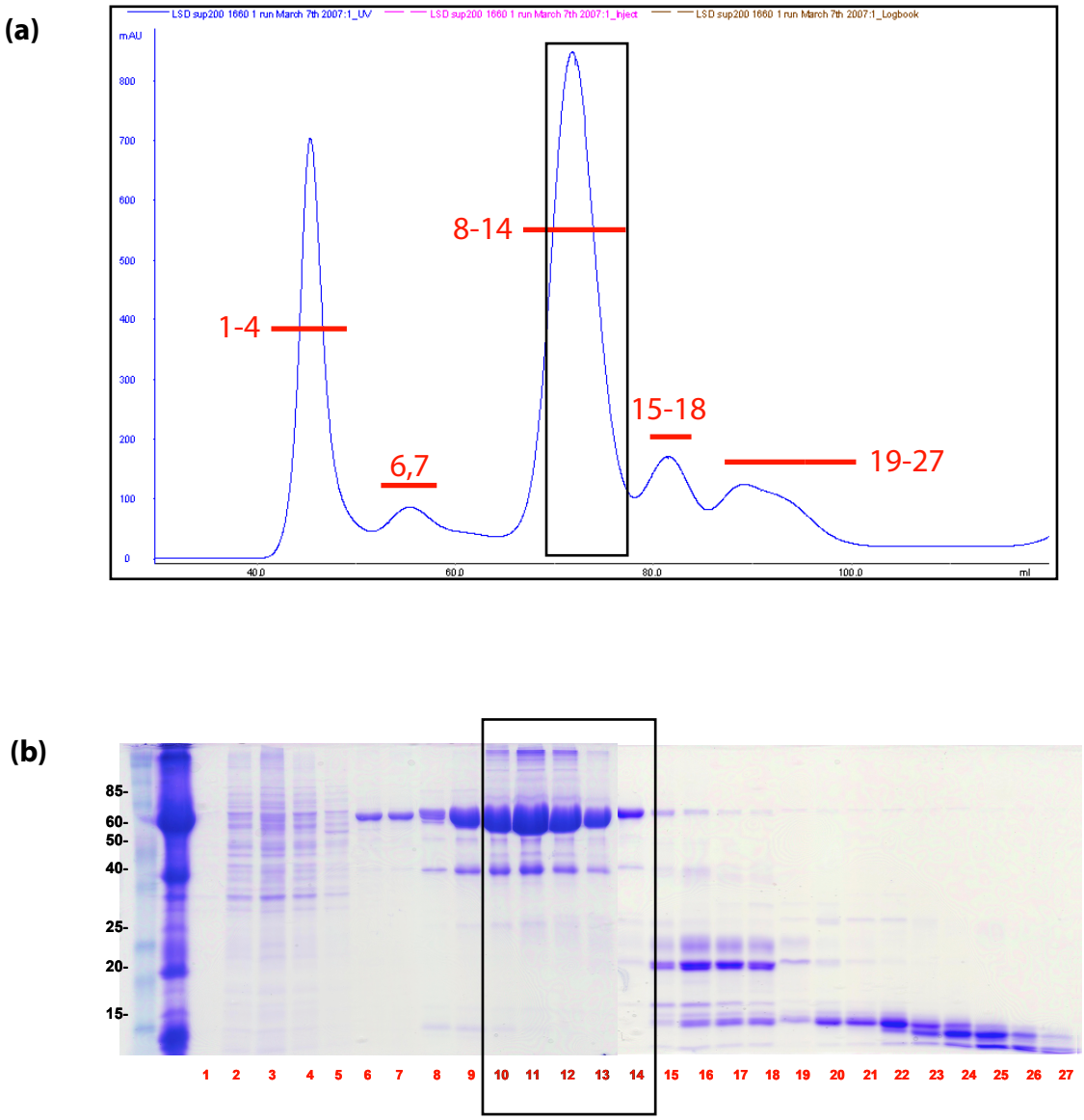


Figure 2.3 Gel Filtration Chromatography on Superdex 200 16/60

LSD1 was further purified using size exclusion. (a) A chromatographic representation of UV trace at 280nm. (b) A 12% denaturing SDS polyacrylamide gel stained with coomassie brilliant blue R-250 illustrating the peak fractions that are indicated by the red line. The black inset on the chromatograph and gel defines the fractions that were pooled and further purified.

Fractions positive for LSD1 where pooled and concentrated using a Millipore 50,000 Da MWCO 5 ml Ultra5 concentrator at 2,000g to a final volume of 2 ml. The concentrated protein was then purified over a 16/60 Superdex 75 column that was pre-equilibrated in a buffer containing 20 mM HEPES pH 8.0, 250 mM NaCl, 5% (v/v) glycerol, and 5 mM DTT at a flow rate of 1.0 ml/min (Figure 2.4).

The positive fractions were pooled and diluted in a buffer containing 20 mM HEPES pH 8.0, 5% (v/v) glycerol,and 5 mM DTT until a NaCl concentration of 50 mM was reached. The diluted protein sample was run on a Mono Q 5/5 column which was pre-equilibrated with 20 mM HEPES pH 8.0, 50 mM NaCl, 5% (v/v) glycerol, and 5 mM DTT at a flow rate of 1.0 ml/min. A linear salt gradient was applied with the final buffer consisting of 20 mM HEPES pH 8.0, 1 M NaCl, 5% (v/v) glycerol, and 5mM DTT (Figure 2.5).

Fractions containing LSD1 were pooled and dialyzed for 4 hours in a buffer containing 20 mM HEPES pH 8.0, 250 mM NaCl, 5% (v/v) glycerol, and 5 mM DTT. The dialyzed sample was then concentrated using a Millipore 50,000 Da MWCO 5 ml Ultra5 concentrator at 2,000g. The protein sample was concentrated to 16 mg/ml for crystallization. Typically, yields were approximately 4 mg of pure protein from 40 liters of *E.coli* culture.

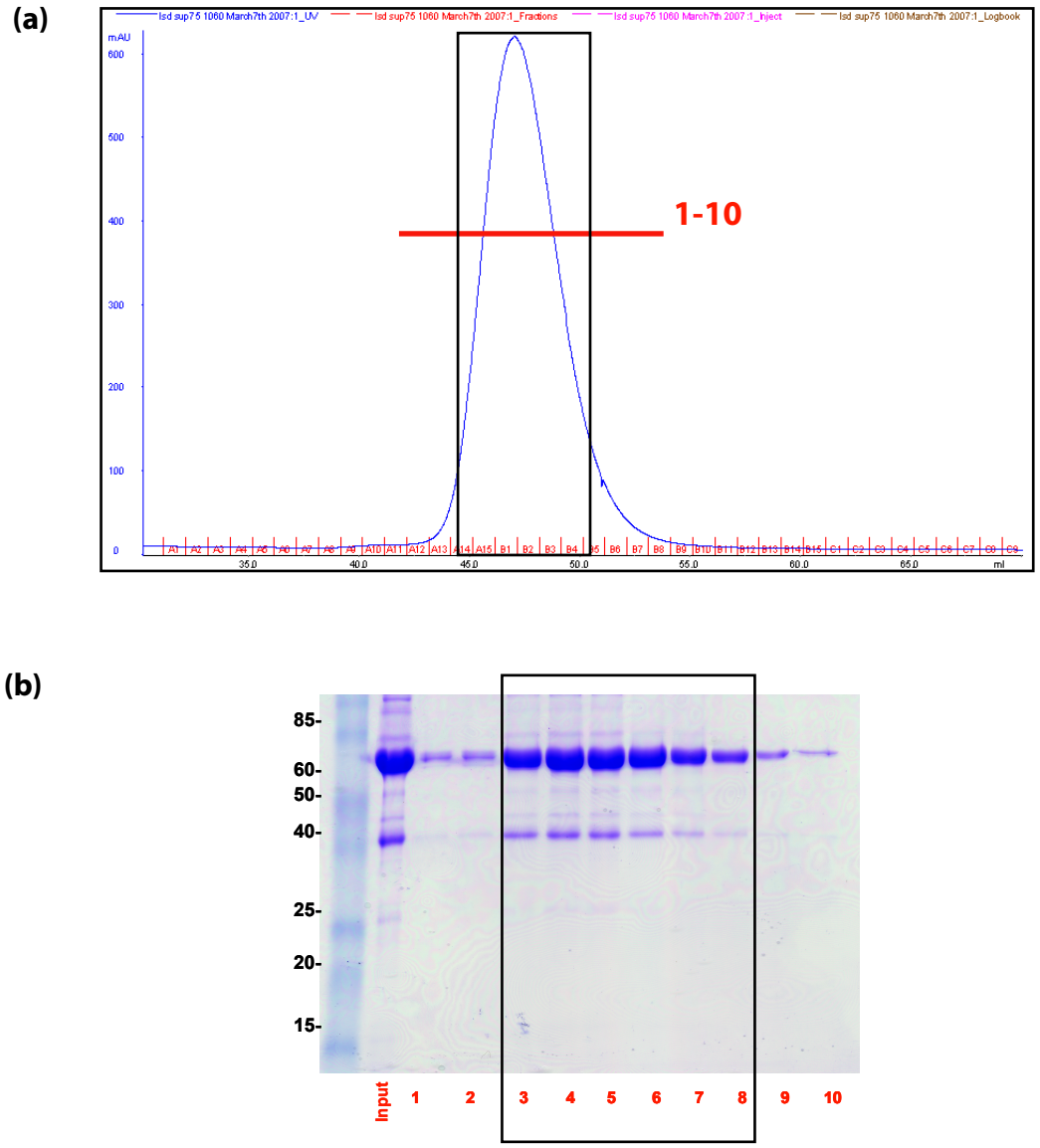
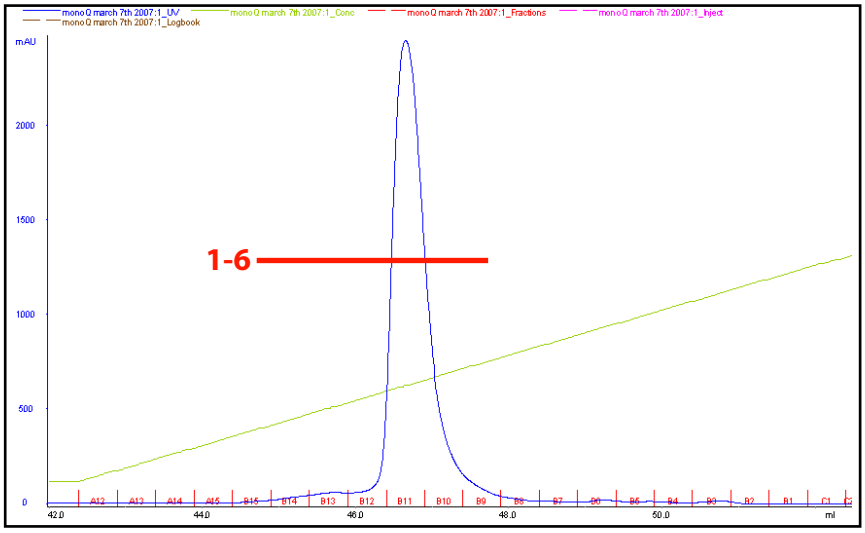


Figure 2.4 Gel Filtration Chromatography on Superdex 75 16/60

LSD1 was further purified using size exclusion. (a) A chromatographic representation of UV trace at 280nm. (b) A 12% denaturing SDS polyacrylamide gel stained with coomassie brilliant blue R-250 illustrating the peak fractions that are indicated by the red line. The black inset on the chromatograph and gel defines the fractions that were pooled and further purified.

(a)



(b)

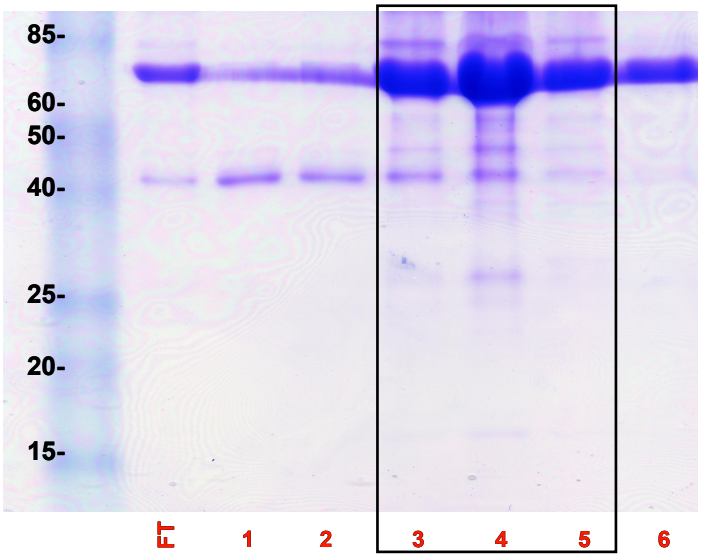


Figure 2.5 Ion Exchange Chromatography on Mono Q 5/5

LSD1 was further purified using ion exchange chromatography. (a) A UV trace at 280nm of the Mono Q 5/5 run. (b) A 12% denaturing SDS polyacrylamide gel stained with coomassie brilliant blue R-250 illustrating the peak fractions that are indicated by the red line. The black inset on the chromatograph and gel defines the fractions that were pooled and used for crystallization.

2.3 Expression of Selenomethionine-Labeled Proteins

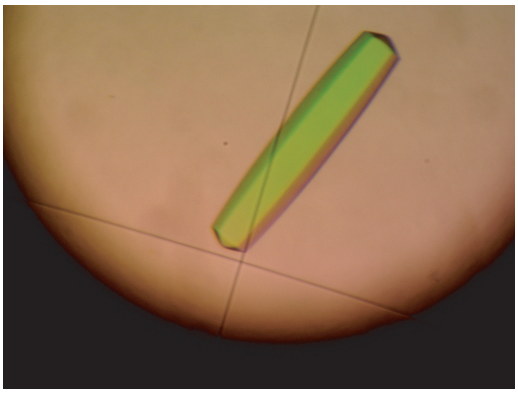
Proteins labeled with selenomethionine were produced using standard protocols in *E. Coli* BL21-CodonPlus(DE3)-RIL cells in a synthetic media that suppresses methionine biosynthesis [105]. Protein expression was induced with 200 μM IPTG at 18° C for 18 hours. The protein purification protocol was identical to the native protein purification protocol decribed above. Typically, the yields were approximately 0.25-0.50 mg of pure protein from 60 liters of *E.coli* culture.

2.4 Crystallization and Structure Determination

Crystals of LSD1 (16 mg/ml) were grown at 21°C in hanging drops containing 1 μl of the protein and 1 μl of a reservoir solution consisting of 9-12% (w/v) PEG 3350, 200 mM di-ammonium tartrate, and 100 mM HEPES pH 6.4-6.7. Crystals grew in the hexagonal space group $P6_122$, with one molecule in the asymmetric unit. They grew to their maximum size of 350 μm x 100 μm x 100 μm within 10 days (Figure 2.6a).

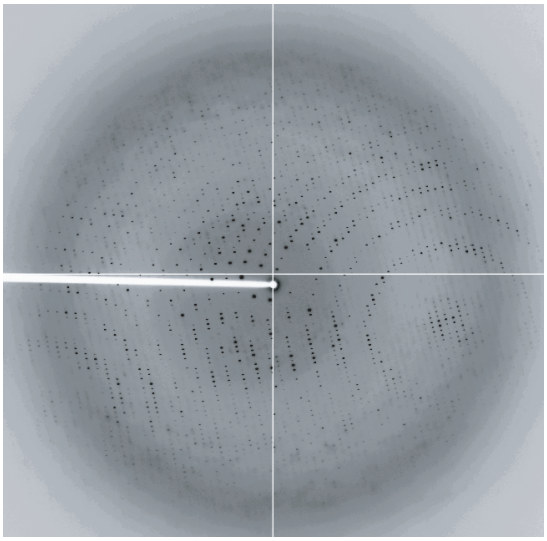
For cryoprotection, crystals were stabilized in 15% (w/v) PEG 3350, 200 mM di-ammonium tartrate, 100 mM HEPES pH 6.6, and 20% (v/v) glycerol and flash frozen in liquid nitrogen-cooled liquid propane. X-ray diffraction data were collected at the Advanced Light Source (ALS), Lawrence

(a)



Hexagonal crystals

(b)



X-ray diffraction pattern

Figure 2.6 Representative Crystal of Human LSD1 and its Diffraction Pattern

(a) A single crystal of human LSD1 shown in its crystallization drop. The crystal has a yellow hue due to the presence of its FAD cofactor. LSD1 crystals belong to the hexagonal space group P6122. (b) Diffraction pattern of the hexagonal crystal shown in (a). The typical diffraction limit of LSD1 crystals is approximately 2.7 Å. The shown diffraction pattern was collected at beamline 8.2.1 at the Advanced Light Source (ALS), Lawrence Berkeley National Laboratory (LBNL). The image corresponds to an oscillation angle of 1.0°, measured over a time period of 15 seconds.

Berkeley National Laboratory (LBNL), beamline 8.2.1 (Figure 2.6b; Table 1 and 2).

X-ray intensities were processed using the HKL2000 denzo/scalepack package (Figure 2.6b) [106]. The CCP4 program package was used for subsequent calculations [107].

Initial phases were determined to approximately 4.0 Å using a $[Ta_6Br_{12}]^{2+}$ cluster derivative and multiwavelength anomalous dispersion (MAD) measurements [108, 109]. These phases were used to locate the heavy atom sites of sodium ethylmercurithiosalicylate (EMTS) and potassium osmate (K_2OsO_4) derivatized and selenomethionine (SeMet) labeled protein crystals (Figure 2.7). Phasing from a SeMet MAD data set was conducted in SHARP [110]. The phases where improved by density modification in DM with solvent flattening and histogram matching. This phasing procedure yielded an electron density map of high quality (Figure 2.8) [107]. A model lacking only two disordered loops was built with the program O [111]. The selenium, osmium and mercury sites aided in the identification of side chain positions. The model was refined using CNS [112]. The stereochemical quality of the model was assessed with PROCHECK [113]. There are no residues in the disallowed region of the Ramachandran plot (Figure 2.9).

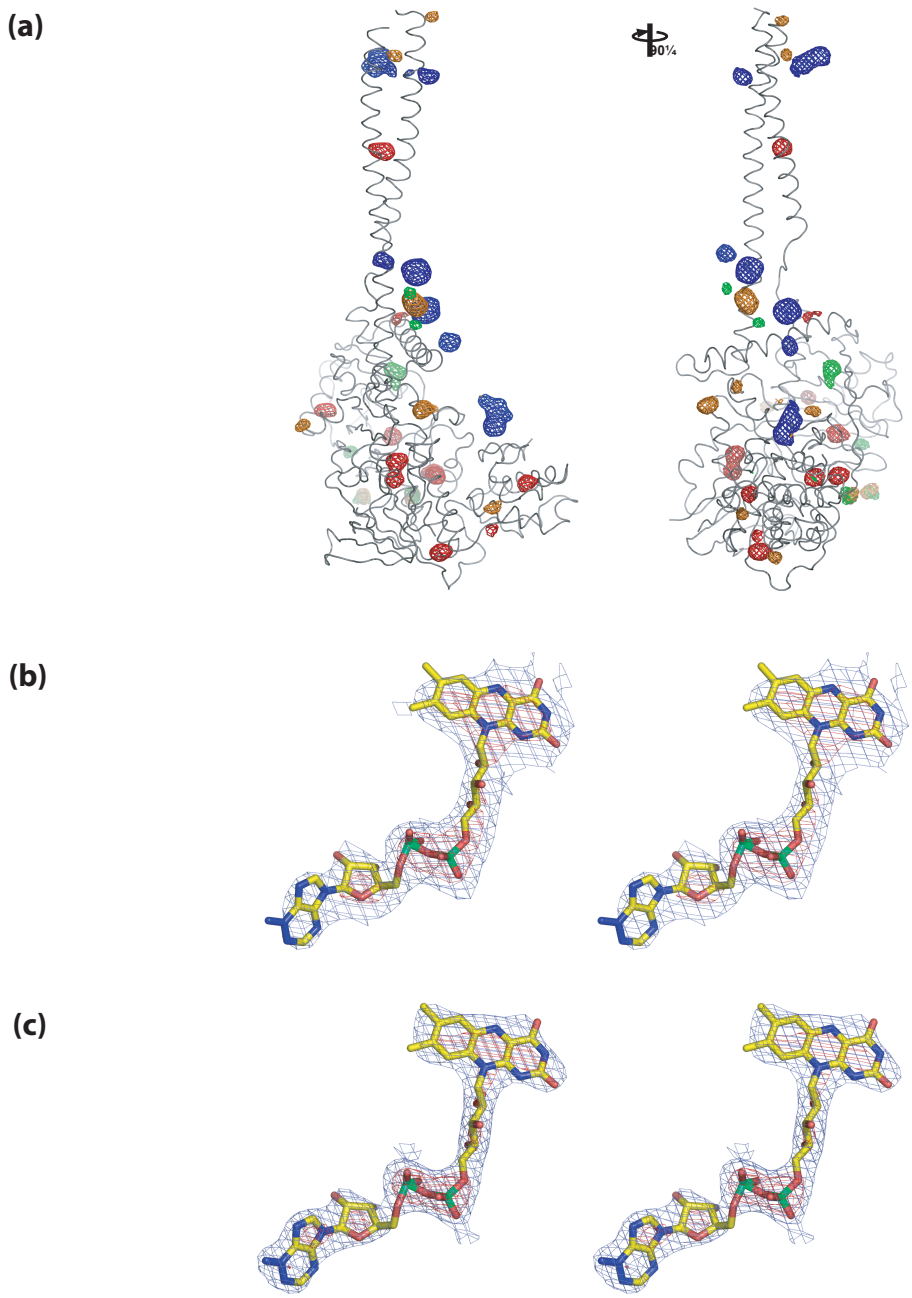


Figure 2.7 Electron Density Maps

(a) Anomalous difference Fourier maps, selenium (red), osmium (orange), mercury (green) and $[Ta_6Br_{12}]^{2+}$ cluster (blue). A 90° rotated view is shown on the right. (b) Experimental electron density map rendered at 4.5 s (red) and 1.0 s (blue) around the FAD co-factor. (c) Final 2|fo|-|fc| electron density map, rendered at the same s levels as in panel b.

Table 1 Crystallographic analysis

	Crystal 1 EMTS		Crystal 2 Se Met	
Data collection				
Space group	$P6_122$		$P6_122$	
Cell dimensions				
a, *b*, *c* (Å)	a=b=187.9, c=108.7,		a=b=187.9, c=108.7,	
α, β, γ (°)	α=β=90°, γ=120°		α=β=90°, γ=120°	
	Hg Peak	*Peak*	*Inflection*	*Remote*
Wavelength	1.0080	0.9796	0.9797	0.9537
Resolution (Å)	20.0-2.90	50.0-3.10	50.0-3.15	50.0-3.65
R_{sym} or R_{merge}	8.8 (65.1)	9.2 (55.0)	10.6 (59.4)	12.0 (54.7)
$I / \sigma I$	19.9 (2.0)	16.3 (1.8)	14.7 (1.8)	11.6 (1.9)
Completeness (%)	98.4 (91.5)	97.9 (82.1)	96.2 (71.9)	99.2 (95.5)
Redundancy	15.3	9.1	8.9	8.4
Refinement				
Resolution (Å)	20.0-2.90			
No. reflections	38,108			
	4,087 (8.6%)			
R_{work} / R_{free}	28.1 / 22.9			
No. atoms	5,153			
Protein	5,077			
Ligand/ion	56			
Water	20			
B-factors	82			
Protein	83			
Ligand/ion	72			
Water	61			
R.m.s deviations				
Bond lengths (Å)	0.008			
Bond angles (°)	1.4 / 1.8			

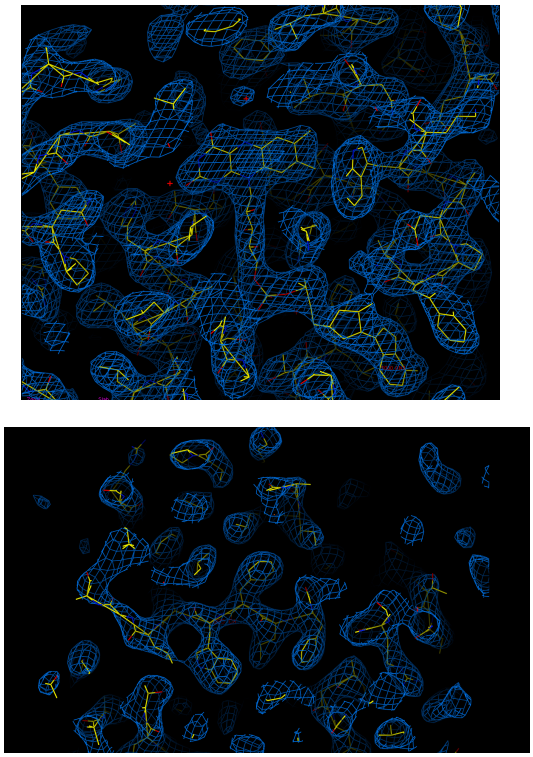
*Highest-resolution shell is shown in parentheses.

Table 2 Crystallographic analysis

		Crystal 3 $[Ta_6Br_{12}]^{2+}$		Crystal 4 EMTS	Crystal 5 K_2OsO_4	Crystal 6 Native
Data collection						
Space group		$P6_122$		$P6_122$	$P6_122$	$P6_122$
Cell dimensions						
a, *b*, *c* (Å)		a=b=189.1, c=107.2		a=b=186.8, c=108.4	a=b=189.9, c=108.6	a=b=188.7, c=107.6
α, β, γ (°)		α=β=90°, γ=120°		α=β=90°, γ=120°	α=β=90°, γ=120°	α=β=90°, γ=120°
	Ta Peak	*Ta Inflection*	*Br Peak*	*Hg Peak*	*Os Peak*	
Wavelength	1.2552	1.2557	0.9202	1.0080	1.1401	0.9796
Resolution (Å)	20.0-3.90	20.0-3.90	20.0-4.10	20.0-4.00	20.0-4.90	20.0-3.15
R_{sym} or R_{merge}	9.9 (48.8)	7.5 (47.1)	8.5 (64.1)	10.7 (67.5)	6.4 (38.5)	6.1 (68.6)
$I / \sigma I$	19.0 (3.6)	23.6 (3.1)	21.6 (2.0)	12.5 (2.3)	24.4 (3.2)	28.4 (2.2)
Completeness (%)	99.5 (99.5)	99.2 (96.7)	96.8 (83.0)	94.1 (95.1)	95.5 (88.6)	99.6 (98.2)
Redundancy	7.4	7.4	12.9	7.4	8.3	4.8

*Highest-resolution shell is shown in parentheses.

(a)



(b)

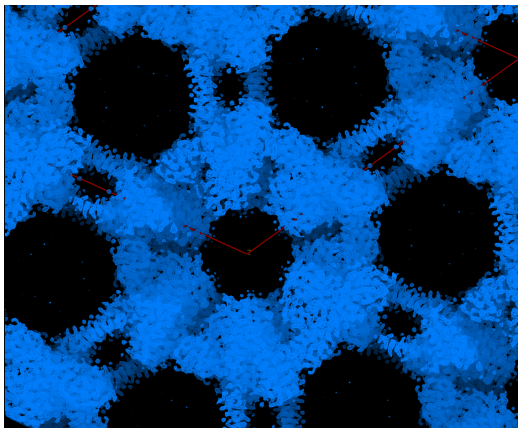


Figure 2.8 Representative 2|Fo|-|Fc| Electron Density

(a) Electron density maps at 2.8 Å resolution. The map was calculated using Fourier coefficients (2|Fo|-|Fc|)exp(-ia), where |Fo| is the observed structure factor amplitude and |Fc| and a are the structure factor amplitude and phase angle calculated from the current model. Blue contours indicate electron density at a 1.5 s level. The protein model and cofactor are shown in stick representation. (b) Electron density map with view of crystal packing and hexagonal crystal symmetry along the 6_1 screw axis in the Z direction.

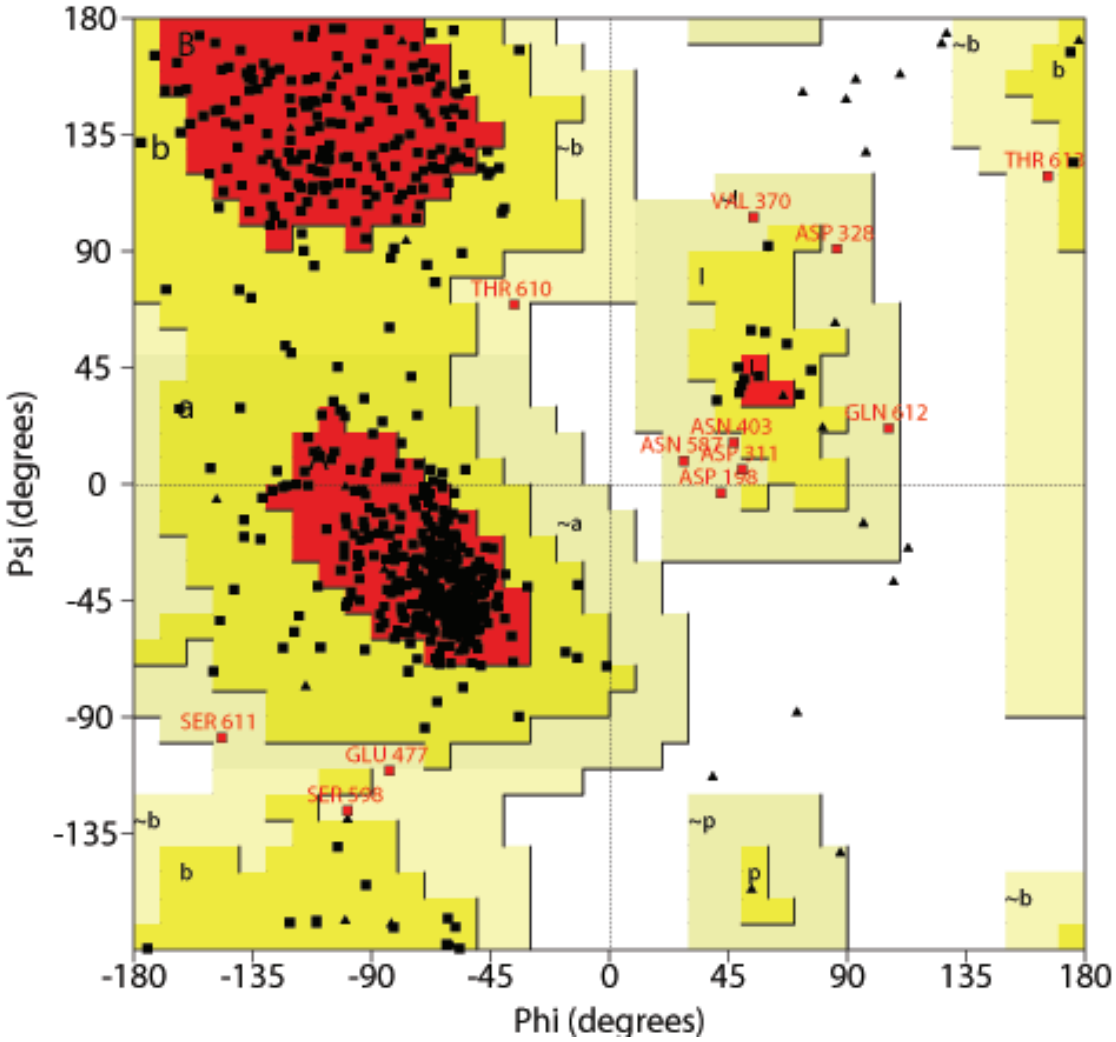


Figure 2.9 Ramachandran Plot of the Human LSD1 Structure

Glycine residues are represented by black triangles and all other residues are represented by black or red squares. The most favorable regions are shaded red, the allowed regions are shaded bright yellow and the generously allowed regions are shaded a pale yellow. Disallowed regions are represented as white.

2.5 Illustrations And Figures

Figures were generated using PyMOL [114]. Molecular surfaces were calculated using MSMS [115]. Sequence alignments were generated using ClustalX [116].

2.6 Photometric Activity Assay

LSD1 mutants were generated by QuikChange mutagenesis (Stratagene) and confirmed by DNA sequencing. Mutant proteins were expressed in 32 liter cultures and purified to homogeneity as described above, but without proteolytic cleavage of the N-terminal hexahistidine tag. The purity and concentrations of the mutant proteins were determined by SDS-PAGE. Western blot analysis of the purified mutant protein was performed using an anti-hexahistadine antibody using standard protocols (Figure 2.10). The purified protein was stored at -80ºC in 20 μl aliquots in a buffer containing 12.5 mM HEPES pH 8.0, 125 mM NaCl, 50% (v/v) glycerol, and 5 mM DTT.

Initial velocity measurements were performed under aerobic conditions using a peroxidase-coupled assay that monitors hydrogen peroxide production as described previously [50]. Briefly, the time courses of the reactions were measured in a Shimadzu UV-1650PC UV visible spectrophotometer equipped with a thermostated cell holder (T=25ºC). The reactions were initiated by addition of 1 μM LSD1 protein and the mixtures

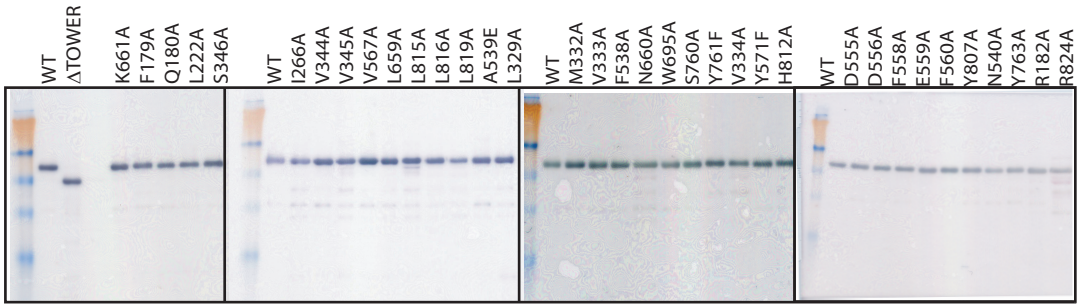


Figure 2.10 Western Blot Analysis of Purified LSD1 Mutants

Hexa-histidine tagged LSD1 mutants where run on 12% denaturing SDS polyacrylamide gels and transferred to PVDF membranes using standard protocols. The PVDF membrane was then immunoblotted for anti-hexa-histidine containing proteins.

contained 50 mM HEPES/NaOH buffer, pH 7.5, 0.1 mM 4-aminoantipyrine (Sigma Aldrich, #A4382), 1 mM 3,5-dichloro-2-hydroxybenzenesulfonic acid (Sigma Aldrich, #238821), 4 units of horseradish peroxidase (Sigma Aldrich, #P8415), and variable concentrations (1-320 μM) of methylated H3-K4 peptide in a total volume of 150 μl. Absorbance changes were monitored at 515 nm, and an extinction coefficient of 26,000 M^{-1} cm^{-1} was used to calculate initial velocity values.

RESULTS AND DISCUSSION

3.1 Overall Architecture of Human LSD1

3.1.1 Domain Organization of Human LSD1

The LSD1 polypeptide chain can be divided into several functional regions (Figure 3.1). The N-terminal region exhibits no predicted structural elements and is not necessary for catalysis, but it contains a putative nuclear localization signal [51]. This region is followed by a SWIRM (from Swi3p, Rsc8p, and Moira) domain, which is found in several histone-interacting proteins [117]. The remainder of the protein is comprised of an oxidase domain, which is capable of demethylating specific histone tail lysine residues in an FAD-dependent manner [44]. Interestingly, the oxidase domain contains an atypical insertion of unknown function not found in other oxidases [4].

3.1.2 Architectural Overview of the Human LSD1 Crystal Structure

The LSD1 polypeptide chain folds into three distinct structural domains. Together, they form a highly asymmetric molecule, which is ~60 Å wide and ~140 Å high (Figure 3.2). Two of them, the N-terminal SWIRM domain and the C-terminal oxidase domain, closely pack against each other to form a core structure from which a ~90 Å-high tower domain protrudes.

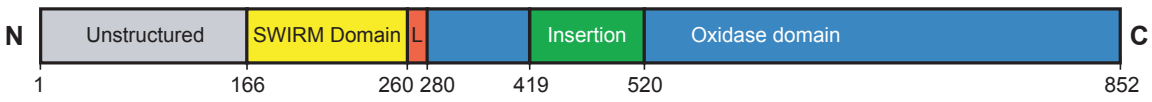


Figure 3.1 Domain Structure of Human LSD1

Domain structure of human LSD1. The unstructured N-terminal region is shown in gray, the SWIRM domain in yellow, the SWIRM-oxidase connector in red, the oxidase domain in blue, and the unique helical insertion not found in other amine oxidases in green.

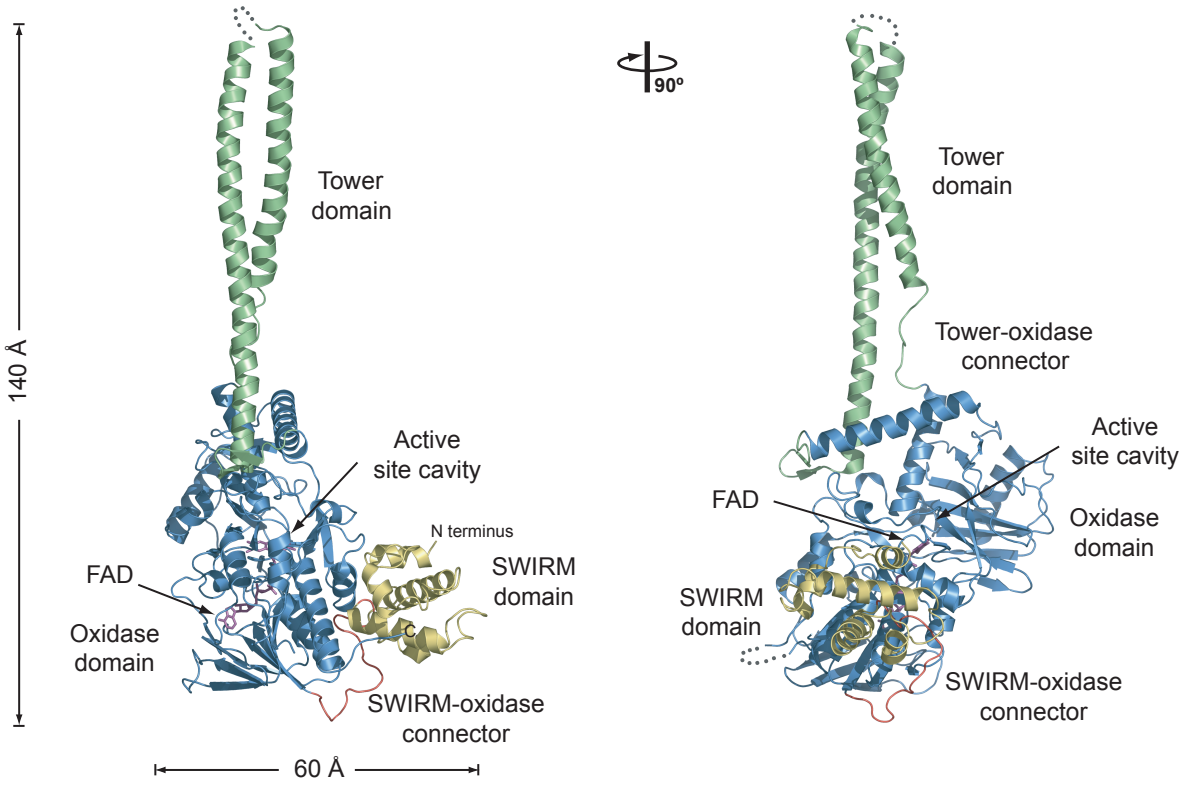


Figure 3.2 Crystal Structure of Human LSD1

The structure of LSD1 is shown in ribbon representation. The SWIRM domain is colored in yellow, the SWIRM-oxidase connector in red, the oxidase domain in blue, and the helical insertion in green. Dimensions are indicated. A 90º rotated view is shown on the right.

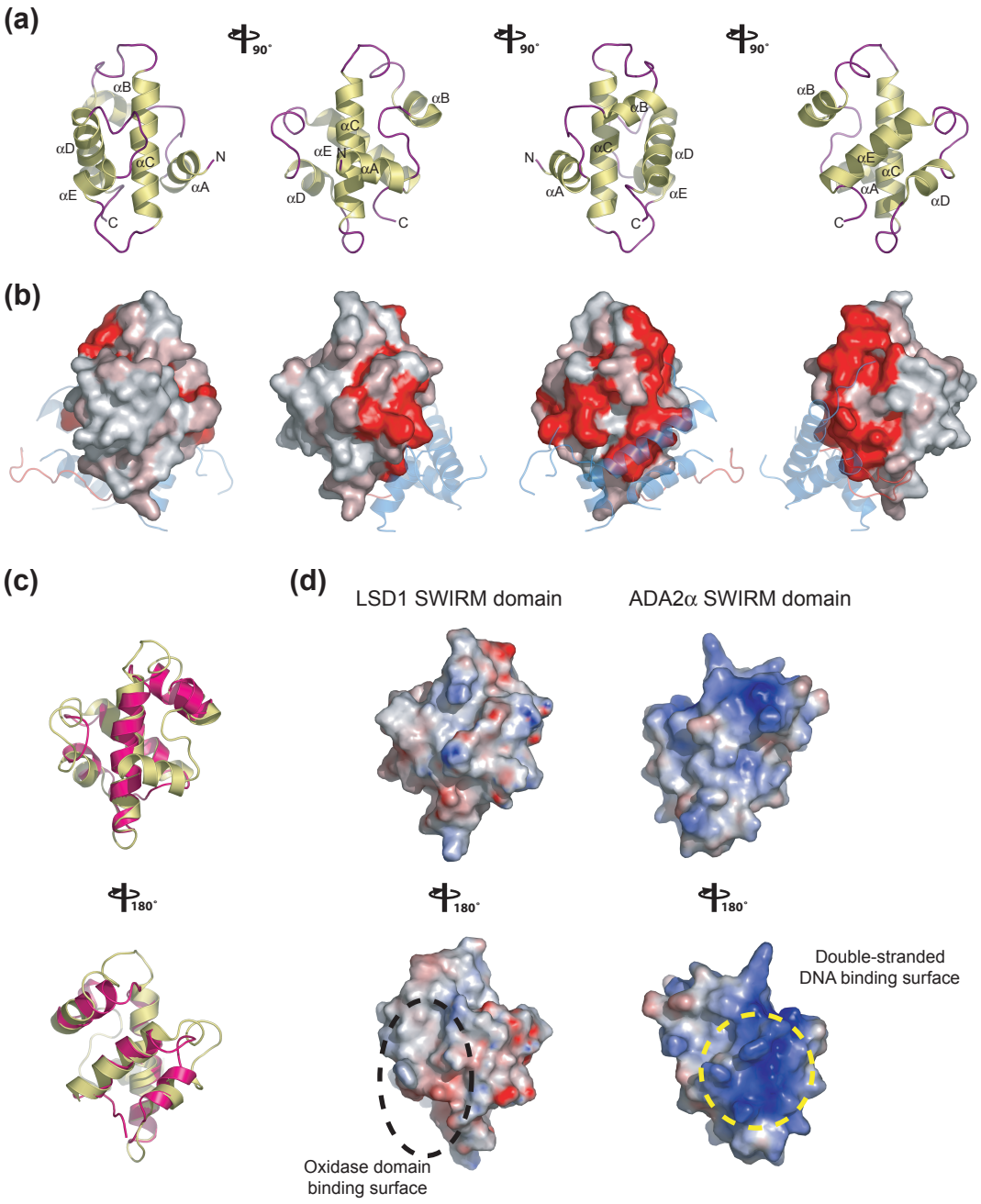
The first domain is the ~100-residue SWIRM domain which adopts a completely helical histone fold. An ~20-residue connector segment links this domain to the C-terminal oxidase domain. The oxidase domain adopts an amine oxidase fold which is present in a number of flavoenzymes, and can be divided into two separate functional lobes. One lobe is required for non-covalent FAD-binding, and the other is required for substrate binding and recognition. The substrate-binding lobe contains a large, open active site cavity, which exposes the isoalloxazine ring of the FAD within its deepest confines. In addition, a ~100 residue insertion which is not found in other amine oxidases forms two long, anti-parallel helices termed the tower domain (Figure 3.2).

3.2 The LSD1 SWIRM Domain

The SWIRM domain, found in a number of proteins involved in chromatin remodeling and histone modification, has been implicated in DNA binding. Recently, the structures of two SWIRM domains, from ADA2α and Swi3, have been determined [118, 119]. Consistent with the previously determined SWIRM domain structures, the LSD1 SWIRM domain consists of a helical bundle containing a long central helix, αC, which separates two smaller helix-turn-helix motifs αA/B and αD/E (Figure 3.3). The interface between the SWIRM and oxidase domains buries a

Figure 3.3 The SWIRM domain of LSD1.

(a) Ribbon representation of the LSD1 SWIRM domain shown from different angles, with α-helices shown in yellow and coil regions in purple. (b) Surface representation shown from the same angles as in panel (a). The surface is colored from gray (< 70% identity) to red (100% identity), according to the sequence conservation. The interacting structural elements of the oxidase domain are shown transparently in blue (helices) and red (SWIRM-oxidase connector). (c) Superposition of the SWIRM domains from LSD1 (yellow) and Ada2α (magenta) in ribbon representation. A 180° rotated view is shown on the bottom. (d) Surface representations of the SWIRM domains of LSD1 and Ada2α in the same orientation as in panel (c). The electrostatic potential is plotted from red to blue onto the surface between -10kBT/e and +10kBT/e. The surfaces that are involved in oxidase domain binding (LSD1) and that have been implicated in double-stranded DNA binding (Ada2α) are indicated.

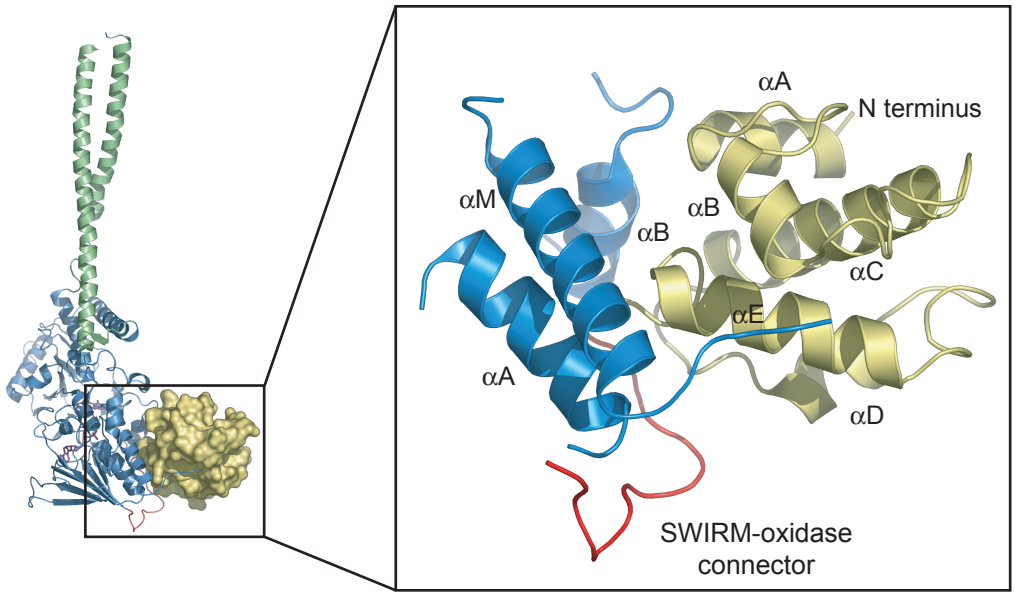

(a)
90°
90°
90°
αB
αC
αD
αE
αA
N
C
(b)
(c)
180°
(d)
LSD1 SWIRM domain
ADA2α SWIRM domain
180°
180°
Double-stranded
DNA binding surface
Oxidase domain
binding surface

surface area of approximately 1,680 $Å^2$ and involves three helices from both the SWIRM (αA^{SW}, αB^{SW} and αE^{SW}) and oxidase (αA^{OX}, αB^{OX} and αM^{OX}) domains (Figure 3.4a). In addition, the SWIRM domain engages the C-terminal tail of the oxidase domain (Figure 3.4a). This interface exhibits a high degree of hydrophobicity, and the two domains are primarily held together by an extensive network of van der Waals contacts. Through the interaction of both domains, a highly conserved cleft in the vicinity of the active site cavity is formed (Figure 3.4b).

The SWIRM domains of other histone modification proteins have been shown to bind DNA. DNA binding in these proteins is achieved by a positively charged patch on the SWIRM domain surface, which is comprised of a series of conserved residues. In contrast to these other SWIRM domains, the residues that typically compose this DNA binding patch are not conserved in LSD1 (Figure 3.3) [118, 119]. We predict that the SWIRM domain of LSD1 is not directly involved in DNA binding. However, the residues that are involved in oxidase domain binding as well as those found in the cleft between the oxidase and SWIRM domains are essentially invariant across species, while the remaining surface residues are variable (Figures 3.3 and 3.5).

Figure 3.4 Structure of the SWIRM Domain of LSD1. (a) Ribbon representation of the SWIRM-oxidase interface. The SWIRM domain is shown in yellow, the portion of the oxidase domain involved in the interaction in blue, and the SWIRM-oxidase connector in red. (b) Surface rendition of a close-up view of the cleft formed by the SWIRM and oxidase domains. The surface is colored according to panel a. The residues that form the cleft between the two domains are indicated, and their contribution to the surface is colored in orange. A ~60° rotated view is shown on the right.

(a)


αA
N terminus
αM
αB
αB
αC
αE
αA
αD
SWIRM-oxidase
connector

(b)

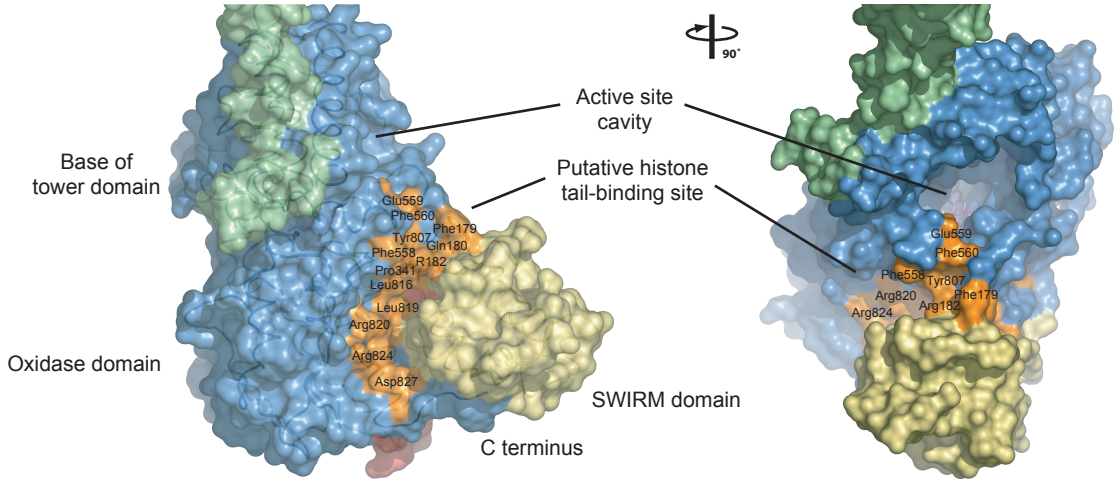
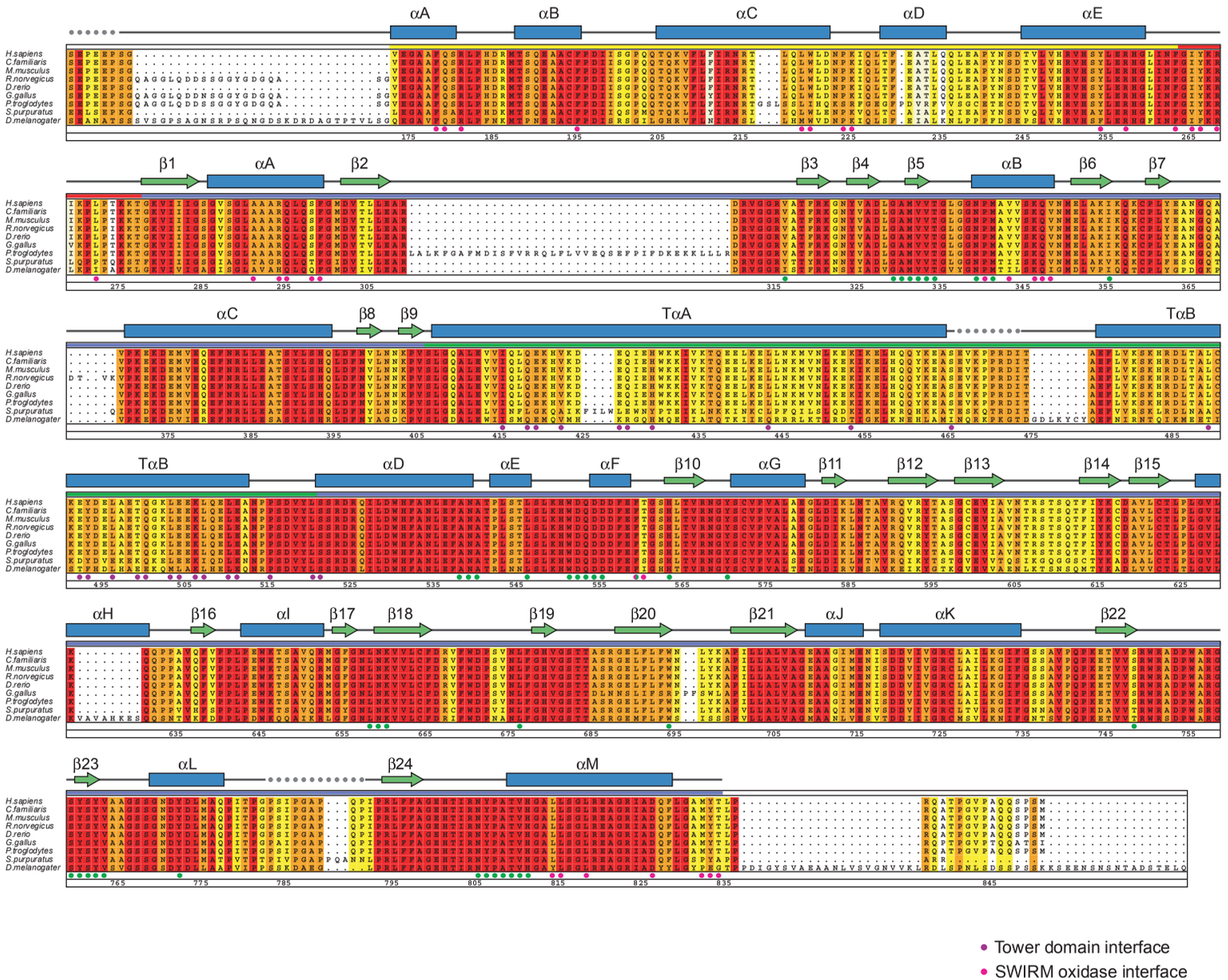

90°
Active site
cavity
Base of
tower domain
Putative histone
tail-binding site
Glu559
Phe560
Phe179
Tyr807
Gln180
Phe558
R182
Pro341
Leu816
Leu819
Arg820
Arg824
Asp827
Oxidase domain
SWIRM domain
C terminus
Glu559
Phe560
Phe558
Tyr807
Arg820
Phe179
Arg824
Arg182

Figure 3.5 Multi-Species Sequence Alignment of LSD1 Homologs

Alignment of LSD1 homologs from H. sapiens (GI:3043726), C. familiaris (GI:73950246), M. musculus (GI:37360004), R. norvegicus (GI:62649867), D. rerio (GI:68387817), G. gallus (GI:50759643), P. troglodytes (GI:55586433), S. purpuratus (GI:72031239), and D. melanogaster (GI:20151661). The secondary structure is indicated above the sequence as blue cylinders (alpha helices), green arrows (beta sheets), gray lines (coil regions), and gray dots (disordered residues). The numbering below the alignment is relative to human LSD1. Overall sequence conservation at each position is shaded in a color gradient from yellow (70% similarity) to dark red (100% identical). The participation of various residues in the formation of the SWIRM-oxidase interface (magenta dots), active site cavity (green dots), and tower domain interface (purple dots) is indicated below the aligned sequences.


αA
αB
αC
αD
αE
β1
β2
β3
β4
β5
β6
β7
β8
β9
TαA
TαB
αF
β10
αG
β11
β12
β13
β14
β15
αH
β16
αI
β17
β18
β19
β20
β21
αJ
αK
β22
β23
αL
β24
αM
H.sapiens
C.familiaris
M.musculus
R.norvegicus
D.rerio
G.gallus
P.troglodytes
S.purpuratus
D.melanogaster
Tower domain interface
SWIRM oxidase interface
Active site surface residues

3.3 Biochemical Analysis of Mutants located in the SWIRM-Amine Oxidase Interface

We performed a series of biochemical experiments to determine whether the SWIRM domain has an impact on the catalytic capacity of the enzyme. Any mutation that weakens the hydrophobic interactions in the interface between the SWIRM and oxidase domains severely reduces catalytic activity (Figure 3.6). Removal of the SWIRM domain renders the protein insoluble when expressed in *E.coli*. Given the proximity of the SWIRM domain to the active site of the enzyme, as well as the presence of the highly conserved cleft formed by the SWIRM and oxidase domains, it is intriguing to speculate that this cleft engages the histone tail substrate, thus explaining the necessity of the SWIRM domain for catalysis. The results of the biochemical analysis will be discussed in detail below.

3.4 The Active Site Cavity of LSD1

The oxidase domain of LSD1 consists of two functional lobes, one that binds FAD and another that recognizes and binds substrate (Figure 3.7a). The active site cavity of LSD1 is located within the substrate-binding lobe of the oxidase domain and is substantially more spacious and open than in other FAD-dependent amine oxidases (Figure 3.7a,b). Polyamine oxidase (PAO), for example, contains a 30 Å long catalytic tunnel with only two small openings on the surface of the oxidase domain, which ensures

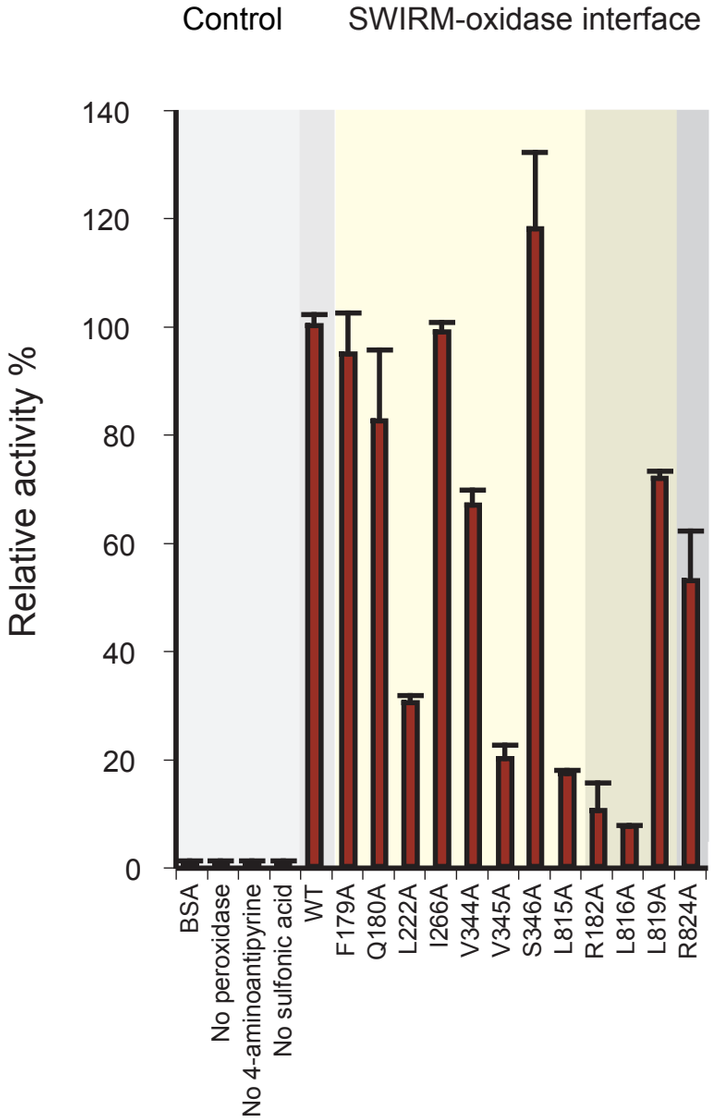
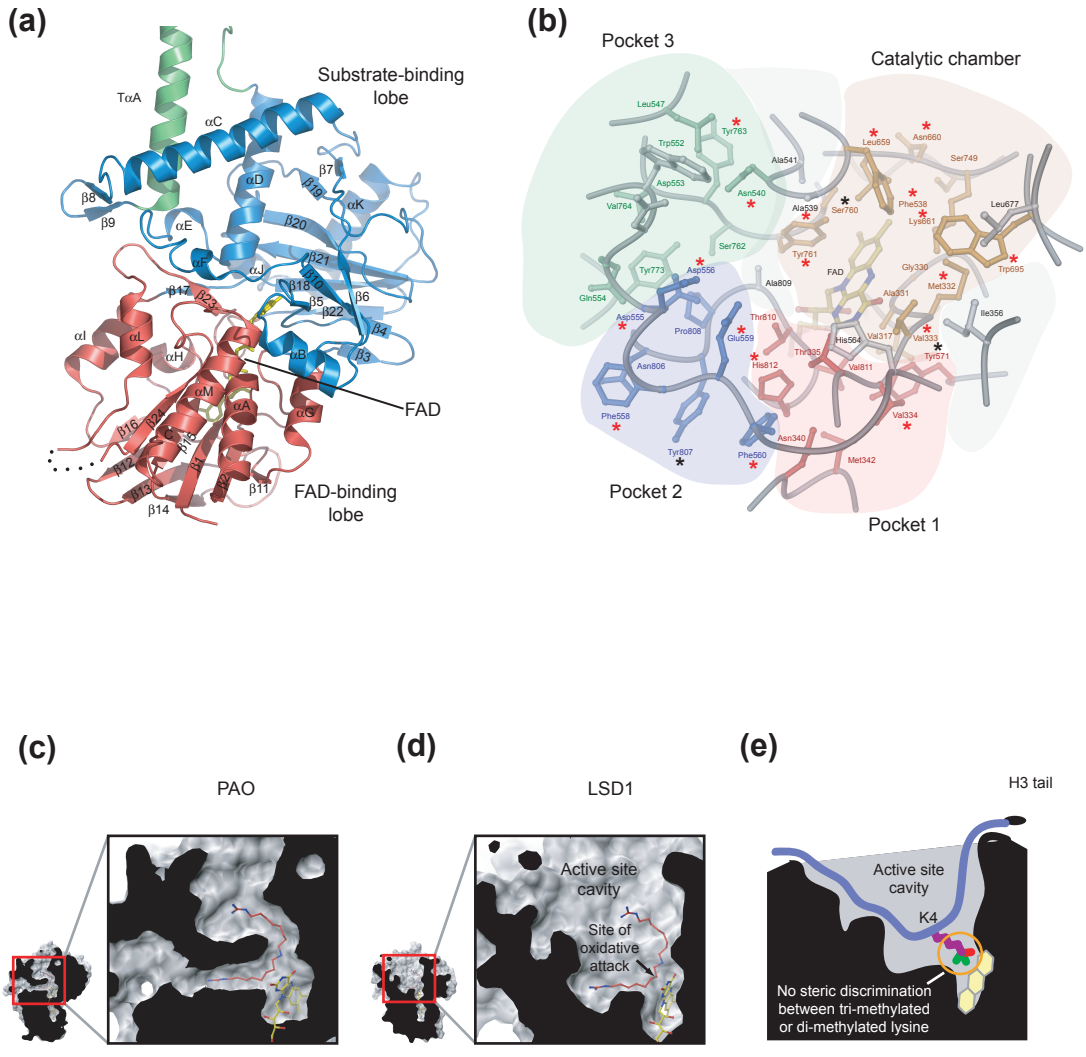


Figure 3.6 Biochemical Analysis of LSD1 Mutants Located in the SWIRM-Oxidase Interface

Initial velocities of LSD1 mutants. Initial velocities were determined in a photometric activity assay, using 1 μM LSD1 mutant protein and 40 μM di-methylated H3K4 peptide, and normalized against wildtype LSD1 activity. Each data point represents at least three independent measurements. The error bars show the standard deviation. LSD1 mutants are grouped according to their location in the protein the SWIRM-oxidase interface (yellow),residues on the periphery of the interface that also contribute to the surface of the cleft (brown), and putative histone tail-binding groove (gray) are indicated.

Figure 3.7 Structure of the Oxidase Domain and Active Site Cavity of LSD1

(a) Ribbon representation of the oxidase domain of LSD1. The FAD-binding domain is shown in red, the substrate-binding domain in blue, and the tower domain in green. The FAD molecule is shown in yellow stick representation. (b) The catalytic importance of residues in the active site cavity. The locations of the various pockets within the active site cavity are illustrated. Mutated residues that reduce the activity by more than 50% (red stars) and residues that had little effect on the activity (black stars) are indicated. (c) The active site of PAO shown in a cutaway surface rendition to illustrate the U-shaped catalytic channel. The PAO inhibitor Guazatine is shown in red. (d) The active site cavity of LSD1 shown in a cutaway rendition in the same orientation as in panel (c). The PAO inhibitor Guazatine is modeled into the active site of LSD1 as a marker for the approximate position of the ε nitrogen atom of the methylated lysine substrate. (e) Schematic representation of the active site cavity of LSD1 illustrating the approximate positioning of the histone tail and methylated lysine 4. The backbone of the histone tail is illustrated in blue, the side chain of H3K4 in purple, and the methyl groups in green and red.


(a)
Substrate-binding lobe
FAD
FAD-binding lobe
(b)
Pocket 3
Catalytic chamber
Pocket 2
Pocket 1
(c)
PAO
(d)
LSD1
Active site cavity
Site of oxidative attack
(e)
H3 tail
Active site cavity
K4
No steric discrimination between tri-methylated or di-methylated lysine

substrate specificity for elongated polyamines like spermine (Figure 3.7c) [120]. In contrast, the active site of LSD1 consists of a ~15 Å deep cavity spanning ~25 Å at its widest opening (Figure 3.7d). The surface of the cavity is comprised of ~50 residues which are essentially invariant across different species, underlining the importance of each residue. Based on sheer size, it is immediately obvious that more than only the side chain of the substrate lysine is inserted into the active site. Indeed, several residues of the histone tail must enter the cavity in order to properly position the methylated substrate lysine residue (Figure 3.7a).

Inside the active site cavity, there are four major invaginations with distinct chemical properties that could be utilized for the specific binding of side chains (Figure 3.7a). Furthermore, it seems plausible that these pockets are capable of sensing not only the histone tail sequence, but also its chemical modifications. The first pocket exposes the isoalloxazine ring of the FAD co-factor and forms the catalytic chamber of LSD1. The residues that create this hydrophobic chamber and that lie in direct vicinity of the isoalloxazine ring are Val317, Gly330, Ala331, Met332, Val333, Phe538, Leu659, Asn660, Lys661, Trp695, Ser749, Ser760, and Tyr761. These residues are required for the exact positioning of both the FAD co-factor and the substrate lysine. In other oxidases, the isoalloxazine ring is usually found in conjunction with a positively charged residue which is required for proper FAD-binding and stabilization of the anionic form of the

reduced co-factor. The residue that fulfils this function in LSD1 is Lys661. The remaining three pockets of the active site cavity are likely required for the binding of histone tail residues directly adjacent to the substrate lysine. These additional pockets would also enable LSD1 to 'read' the chemical modifications of the histone tail residues, thereby selecting the optimal substrate. The first putative side-chain binding pocket, which is next to the catalytic chamber, is generated by Val334, Thr335, Asn340, Met342, Tyr571, Thr810, Val811, and His812. The second binding pocket is generated by Phe558, Glu559, Phe560, Asn806, Tyr807, and Pro808, and the third is generated by Asn540, Leu547, Trp552, Asp553, Gln554, Asp555, Asp556, Ser762, Tyr763, Val764, and Tyr773. In addition, Ile356, Ala539 and Ala809 lie in between these pockets. Each pocket creates a distinct chemical environment potentially capable of sensing different residues and their modifications.

Recently, the structure of the LSD1-CoREST complex has been determined in a ternary complex with a H3 tail peptide substrate analog. A 2.7 Å crystal structure was obtained after treating the LSD1-CoREST complex with a suicide inhibitor that was covalently tethered to the flavin ring [121]. The inhibitor contained the first 21 amino acid residues of histone H3 N terminal tail where the lysine residue at position four was modified to contain a *N*-methylpropargylamine group [121]. Only the first seven residues of the H3 tail contained interpretable density [121]. The

first seven residues adopted an unusual fold consisting of three gamma turns, which involve hydrogen bond interactions between the n and n+2 residues in a protein chain [121]. These sets of interactions form pseudo seven membered rings (Figure 3.8a) [121]. The crystal structure revealed a series of van der Waals interactions and hydrogen bonds between the H3 peptide and the pockets in close proximity to the active site of LSD1 (Figure 3.8b,c) [121]. These interactions are consistent with the predictions set forth by our structure in conjunction with our biochemical analysis. The amino terminal end of the H3 peptide makes an electrostatic interaction with Asp555 of LSD1 (Figure 3.8b) [121]. Arg2 of the H3 peptide form a salt bridge with Asp556 and the aliphatic portion of the arginine residue makes contacts with Trp552 of LSD1 (Figure 3.8b) [121]. Thr3 of H3 interacts with Asn536 (Figure 3.8c) [121]. The side chain of Gln5 of H3 makes a hydrogen bond with the side chain amide group of Asn535 (Figure 3.8c) [121]. The backbone carbonyl group of Gln5 also makes a hydrogen bond with the side chain amide group of Gln358 (Figure 3.8c) [121]. It is believed that the specific interaction between H3 residues Arg2, Thr3, and Gln5 are mainly responsible for the specificity towards H3K4. Weak density for amino acid residues 8-21 of the H3 was observed and was detected extending downwards towards the SWIRM-Oxidase interface as predicted by our study [121].

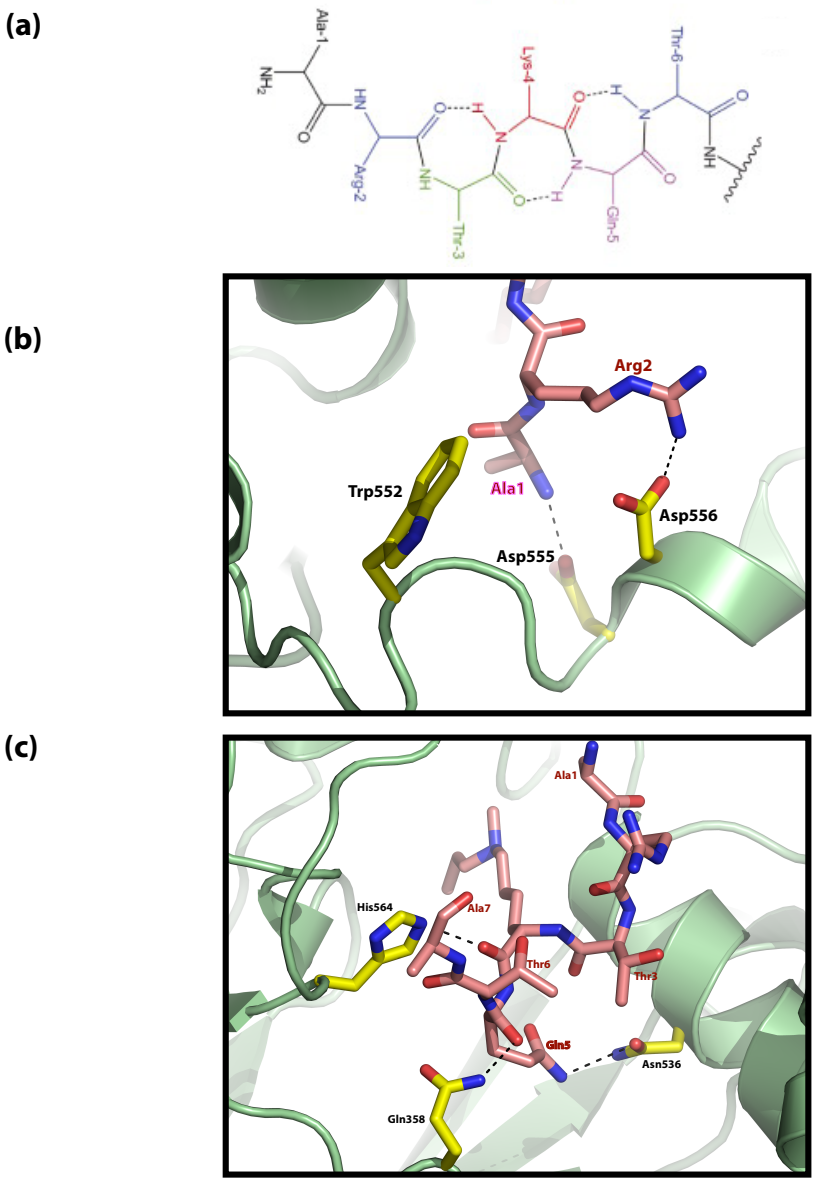


Figure 3.8 Complex Crystal Structure of LSD1 Bound to SubstrateH3 Peptide

(a) Chemical representation of the first six amino acids seen in the density of a 21-residue histone H3 N-terminal peptide forming the unusual gamma turn structure (122). (b) and (c) Close up view of LSD1 in ribbon representation with residues labeled in black and histone H3 as stick representation with residues labeled in red (PDB accession code 2UXN).

3.5 Enzymatic Discrimination of Tri-methylated Lysine

An important aspect of LSD1's enzymatic activity is its inability to act on tri-methylated H3K4 [44]. Yet, based on our LSD1 structure, the catalytic chamber is just large enough to accommodate a tri-methyl lysine residue. In fact, we found that in an *in vitro* mono-methylated H3-K4 demethylation assay, a tri-methylated H3-K4 peptide acts as a competitive inhibitor (Figure 3.9), which argues that the tri-methylated peptide readily binds in the active site but cannot be converted.

Two possible enzymatic mechanisms have been proposed for the demethylation reaction of LSD1 (Figure 3.10a,b) [123]. The first starts with an oxidation step in which a single electron is transferred. A carbon radical intermediate is formed, which becomes a carbinolamine via the addition of a hydroxy radical (Figure 3.10a). The alternate pathway starts with a two-electron oxidation process via an aminium cation. The cation is hydrolyzed to the common intermediate carbinolamine, which readily breaks down to formaldehyde and lysine (Figure 3.10b). This second pathway depends on the availability of the free electron pair at the methylated lysine residue. Therefore, a tri-methylated lysine could not be demethylated via such a reaction pathway. Regardless of the pathway, other flavoenzyme complex structures illustrate that, despite the great diversity of reactions catalyzed, the atom that is subject to oxidative attack is located ~3.5 Å away from N5 of the flavin ring, with an angle of 96-117º

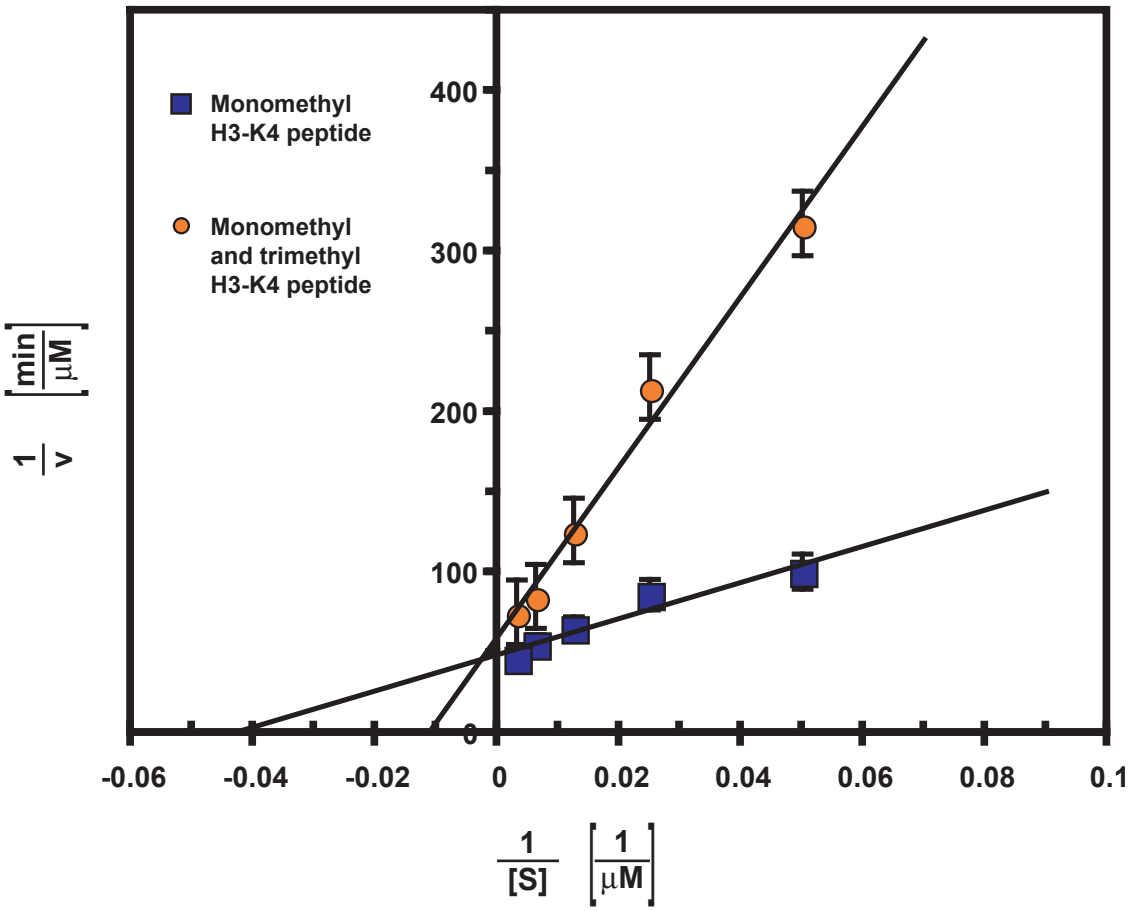


Figure 3.9 Tri-methylated H3K4 Peptide Acts as a Competitive Inhibitor of LSD1 Activity

Initial reaction velocities of LSD1 were determined in a photometric activity assay at various concentrations of mono-methylated H3K4 peptide in the absence and presence of tri-methylated H3K4 peptide. Each data point represents at least three independent measurements. The error bars show the standard deviation. The data were fitted by linear regression analysis.

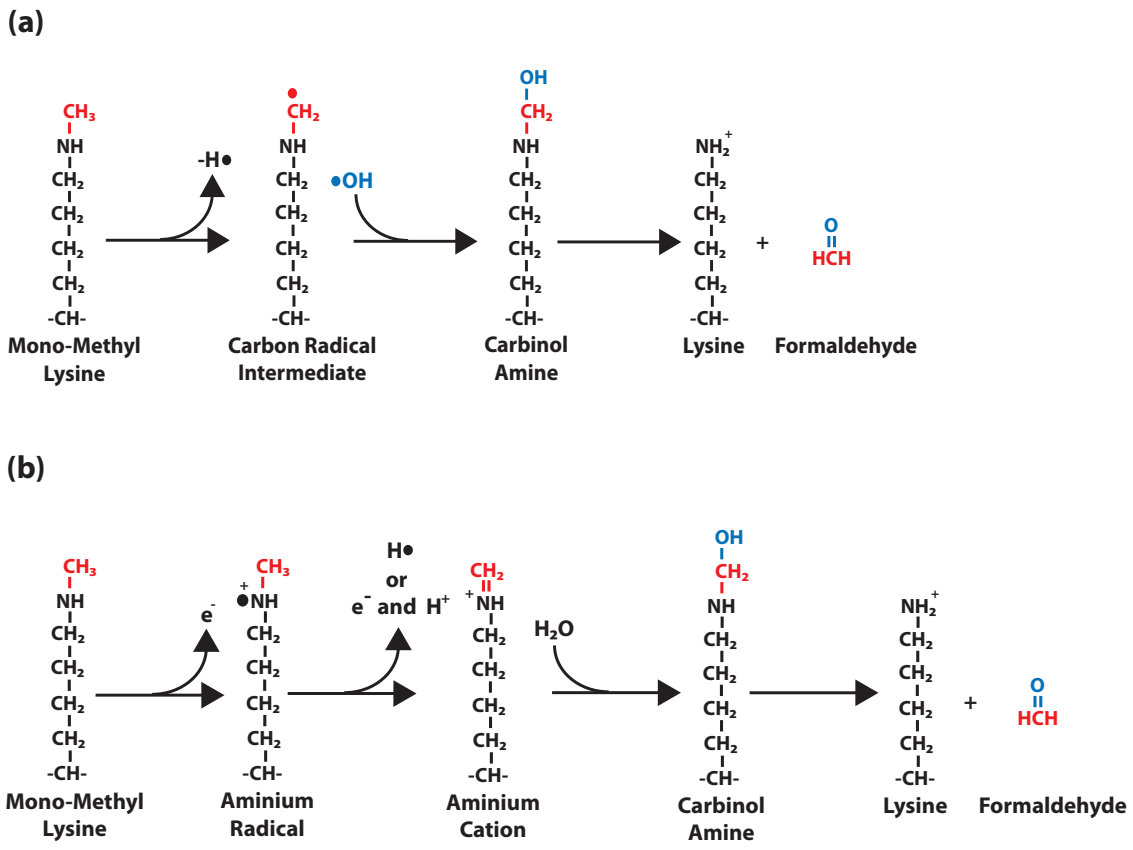


Figure 3.10 Proposed Chemical Mechanisms of LSD1 Demethylation Activity

The two proposed demethylation reaction mechanisms of LSD1 are illustrated in (a) and (b) as discussed in the text.

[126]. If we consider this finding, it is immediately obvious that the catalytic chamber of LSD1 is incapable of discriminating between different methylation states of the substrate lysine in the state crystallized (Figure 3.7d,e). Based on these observations, we conclude that the inability of LSD1 to demethylate tri-methylated lysine residues is most likely due to the chemical mechanism itself rather than steric discrimination, which rules out the radical pathway but supports the two-electron transfer pathway.

3.6 The Tower Domain

The tower domain is formed by two long alpha helices, TαA and TαB, which pack against each other without coiling (Figure 3.11). The first helix is ~90 Å long and formed as an LSD1-specific extension which has not yet been observed in other oxidases. The second helix, TαB, is ~60 Å long and connects to TαA at the top of the tower via a short disordered loop. The two helices bind tightly to each other by hydrophobic interactions. A highly conserved ~10-residue connector segment attaches TαB to helix αD of the substrate binding lobe. Helices αD^{OX} and TαA form the base of the tower and are stabilized by interactions with helix αC^{OX}, which packs perpendicularly to both helices. Helix αD^{OX} is involved in and essential to the formation of the catalytic chamber by providing Phe538. In addition,

helix αD^{OX} contributes hydrogen bonds to Tyr761 that position it close to the isoalloxazine ring of the FAD co-factor.

The morphology of the catalytic chamber is crucial for the precise positioning of the methylated amino headgroup of the substrate lysine, and thus for catalysis. Therefore, the tower domain is directly connected to the catalytic center of the enzyme. It is interesting to speculate that the binding of other proteins to the tower domain, in particular to the tower-oxidase connector, regulates the size of the chamber, which in turn would allosterically regulate the catalytic activity of LSD1 (Figure 3.11). Indeed, it is reasonable to assume that the tower domain is a binding platform for members of the CoREST or other histone modification complexes, which was shown to be the case [67]. The hetero-dimeric complex between LSD1 and a fragment consisting of the linker between SANT1 and SANT2 and the SANT2 domain of CoREST was crystallized (figure 3.12) [67].

The structure reveals that the interaction between CoREST and LSD1 is mediated via the linker, which forms a helix that packs against the tower domain helices of LSD1 [67]. The SANT domain rests close to the apex of the tower domain (Figure 3.12) [67]. Unfortunately, this CoREST fragment did not induce any significant changes in the structure of LSD1.

In the biochemical experiments described in greater detail below, we have found that a towerless mutant, $LSD1^{\Delta TOWER}$, has greatly reduced catalytic activity. At present, little evidence for an allosteric regulation of

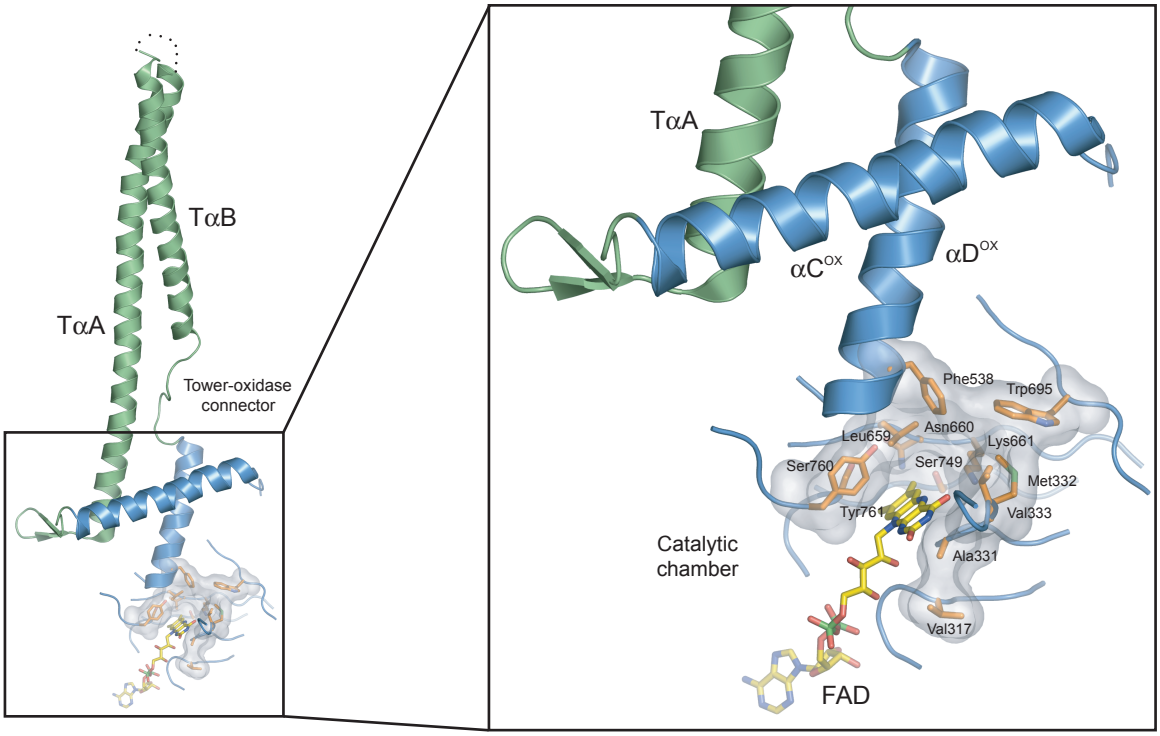


Figure 3.11 Allosteric Regulation of LSD1.

The base of the LSD1 tower domain is crucially involved in the formation of the catalytic chamber. The tower α-helices are show in green, the oxidase domain α-helices in blue. The catalytic chamber is shown in a transparent surface rendition, with critical residues shown in orange stick representation. The FAD molecule is shown in yellow stick representation.

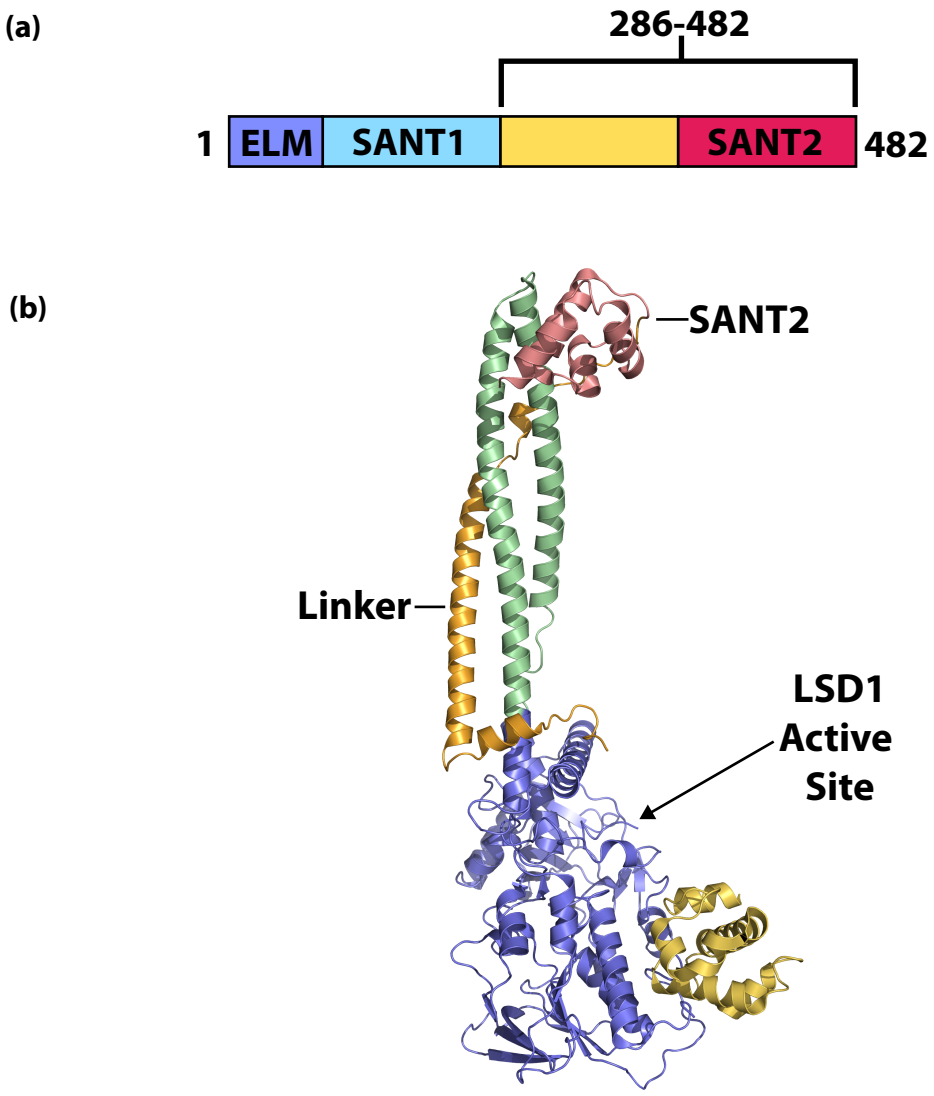


Figure 3.12 Structure of the Hetero-Dimeric Human LSD1-CoREST Complex

(a) Domain organization of human CoREST. A CoREST fragment comprising residues 286-482 was used for co-crystallization with LSD1. (b) Ribbon representation of the crystal structure of human LSD1-CoREST complex (PDB accession code 2IW5). The SWIRM domain of LSD1 is colored in yellow; the oxidase domain of LSD1 is colored in blue; the tower domain of LSD1 is colored in green; the linker domain of CoREST is colored in orange; the SANT2 domain of CoREST is colored in red.

LSD1's activity exists, and the significance of the tower domain merits further examination.

3.7 Biochemical Characterization

In order to assess the importance of the various domains of LSD1 in catalysis and histone recognition, we generated and expressed 47 LSD1 mutants. These mutants can be divided into four categories: (1) mutants within the active site cavity, (2) mutants that interfere with the interaction between the SWIRM and oxidase domains, (3) mutants in the cleft between the SWIRM and oxidase domains, and (4) tower domain deletion mutants. We were able to purify 37 of these mutants to homogeneity in milligram amounts, and their activities were measured in an *in vitro* photometric H3K4 peptide demethylation assay [125, 126]. All of the point mutants we examined were indistinguishable from the crystallized wildtype protein on a gel-filtration column and in their ability to bind FAD. The results from these experiments are summarized in Figure 3.13.

Within the active site cavity, we mutated 23 residues, including all 10 residues in the catalytic chamber, several residues in each of the three additional putative peptide-binding pockets, and one residue, A539E, which obstructs the active site cavity. Of the 10 mutated residues in the catalytic chamber, 9 had greatly reduced activity (figure 3.13). These results strongly support the importance of each residue in the catalytic

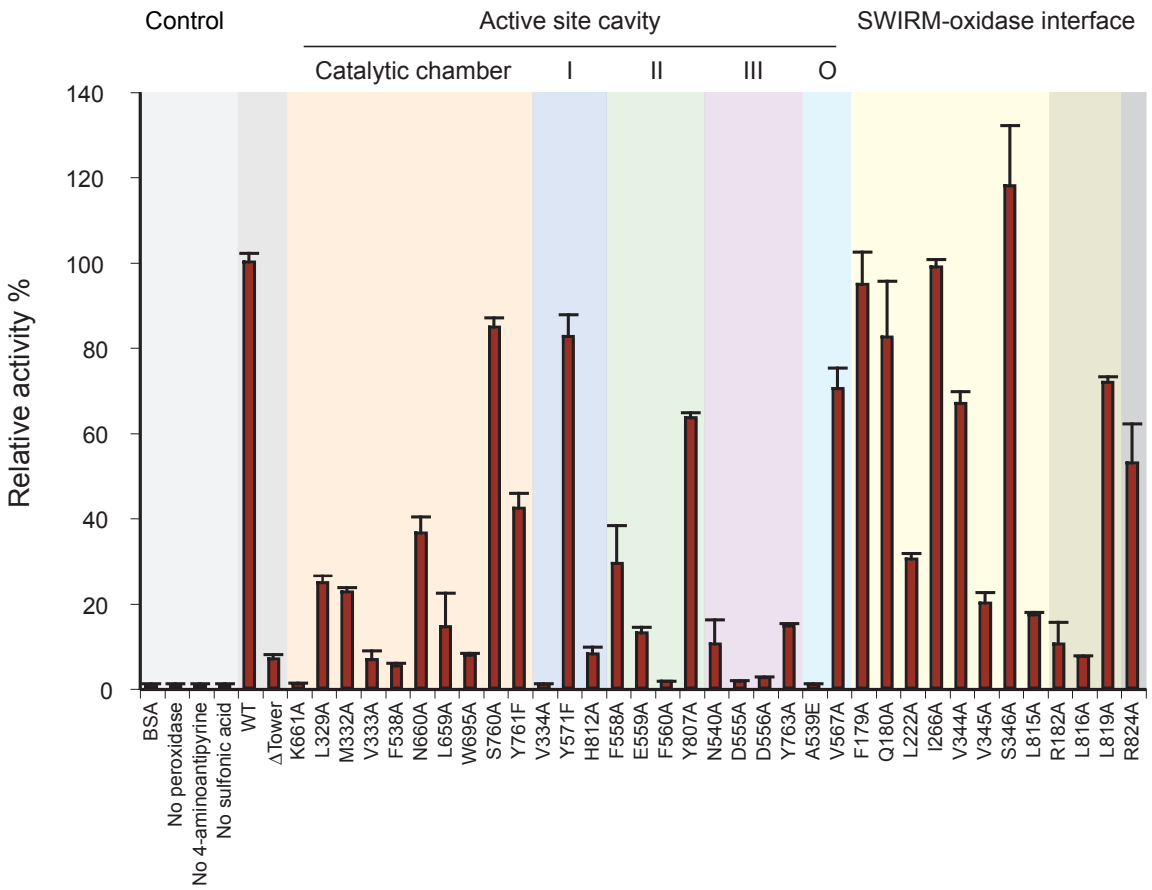


Figure 3.13 Initial Velocities of LSD1 Mutants

Initial velocities were determined in a photometric activity assay, using 1 μM LSD1 mutant protein and 40 μM dimethylated H3K4 peptide, and normalized against wildtype LSD1 activity. Each data point represents at least three independent measurements. The error bars show the standard deviation. LSD1 mutants are grouped according to their location in the protein: mutants in the active site cavity (catalytic chamber in orange, pocket I in blue, pocket II in green, pocket III in purple, and other residues in the active site cavity (O) in cyan), the SWIRM-oxidase interface (yellow), residues on the periphery of the interface that also contribute to the surface of the cleft (brown), and putative histone tail-binding groove (gray) are indicated.

chamber for the precise positioning of the flavin ring with respect to the methylated amino group of the substrate lysine. Similar to the residues in the catalytic chamber, we identified at least two residues in each of the three additional putative peptide-binding pockets that also had greatly reduced catalytic activity (figure 3.13). These residues are V334A and H812A in pocket I, F558A, E559A, and F560A in pocket II, and N540A, D555A, D556A, and Y763A in pocket III. Finally, the cavity obstruction mutant A539E had no detectable catalytic activity. Together, these experiments strongly support the hypothesis that the pockets within the active site cavity play a crucial role in the recognition and precise positioning of the substrate peptide and thus suggest that these pockets may be capable of sensing additional modifications of the histone tail.

We mutated 11 residues within the interface between the SWIRM and oxidase domains. Many mutants within the interface had greatly reduced activity. In general, mutants closer to the center of the interface (L222A, V345A, L815A) had a greater effect on the activity than mutants on the periphery (F179A, Q180A, I266A, V344, S346A) (figure 3.13). However, residues on the periphery of the interface that also contribute to the surface of the cleft (R182A, L816A, L819A) had greatly reduced activity (figure 3.13). Furthermore, the R824A mutant which lies in that cleft and comes from the oxidase domain exhibits a substantial decrease in activity (figure 3.13). Hence, mutants in the interface had a great effect on the

demethylation activity, even though the SWIRM domain is not directly involved in the formation of the active site. Finally, the SWIRM domain appears to be critical for the stability of LSD1 because a mutant lacking this domain ($LSD1^{\Delta SWIRM}$) was insoluble in our bacterial expression system.

3.8 Summary and Future Direction

The initial discovery of two histone demethylases, LSD1 and JHDM1, has revealed two families of enzymes that are capable of demethylating mono-methylated and di-methylated lysine residues of histone 3 tails [44, 52]. Here, we present the first crystal structure and a comprehensive biochemical structure-function analysis of the human histone demethylase, LSD1. Based on our experiments, we conclude that the stable interaction between the SWIRM and oxidase domains is critical for catalysis, suggesting a role in substrate recognition by generating an additional binding site for the histone tail. LSD1 contains an active site cavity that is large enough to accommodate several residues of the substrate histone tail and is likely capable of sensing their chemical modifications. However, in the state that we crystallized, the large active site cavity does not seem to provide any structural features that would enable LSD1 to determine the methylation state of the substrate lysine. We speculate that, in the absence of such a steric-sensing mechanism, there must be a chemical basis for the discrimination of tri-methyl lysine as a substrate. This implies

that, of the two possible enzymatic mechanisms that have been proposed for LSD1, demethylation occurs via the two-electron-transfer pathway [123]. On the other hand, JHDM1 has been found to contain a hydroxylase domain as the catalytic core. This enables the enzyme to attach a hydroxyl radical directly to the methyl group of the methylated lysine and should therefore allow it to demethylate tri-methylated lysine.

Recent genetic and biochemical analysis of LSD1 has unveiled a palette of functional diversity. Experimental evidence has indicated that LSD1 plays a role in various biological processes such as maintenance of boundary elements adjacent to centromeric heterochromatic regions to a pivotal role in mammalian development. The diverse processes that LSD1 regulates present a series of interesting questions which need to be attended to in order to fully understand LSD1 mechanistically and functionally. How do other proteins in the CoREST, ZEB1/CtBP complex and other histone modification complexes help target LSD1 to the appropriate locations in the chromatin? The complexes LSD1 has been shown to reside in, consist of a core complex with various polypetides being interchangeable. Many of the interchangeable polypeptides contain various domains known to interact with histones in modification dependent and independent fashions. Do these effector modules influence the activity of LSD1? Do some act as substrate presentation modules for LSD1? Do interactions with the androgen receptor, the estrogen receptor

or components of the MLL1 activator complex switch the specificity of the enzyme from H3K4 to H3K9? Do post-translational modifications switch LSD1 activity from H3K4 to H3K9? Is LSD1 allosterically regulated by a protein cofactor or a modification event? Our structural and mutational analysis of LSD1 opens the door to answering this next generation of questions.

References

1. Luger K: Structure and dynamic behavior of nucleosomes. *Curr. Opin. Genet. Dev.* (2003) **13**(2):127-35.

2. Schalch T, Duda S, Sargent DF, Richmond TJ: X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* (2005) **436**(7047):138-41.

3. Strahl BD, Allis CD: The language of covalent histone modifications. *Nature* (2000) **403**(6765):41-5.

4. Jenuwein T, Allis CD: Translating the histone code. *Science* (2001) **293**(5532):1074-80.

5. Spotswood HT, Turner BM: An increasingly complex code. *J. Clin. Invest*. (2002) **110**(5):577-82.

6. Turner BM.: Cellular memory and the histone code. *Cell* (2002) **111**(3):285-91.

7. Tamkun JW, Deuring R, Scott MP *et al:* brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell (*1992) **68**(3):561-72.

8. Haynes SR, Dollard C, Winston F *et al:*. The bromodomain: a conserved sequence found in human, Drosophila and yeast proteins. *Nucleic Acids Res.* (1992) **20**(10):2603.

9. Zeng L, Zhou MM. Bromodomain: an acetyl-lysine binding domain. *FEBS Lett.* (2002) **513**(1):124-8.

10. Dhalluin C, Carlson JE, Zeng L *et al:*. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* (1999) **399**(6735):491-6.

11. Jacobson RH, Ladurner AG, King DS, Tjian R: Structure and function of a human TAFII250 double bromodomain module. *Science* (2000) **288**(5470):1422-5.

12. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, ZhouMM. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399: 491–496.

13. Jacobson RH, Ladurner AG, King DS, Tjian R. (2000). Structure and function of a human TAFII250 double bromodomain module. *Science* 288: 1422–1425.

14. Paro R, Hogness DS: The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila. *Proc. Natl. Acad. Sci. U S A*. (1991) **88**(1):263-7.

15. Eissenberg JC: Molecular biology of the chromo domain: an ancient chromatin module comes of age. *Gene* (2001) **275**(1):19-29.

16. Bannister AJ, Zegerman P, Partridge JF *et al:*. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* (2001) **410**(6824):120-4.

17. Lachner M, O'CArroll D, Rea S *et al:*. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* (2001) **410**(6824):116-20.

18. Flanagan JF, Mi LZ, Chruszcz M *et al:*. Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature.*(2005) **438**(7071):1181-5.

19. Ball LJ, Murzina NV, Broadhurst RW, Raine AR, Archer SJ, Stott FJ, Murzin AG, Singh PB, Domaille PJ, Laue ED. Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. *EMBO J* 1997; 16:2473–481.

20. Brasher SV, Smith BO, Fogh RH, Nietlispach D, Thiru A, Nielsen PR, Broadhurst RW, Ball LJ, Murzina NV, Laue ED. The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *EMBO J* 2000;19:1587–1597.

21. Jacobs SA, Khorasanizadeh S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* 2002;295: 2080–2083.

22. Nielsen PR, Nietlispach D, Mott HR, Callaghan J, Bannister A, Kouzarides T, Murzin AG, Murzina NV, Laue ED. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* 2002; 416:103–107.

23. Huang Y, Fang J, Bedford MT *et al:*. Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* (2006) **312**(5774):748-51.

24. Kim J, Daniel J, Espejo A *et al:*. Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* (2006) **7**(4):397-403.

25. Cote J, Richard S.: Tudor domains bind symmetrical dimethylated arginines. *J. Biol. Chem.* (2005) **280**(31):28476-83.

26. Huyen Y, Zgheib O, Ditullio RA Jr, Gorgoulis VG, Zacharatos P, Petty TJ, Sheston EA, Mellert HS, Stavridi ES, Halazonetis TD. Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature*. 2004 Nov 18;432(7015):406-11

27. Wysocka J, Swigut T, Xiao H *et al:*. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* (2006) **442**(7098):86-90.

28. Doyon Y, Cayrou C, Ullah M *et al:*. ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Mol. Cell.* (2006) **21**(1):51-64.

29. Shi X, Hong T, Walter KL, Ewalt M *et al:*. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* (2006) **442**(7098):96-9.

30. Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*. 2006 Jul 6;442(7098):91-5

31. Aasland R, Stewart AF, Gibson T. The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. *Trends Biochem Sci.* (1996) **21**(3):87-8.

32. Boyer LA, Latek RR, Peterson CL: The SANT domain: a unique histone-tail-binding module? *Nat. Rev. Mol. Cell. Biol.* (2004) **5**(2):158-63.

33. Yu J, Li Y, Ishizuka T *et al:*. A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation. *EMBO J.* (2003) **22**(13):3403-10.

34. Grüne T, Brzeski J, Eberharter A, Clapier CR, Corona DF, Becker PB, Müller CW. Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. *Mol Cell*. 2003 Aug;12(2):449-60.

35. Zhang Y, Reinberg D: Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev*. (2001) **15**(18):2343-60.

36. Bannister AJ, Schneider R, Kouzarides T: Histone methylation: dynamic or static? *Cell* (2002) **109**(7):801-6.

37. Lachner M, O'Sullivan RJ, Jenuwein T: An epigenetic road map for histone lysine methylation. *J. Cell Sci*. (2003) **116**(Pt 11):2117-24.

38. Bauer UM, Daujat S, Nielsen SJ *et al:*. Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep.* (2002) **3**(1):39-44.

39. Chen D, Ma H, Hong H *et al:*. Regulation of transcription by a protein methyltransferase. *Science* (1999) **284**(5423):2174-7.

40. Strahl BD, Briggs SD, Brame CJ *et al:*. Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol*. (2001) **26**;11(12):996-1000.

41. Wang H, Huang ZQ, Xia L *et al:*. Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* (2001) **293**(5531):853-7.

42. Cheng X, Collins RE, Zhang X: Structural and sequence motifs of protein (histone) methylation enzymes. *Annu. Rev. Biophys. Biomol. Struct*. (2005) **34**:267-94.

43. Wang Y, Wysocka J, Sayegh J *et al:*. Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* (2004) **306**(5694):279-83.

44. Shi Y, Lan F, Matson C, Mulligan P *et al:*. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* (2004) **119**(7):941-53.

45. You A, Tong JK, Grozinger CM, Schreiber SL. CoREST is an integral component of the CoREST- human histone deacetylase complex. *Proc. Natl. Acad. Sci. U S A*. (2001) **98**(4):1454-8.

46. Humphrey GW, Wang Y, Russanova VR *et al:*. Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. *J. Biol. Chem.* (2001) **276**(9):6817-24.

47. Ballas N, Battaglioli E, ATOUF F *et al:*. Regulation of neuronal traits by a novel transcriptional complex. *Neuron* (2001) **31**(3):353-65.

48. Hakimi MA, Bochar DA, Chenoweth J *et al:*. A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc. Natl. Acad. Sci. U S A.* (2002) **99**(11):7420-5.

49. Shi Y, Sawada J, Sui G *et al:*. Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature* (2003) **422**(6933):735-8.

50. Forneris F, Binda C, Vanoni MA *et al:.* Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Lett*. (2005) (10):2203-

7.

51. Kubicek S, Jenuwein T: A crack in histone lysine methylation. *Cell* (2004) **19**(7):903-6.

52. Tsukada Y, Fang J, Erdjument-Bromage H *et al:*. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* (2006) **439**(7078):811-6.

53. Chen Z, Zang J, Whetstine J *et al:*. Structural insights into histone demethylation by JMJD2 family members. *Cell* (2006) **125**(4):691-702.

54. Hausinger RP: FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes.
Crit Rev Biochem Mol. Biol. (2004) **39**(1):21-68.

55. Klose RJ, Kallin EM, Zhang Y: JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* (2006) **7**(9):715-27.

56. Yamane K, Toumazou C, Tsukada Y *et al:*. JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* (2006) **125**(3):483-95.

57. Whetstine JR, Nottke A, Lan F *et al:*. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* (2006) **125**(3):467-81.

58. Fodor BD, Kubicek S, Yonezawa M *et al:*. Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes Dev.* (2006) **20**(12):1557-62.

59. Christensen J, Agger K, Cloos PA *et al:.* RBP2 Belongs to a Family of Demethylases, Specific for Tri-and Dimethylated Lysine 4 on Histone 3. *Cell*. (2007) **128**(6):1063-76.

60. Seward DJ, Cubberley G, Kim S *et al:.* Demethylation of trimethylated histone H3 Lys4 in vivo by JARID1 JmjC proteins. *Nat. Struct. Mol. Biol.* (2007) **14**(3):240-2.

61. Shi YJ, Matson C, Lan F *et al:*. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol. Cell*. (2005) **19**(6):857-64.

62. Lee MG, Wynder C, Cooch N, Shiekhattar R: An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* (2005) **437**(7057):432-5.

63. Stavropoulos P, Blobel G, Hoelz A. Crystal structure and mechanism of

human lysine-specific demethylase-1. *Nat. Struct. Mol. Biol.* (2006) **13**(7):626-32.

64. Ding Z, Gillespie LL, Paterno GD: Human MI-ER1 alpha and beta function as transcriptional repressors by recruitment of histone deacetylase 1 to their conserved ELM2 domain. *Mol. Cell. Biol.* (2003) **23**(1):250-8.

65. Wang L, Rajan H, Pitman JL, Mckeown M, Tsai CC: Histone deacetylase-associating Atrophin proteins are nuclear receptor corepressors. *Genes Dev.* (2006) **20**(5):525-30.

66. Boyer LA, Langer MR, Crowley KA *et al:*. Essential role for the SANT domain in the functioning of multiple chromatin remodeling enzymes. *Mol. Cell.* (2002) **10**(4):935-42.

67. Yang M, Gocke CB, Luo X *et al:*. Structural basis for CoREST-dependent demethylation of nucleosomes by the human LSD1 histone demethylase. *Mol. Cell.* (2006) **23**(3):377-87.

68. Lan F, Collins RE, De Cegli R, Alpatov R, Horton JR, Shi X, Gozani O, Cheng X, Shi Y. Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature*. 2007 Aug 9;448(7154):718-22.

69. Metzger E, Wissmann M, Yin N *et al:*. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* (2005) **437**(7057):436-9.

70. Perillo B, Sasso A, Abbondanza C, Palumbo G. 17beta-estradiol inhibits apoptosis in MCF-7 cells, inducing bcl-2 expression via two estrogen-responsive elements present in the coding sequence. *Mol Cell Biol.* 2000 Apr;20(8):2890-901.

71. Perillo B, Ombra MN, Bertoni A, Cuozzo C, Sacchetti S, Sasso A, Chiariotti L, Malorni A, Abbondanza C, Avvedimento EV. DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science.* 2008 Jan 11;319(5860):202-6.

72. Grollman AP, Moriya M. Mutagenesis by 8-oxoguanine: an enemy within. *Trends Genet.* 1993 Jul;9(7):246-9.

73. Roldán-Arjona T, Wei YF, Carter KC, Klungland A, Anselmino C, Wang RP, Augustus M, Lindahl T. Molecular cloning and functional expression of a human cDNA encoding the antimutator enzyme 8-hydroxyguanine-DNA glycosylase.
Proc Natl Acad Sci U S A. 1997 Jul 22;94(15):8016-20.

74. Kroeger PE, Osheroff N, Rowe TC. Analysis of topoisomerase II-mediated DNA cleavage in the 5'-region of the Drosophila hsp70 gene. Identification of a novel half-site DNA substrate for topoisomerase II cleavage. *J Biol Chem*. 1993 Aug 5;268(22):16449-57

75. Garcia-Bassets I, Kwon YS, Telese F, Prefontaine GG, Hutt KR, Cheng CS, Ju BG, Ohgi KA, Wang J, Escoubet-Lozach L, Rose DW, Glass CK, Fu XD, Rosenfeld MG. Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. *Cell*. 2007 Feb 9;128(3):505-18.

76. Scully KM, Rosenfeld MG. Pituitary development: regulatory codes in mammalian organogenesis. *Science*. 2002 Mar 22;295(5563):2231-5.

77. Wang J, Scully K, Zhu X, Cai L, Zhang J, Prefontaine GG, Krones A, Ohgi KA, Zhu P, Garcia-Bassets I, Liu F, Taylor H, Lozach J, Jayes FL, Korach KS, Glass CK, Fu XD, Rosenfeld MG. Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature.* 2007 Apr 19;446(7138):882-7.

78. Olson LE, Tollkuhn J, Scafoglio C, Krones A, Zhang J, Ohgi KA, Wu W, Taketo MM, Kemler R, Grosschedl R, Rose D, Li X, Rosenfeld MG. Homeodomain-mediated beta-catenin-dependent switching events dictate cell-lineage determination. *Cell.* 2006 May 5;125(3):593-605.

79. Maier MM, Gessler M. Comparative analysis of the human and mouse Hey1 promoter: Hey genes are new Notch target genes. *Biochem Biophys Res Commun.* 2000 Aug 28;275(2):652-60.

80. H. Hock, M.J. Hamblen, H.M. Rooke, J.W. Schindler, S. Saleque, Y. Fujiwara and S.H. Orkin, Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells, *Nature* **431** (2004), pp. 1002–1007.

81. Wang, K. Scully, X. Zhu, L. Cai, J. Zhang, G.G. Prefontaine, A. Krones, K.A. Ohgi, P. Zhu and I. Garcia-Bassets *et al.*, Opposing LSD1 complexes function in developmental gene activation and repression programmes, *Nature* **446** (2007), pp. 882–887

82. H. Hock, M.J. Hamblen, H.M. Rooke, D. Traver, R.T. Bronson, S. Cameron and S.H. Orkin, Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation, *Immunity* **18** (2003), pp. 109–120

83. H. Karsunky, H. Zeng, T. Schmidt, B. Zevnik, R. Kluge, K.W. Schmid, U. Duhrsen and T. Moroy, Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1, *Nat. Genet.* **30** (2002), pp.

295–300.

84. Fiolka, R. Hertzano, L. Vassen, H. Zeng, O. Hermesh, K.B. Avraham, U. Duhrsen and T. Moroy, Gfi1 and Gfi1b act equivalently in haematopoiesis, but have distinct, non-overlapping functions in inner ear development, *EMBO Rep.* **7** (2006), pp. 326–333.

85. H.L. Grimes, T.O. Chan, P.A. Zweidler-McKay, B. Tong and P.N. Tsichlis, The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal, *Mol. Cell. Biol.* **16** (1996), pp. 6263–6272.

86. J. Zhu, L. Guo, B. Min, C.J. Watson, J. Hu-Li, H.A. Young, P.N. Tsichlis and W.E. Paul, Growth factor independent-1 induced by IL-4 regulates Th2 cell proliferation, *Immunity* **16** (2002), pp. 733–744

87. Hock and S.H. Orkin, Zinc-finger transcription factor Gfi-1: versatile regulator of lymphocytes, neutrophils and hematopoietic stem cells, *Curr. Opin. Hematol.* **13** (2006), pp. 1–6.

88. Saleque S, Kim J, Rooke HM, Orkin SH. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. *Mol Cell*. 2007 Aug 17;27(4):562-72.

89. Lan F, Zaratiegui M, Villén J, Vaughn MW, Verdel A, Huarte M, Shi Y, Gygi SP, Moazed D, Martienssen RA, Shi Y. S. pombe LSD1 homologs regulate heterochromatin propagation and euchromatic gene transcription. *Mol Cell.* 2007 Apr 13;26(1):89-101.

90. J.-I. Nakayama, J.C. Rice, B.D. Strahl, C.D. Allis and S.I.S. Grewal, Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly, *Science* **292** (2001), pp. 110–113.

91. G.D. Shankaranarayana, M.R. Motamedi, D. Moazed and S.I. Grewal, Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast, *Curr. Biol.* **13** (2003), pp. 1240–1246.

92. Rudolph T, Yonezawa M, Lein S, Heidrich K, Kubicek S, Schäfer C, Phalke S, Walther M, Schmidt A, Jenuwein T, Reuter G. Heterochromatin formation in Drosophila is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. *Mol Cell.* 2007 Apr 13;26(1):103-15.

93. K. Nishioka, S. Chuikov, K. Sarma, H. Erdjument-Bromage, C.D. Allis, P. Tempst and D. Reinberg, Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation, *Genes Dev.* **16** (2002), pp. 479–489

94. Huang J, Perez-Burgos L, Placek BJ, Sengupta R, Richter M, Dorsey JA, Kubicek S, Opravil S, Jenuwein T, Berger SL. Repression of p53 activity by Smyd2-mediated methylation. *Nature*. 2006 Nov 30;444(7119):629-32.

95. Huang J, Sengupta R, Espejo AB, Lee MG, Dorsey JA, Richter M, Opravil S, Shiekhattar R, Bedford MT, Jenuwein T, Berger SL. p53 is regulated by the lysine demethylase LSD1. *Nature.* 2007 Sep 6;449(7158):105-8.

96. Iwabuchi, K., Bartel, P. L., Li, B., Marraccino, R. & Fields, S. Two cellular proteins that bind to wild-type but not mutant p53. *Proc. Natl Acad. Sci. USA* 91, 6098–6102 (1994)

97. Sengupta S, Robles AI, Linke SP, Sinogeeva NI, Zhang R, Pedeux R, Ward IM, Celeste A, Nussenzweig A, Chen J, Halazonetis TD, Harris CC. Functional interaction between BLM helicase and 53BP1 in a Chk1-mediated pathway during S-phase arrest. *J Cell Biol*. 2004 Sep 13;166(6):801-13.

98. Ward I, Kim JE, Minn K, Chini CC, Mer G, Chen J. The tandem BRCT domain of 53BP1 is not required for its repair function. *J Biol Chem.* 2006 Dec 15;281(50):38472-7

99. Iwabuchi K, Li B, Massa HF, Trask BJ, Date T, Fields S. Stimulation of p53-mediated transcriptional activation by the p53-binding proteins, 53BP1 and 53BP2. *J Biol Chem*. 1998 Oct 2;273(40):26061-8.

100. Lee MG, Wynder C, Bochar DA *et al:*. Functional interplay between histone demethylase and deacetylase enzymes. *Mol. Cell. Biol*. (2006) **26**(17):6395-402.

101. Binda C, Mattevi A, Edmondson DE: Structure-function relationships in flavoenzyme-dependent amine oxidations: a comparison of polyamine oxidase and monoamine oxidase. *J. Biol. Chem.* (2002) **277**(27):23973-6.

102. Youdim MB, Edmondson D, Tipton KF: The therapeutic potential of monoamine oxidase inhibitors. *Nat. Rev. Neurosci.* (2006) **7**(4):295-309.

103. Lee MG, Wynder C, Schmidt DM *et al:*. Histone H3 lysine 4 demethylation is a target of nonselective antidepressive medications. *Chem. Biol.* (2006) **13**(6):563-7.

104. Culhane JC, Szewczuk LM, Liu X *et al:*. A mechanism-based inactivator for histone demethylase LSD1. *J. Am. Chem. Soc.* (2006) **128**(14):4536-7.

105. Doublié S. Preparation of selenomethionyl proteins for phase determination.

Methods Enzymol. 1997;276:523-30.

106. Otwinowski, Z. & Minor, W. Processing of X-Ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326 (1997).

107. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763 (1994).

108. Hoelz, A., Nairn, A.C. & Kuriyan, J. Crystal structure of a tetradecameric assembly of the association domain of Ca2+/calmodulin-dependent kinase II. *Mol. Cell* **11**, 1241-1251 (2003).

109. Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O. & Darst, S.A. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285-1290 (2002)

110. La Fortelle, E.D. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. *Methods Enzymol.* **276**, 476−494 (1997).

111. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crysallogr. Sect. A Found. Crystallogr.* **47**, 110-119 (1991).

112. Brünger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905-921 (1998).

113. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291 (1993).

114. DeLano, W.L. (2002). The PyMOL User's Manual. Delano Scientific, San Carlos, CA.

115. Sanner, M.F., Olson, A.J. & Spehner, J.C. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers* **38**, 305-320 (1996).

116. Jeanmougin, F., Thompson, J.D., Guoy, M., Higgins, D.G. & Gibson, T.J. Multiple sequence alignment with Clustal X. *Trends Biochem.* **23**, 403-405 (1998).

117. Aravind, L. & Iyer, L.M. The SWIRM domain: a conserved module found in chromosomal proteins points to novel chromatin-modifying activities. *Genome Biol.* **3**, RESEARCH0039.

118. Qian C, Zhang Q, Li S, Zeng L, Walsh MJ, Zhou MM. Structure and chromosomal DNA binding of the SWIRM domain. *Nat Struct Mol Biol*. 2005 Dec;12(12):1078-85.

119. Da G, Lenkart J, Zhao K, Shiekhattar R, Cairns BR, Marmorstein R. Structure and function of the SWIRM domain, a conserved protein module found in chromatin regulatory complexes. *Proc Natl Acad Sci U S A*. 2006 Feb 14;103(7):2057-62.

120. Binda C, Angelini R, Federico R, Ascenzi P, Mattevi A. Structural bases for inhibitor binding and catalysis in polyamine oxidase. *Biochemistry.* 2001 Mar 6;40(9):2766-76.

121. Yang M, Culhane JC, Szewczuk LM, Gocke CB, Brautigam CA, Tomchick DR, Machius M, Cole PA, Yu H. Structural basis of histone demethylation by LSD1 revealed by suicide inactivation. *Nat Struct Mol Biol*. 2007 Jun;14(6):535-9.

122. Culhane JC, Cole PA. LSD1 and the chemistry of histone demethylation. *Curr Opin Chem Biol*. 2007 Oct;11(5):561-8.

123. Bannister, A.J. & Kouzarides, T. Reversing histone methylation. *Nature* **436**, 1103-1106

124. Fraaije, M.W. & Mattevi. A. Flavoenzymes: diverse catalysts with recurrent features. *Trends. Biochem. Sci.* **25**, 126-132 (2000).

125. Forneris F, Binda C, Vanoni MA, Mattevi A, Battaglioli E. Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process. *FEBS Lett*. 2005 Apr 11;579(10):2203-7

126. Forneris F, Binda C, Vanoni MA, Battaglioli E, Mattevi A. Human histone demethylase LSD1 reads the histone code. *J Biol Chem*. 2005 Dec 16;280(50):41360-5