npg

ORIGINAL ARTICLE

Structure-function analysis reveals a novel mechanism for regulation of histone demethylase LSD2/AOF1/KDM1b

Qi Zhang[1, *], Shankang Qi[2, *], Mingchu Xu[2], Lin Yu[1], Ye Tao[3], Zengqin Deng[1], Wei Wu[1], Jiwen Li[2], Zhongzhou Chen[1], Jiemin Wong[2]

[1]State Key Laboratory of Agrobiotechnology, China Agricultural University, Beijing 100193, China; [2]Shanghai Key Laboratory of Regulatory Biology, The Institute of Biomedical Sciences and School of Life Sciences, East China Normal University, Shanghai 200241, China; [3]Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences 100049, Beijing, China

LSD2/AOF1/KDM1b catalyzes demethylation of mono- and di-methylated H3K4 and plays an important role in transcriptional regulation and genomic imprinting. Here, we report the high-resolution crystal structures of apo-LSD2 and LSD2 in complex with a peptide that mimics H3K4me2. Three structural domains of LSD2, namely, the novel N-terminal zinc finger, the centrally located SWIRM domain, and the C-terminal oxidase domain, closely pack together to form a boot-shaped structure. The active site cavity in the oxidase domain is large enough to accommodate several residues of the histone H3 tail and cannot discriminate between the different states of H3K4 methylation. The N-terminal zinc-finger domain, composed of a novel C4H2C2-type zinc finger and a specific CW-type zinc finger, is required for demethylase activity and, surprisingly, the binding of cofactor flavin adenine dinucleotide (FAD). In fact, a relay of extensive interactions through the zinc finger-SWIRM-oxidase domains is required for LSD2 demethylase activity and the binding of FAD. These results reveal a novel mechanism for the zinc finger and SWIRM domains in controlling LSD2 demethylase activity and provide a framework for elucidating the regulation and function of LSD2.

Keywords: LSD2; complex structures; C4H2C2 zinc finger; CW zinc finger; SWIRM; FAD; amine oxidase

Introduction

Methylation of lysine (K) residues at the ε-amino group is one of the most common post-translational modifications of histone N-terminal tails [1, 2]. Lysine residues can be mono-, di- and tri-methylated. Histone lysine methylation plays critical roles in virtually all chromatin-templated biological processes, such as transcription and DNA repair [2, 4]. Cumulative evidence indicates that histone lysine methylation exerts diverse biological functions in a site- and methylation state-dependent manner [2-4]. Like other histone modifications,

histone lysine methylation is reversible and dynamically regulated through the balancing activities of histone lysine methyltransferases and demethylases [5-7]. To date, two families of proteins have been found to possess lysine demethylase activity. Members of one family are amine oxidases that can catalyze demethylation of mono- and di-methylated lysine residues using flavin adenine dinucleotide (FAD) as a cofactor [8]. Members of the other family are the JmjC family proteins that catalyze oxidative demethylation reactions with iron and α-ketoglutarate as cofactors [9-11]. As histone methyltransferases and demethylases are often deregulated in human diseases, it is essential to understand their physiological and pathological functions by elucidating their precise structure, regulation and substrate specificity.

The FAD-dependent demethylase family includes the first identified histone lysine demethylase LSD1/KDM1α [8] and its homolog LSD2 (also known as AOF1/KDM1b). Although LSD1 has been extensively studied [8,

12], the enzymatic activity and biological function of LSD2 have just begun to emerge. Several studies [13-15] have demonstrated that LSD2 specifically demethylates mono- and di-methylated H3K4 (H3K4me1/2). In one study [16], LSD2 was also shown to demethylate di-methylated H3K9 (H3K9me2). Notably, demethylation of H3K4 by LSD2 requires the presence of a long stretch of H3 tail sequence and is strongly influenced by other modifications of H3 tails [13]. Biochemical studies indicate that LSD2 associates with transcriptional elongation factors and phosphorylated RNA polymerase II [14, 17]. LSD2 is consistently enriched in the coding regions of actively transcribed genes, and may play a role in maintaining a repressive chromatin environment within gene bodies through its H3K4me1/2 demethylase activity [14]. Notably, a genetic study has linked H3K4 demethylation by LSD2 to the establishment of maternal imprinting in mouse oocytes [18]. Like LSD1, LSD2 contains a SWIRM domain and a C-terminal amine oxidase domain (AOD). However, LSD2 lacks a tower domain, a unique insertion present in the AOD of LSD1 that is required for LSD1 binding of CoREST and LSD1's enzymatic activity [19-21]. Furthermore, LSD2 has a unique N-terminal region that is absent in LSD1, which is predicted to form a CW-type zinc-finger domain [15]. The CW domain is present in several chromatin-remodeling proteins and has been shown to bind methylated histones [22, 23]. Deletion mutant experiments have demonstrated that both the unique N-terminal zinc-finger domain and the SWIRM domain are required for LSD2 demethylase activity [15], but the underlying mechanism remains unknown.

Here, we report the crystal structures and functional analyses of LSD2. Like LSD1, the active-site cavity in the AOD of LSD2 is large enough to accommodate several residues of histone H3 tail substrate and cannot distinguish the different states of H3K4 methylation. In the structure, the N-terminal zinc finger, middle SWIRM and C-terminal oxidase domains of LSD2 were found to be closely packed together. Surprisingly, we observed that a relay of extensive interactions through the zinc finger-SWIRM-AOD is required for the proper conformation of LSD2 and, consequently, its demethylase activity and binding of FAD. As zinc fingers are structural modules frequently present in histone methyltransferases and demethylases, our study may also provide an explanation for the commonly observed co-evolution of zinc fingers with histone methyltransferases and demethylases.

Results

Structure determination of apo-LSD2

We prepared the full-length human LSD2 (Figure 1A) and a series of truncated fragments by overexpression and purification from *Escherichia coli*. After extensive trials, only one fragment of LSD2 (residues 30-822) successfully formed crystals. This recombinant protein contained all three known structural domains of LSD2, namely, the zinc finger, SWIRM and AOD (Figure 1A). Using histones as substrate, this recombinant LSD2 exhibited a robust demethylase activity toward H3K4me2 but not H3K9me2 (Figure 1B). A similar result was obtained in the assay to measure the demethylase activity of the full length LSD2 (Supplementary information, Figure S1). Furthermore, isothermal titration calorimetry (ITC) experiments were performed to determine the binding affinity of LSD2 (residues 30-822) to a synthetic peptide of the H3 tail (residues 1-21). This peptide (referred to as H3K4M) mimics H3K4me2, but cannot be demethylated by LSD2 as it contains a methionine substitution at K4. The binding affinity (K_d) of LSD2 to H3K4M was determined to be approximately 5.5 μM (Figure 1C). Thus, the truncated LSD2 (residues 30-822) is functionally active. For simplicity, we hereafter refer to the structure of LSD2 (residues 30-822) as LSD2 unless otherwise specified.

We solved the structures of the SWIRM domain and the AOD by OASIS [24], using molecular replacement methods with the atomic coordinates of LSD1 (Protein Data Bank (PDB) code 2HKO [25]) as the initial search model. The number and the position of three zinc atoms were determined with anomalous data collected from the zinc adsorption edge at 1.28 Å (Table 1) by the SHELX C/D program [26]. The structure of the zinc fingers was built manually in the program COOT [27]. The final model was refined to 2.13 Å with an R_{work} of 21.2% and an R_{free} of 24.9%.

Overall structure of apo-LSD2 and LSD2 in complex with H3K4M

The apo-LSD2 structure is organized into three distinct domains that pack closely against each other with the SWIRM domain in the middle (Figure 1D and 1E). Overall, apo-LSD2 exhibits a densely packed boot-shaped structure, measuring 90 Å in height, 75 Å in length and 60 Å in width. The structure consists of 33 α-helices and 28 β-sheets (Supplementary information, Figure S2). The zinc-finger domain containing three zinc atoms is located at the N-terminus (Figure 1D). The adjacent domain is the ~100-residue SWIRM domain, which adopts a completely helical histone fold as in LSD1 [28]. The AOD adopts a common amine oxidase fold that is conserved within the family of flavoenzymes. The overall structure only lacks a 22-residue disordered loop (residues 238-259) that links the zinc finger and SWIRM
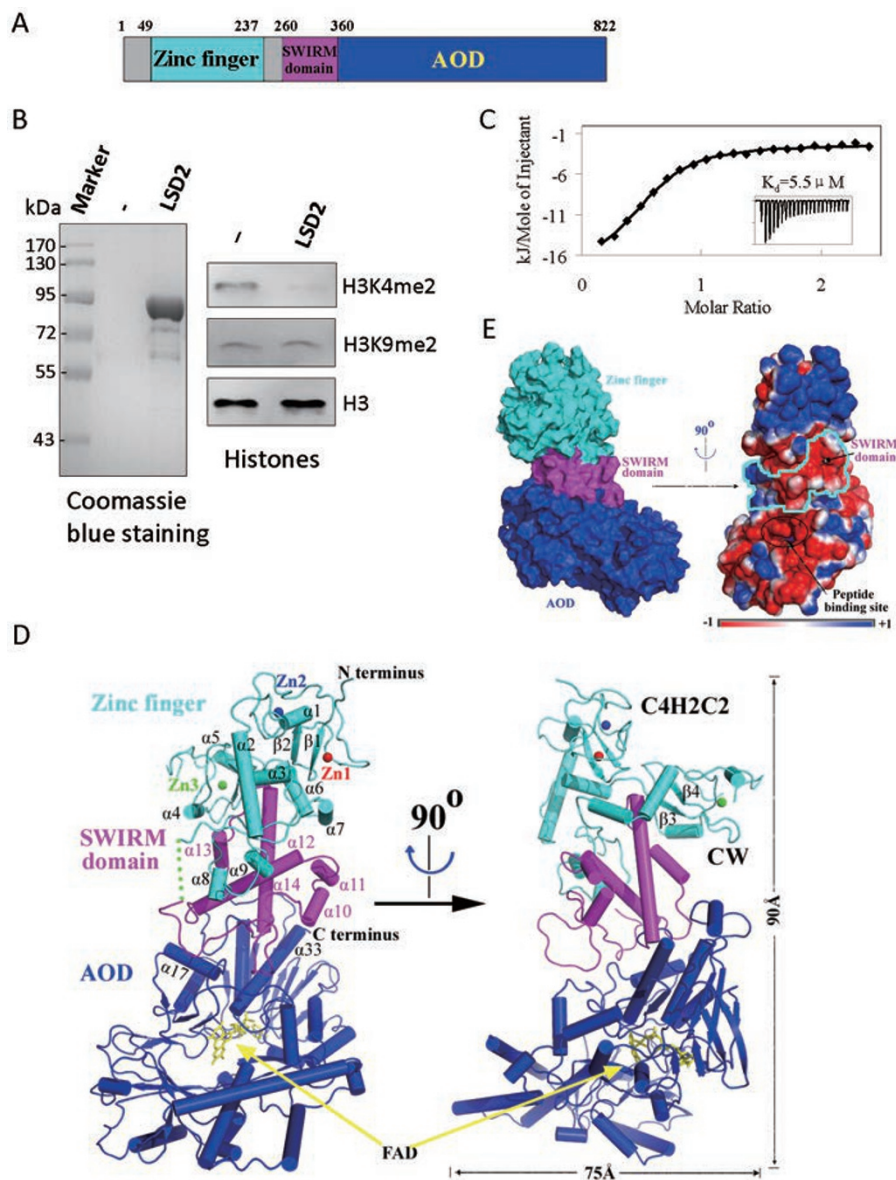
Figure 1 Structure determination of human LSD2. **(A)** Domain organization of LSD2. Color coding: gray, unstructured region; cyan, zinc-finger domain; magenta, SWIRM domain; blue, amine oxidase domain (AOD). **(B)** Coomassie blue staining of LSD2 from *E. coli* (left) and *in vitro* demethylation assay (right). Bulk histones were incubated with or without LSD2 (residues 30-822) for demethylation reactions and analyzed by western blot using antibodies against H3K4me2 and H3K9me2 as indicated. **(C)** Quantification of the binding affinity of LSD2 to H3K4M by ITC. The integrated heat is plotted against the molar ratio of peptide added to LSD2 in the cell. The dissociation constant (K_d) is indicated. **(D)** Structure of LSD2 in tube representation, colored as in **A**. Spheres representing three zinc atoms Zn1, Zn2 and Zn3 are colored red, blue and green, respectively. FAD is shown in yellow stick representation. Dimensions are indicated. A 90°-rotated view is shown on the right. **(E)** Surface view of LSD2 (left) and a 90°-rotated electrostatic surface representation (right). The SWIRM domain and peptide-binding site were indicated by cyan-colored dotted line and black ellipse, respectively. All the structural figures were made using PYMOL [41].

domains.

We also obtained the crystal structure of LSD2 in complex with H3K4M. This complex was diffracted to 2.7 Å in the space group C2 and was solved by molecular re-

placement (Table 1). An electron density corresponding to the residues 1-13 of H3K4M appeared in the active center of the LSD2 molecule at a position similar to that of the LSD1-CoREST-H3K4M complex (Figure 2A) [29].

Table 1 Data collection and refinement statistics of LSD2[*]

Crystal	Peak-LSD2	Native-LSD2	LSD2-H3K4M Complex
Space group	I222	C222₁	C2
wavelength, Å	1.2800	0.9794	0.9792
Resolution, Å	50.0-2.90 (2.95-2.90)[†]	50.0-2.13 (2.17-2.13)[†]	50.0-2.70 (2.74-2.70)[†]
Cell dimensions			
a, b, c, Å	144.3, 170.9, 202.5	93.7, 96.3, 82.8	110.4, 98.5, 101.4
α, β, γ, °	90, 90, 90	90, 90, 90	90, 96.1, 90
Rmerge, %	10.1	8.3	7.2
$I/\sigma I$	12.2 (2.4)[†]	44.8 (2.2)	22.8 (2.3)
Completeness, %	97.6 (94.4)	93.9 (85.6)	98.8 (99.9)
Redundancy	3.3 (3.0)	10.2 (6.6)	4.5 (4.6)
Refinement			
Resolution, Å	50.0-2.90	50.0-2.13	50.0-2.70
No. of reflections	51559 (3642)	41665 (2729)	26437 (1405)
R_{work}/R_{free}, %	18.3/23.0 (29.4/36.0)	20.3/24.6 (30.3/35.5)	20.6/23.7 (25.9/30.5)
No. of atoms			
Protein	11513	5850	5737
Ligand/ion	128	69	56
Water	488	236	33
B-factors			
Protein	44.86	50.28	43.12
Ligand/ion	36.72	40.21	18.24
Water	34.69	52.02	27.96
rms deviations			
Bond lengths, Å	0.009	0.008	0.008
Bond angles, °	1.414	1.359	1.228
Ramachandran plot, %			
Most favored	94.0	96.8	95.8
Additional allowed	5.6	3.2	4.2
Generously allowed	0.4	0.0	0.0
Disallowed	0.0	0.0	0.0

[†]Statistics for highest resolution shell.

[*]Three crystal experiments for each structure.

H3K4M interacts with the acidic pocket of LSD2, with the N-terminus of the H3K4M deeply inserted. It is noteworthy that the structure of the LSD2-H3K4M complex is almost identical to that of apo-LSD2 within the AOD, with a root-mean-square deviation (RMSD) of 0.9 Å between their C-α atoms.

The amine oxidase catalytic domain and the binding of H3 substrate

A comparison of the LSD2 structure with the previously determined LSD1 structure [29] is shown in Figure 2B. In addition to the unique zinc-finger domain (in LSD2) and tower domain (in LSD1), we observed various differences in the SWIRM domain and the AOD between LSD1 and LSD2, which is consistent with a low level of sequence identity within these domains (24% and 33%, respectively). Within the AOD [19, 25, 28], significant local structural differences between LSD1 and LSD2 can be observed, especially in the solvent-exposed surface loops (Figure 2C). For example, most of the loops in LSD2 are shorter and exhibit a different conformation. These differences result in a relatively large RMSD of 2.0 Å over all 462 aligned Cα atoms between LSD2 and LSD1. These local structural differences indicate that the AOD of LSD2 may interact with different proteins and is likely subjected to different mechanisms
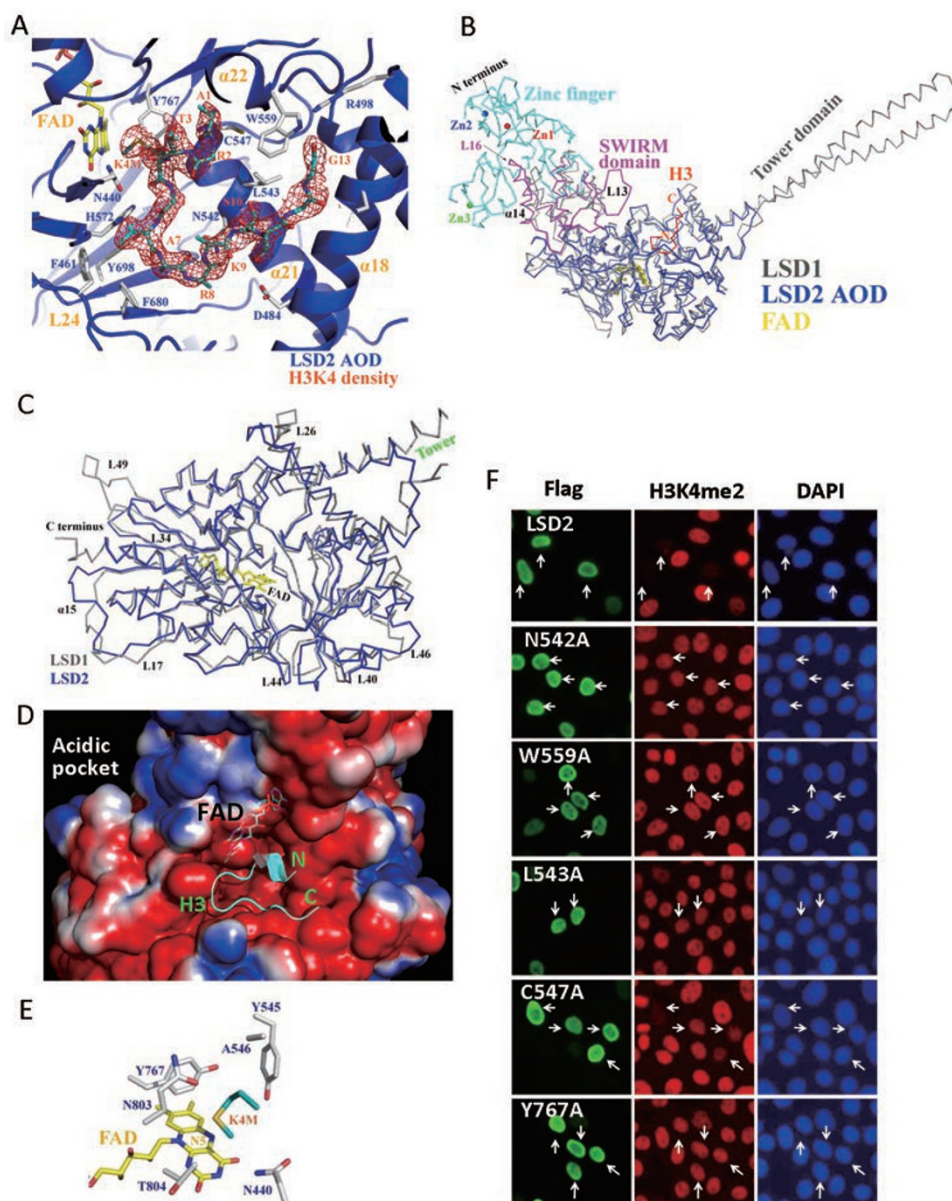
Figure 2 The structure of AOD with the binding of H3K4M. **(A)** $2F_o − F_c$ electron density map of the bound H3K4M at 1.0σ in the LSD2-substrate complex (FAD, yellow; LSD2 main chain, blue; LSD2 side chain, white; H3K4M, cyan). **(B)** Superimposition of the LSD2 structure on LSD1 (PDB 2V1D, grey). Domains of LSD2 are colored as in **1A.** H3K4M peptide is in orange. **(C)** Superimposition of LSD2 AOD on LSD1 AOD. The colors of LSD1, LSD2 and FAD are the same as in **B**. **(D)** H3K4M interacts with the acidic pocket of LSD2. The semi-transparent surface of AOD is colored based on electrostatic potential: red, negative; blue, positive. The N-terminus of H3K4M deeply inserts into the pocket. **(E)** The position of K4M residue in the active site of LSD2. The methyl group of K4M falls exactly in front of the N-5 atom of the flavin. **(F)** Mutations that disrupt the binding of H3 substrate or FAD result in loss of demethylase activity. Flag-tagged wild-type and mutant LSD2 were transfected into HeLa cells and the demethylase activity was determined by immunofluorescence using an anti-H3K4me2 antibody. Arrows mark the cells in which the proteins of interest were overexpressed.

of regulation.

Despite a substantial difference in surface area, the core structure of the AOD is highly conserved between LSD2 and LSD1 (Figure 2C). The LSD2 AOD can be divided into two subdomains: one is required for non-covalent binding of FAD, and the other is responsible for substrate binding and recognition (Figure 2C and 2D). The two subdomains create a large cavity at their inter-

face, which is ~15 Å-deep and ~20 Å-long and forms the catalytic center of LSD2 (Figure 2D). The large size of the catalytic cavity allows the insertion of a long stretch of H3 tail (Figure 2D and 2E), which corresponds to the first 13 residues of H3K4M (Figure 2A). Compared with the apo-LSD2, a few minor structural changes were observed in the AOD of the LSD2-H3K4M complex. For example, the electron density of the side chain of N440 becomes clear in the complex structure (Figure 2A). In addition, loop L24 moves 0.6 Å closer to the active center to enhance the binding of H3K4M, while α18 moves 0.8 Å away to extend the binding pocket. The large size of the LSD2 catalytic cavity also indicates that, like LSD1, the inability of LSD2 to demethylate H3K4me3 is not due to a structural incompatibility, but rather due to its reaction mechanism.

Based on the above complex structure, a number of conserved residues in the AOD of LSD2 are involved in the specific binding of the H3 tail and sensing the post-translational modification of the H3 tail peptide. For example, N542 forms a hydrogen bond with the backbone carbonyl oxygen of A7 in the H3K4M, whereas the side chain of W559 binds R2 in the H3K4M through a cation-π interaction. Indeed, mutations of the aforementioned two residues resulted in the loss of LSD2 demethylase activity, as revealed by immunofluorescent staining in HeLa cells using an H3K4me2-specific antibody (Figure 2F). In addition, from the structure it is clear that L543 and C547 form a hydrophobic interaction with the side chain of W559, which locks the conformation of the side chain of W559. Consistent with this observation, we found that both L543A and C547A mutants exhibit substantially decreased demethylase activity (Figure 2F). Furthermore, from the structure it is clear that Y767 is involved in binding and orienting FAD toward the catalytic center (Figure 2A) and, as expected, the Y767A mutant is enzymatically inactive (Figure 2F).

A previous study has [13] shown that both H3R2 methylation and H3T3 phosphorylation completely block H3K4 demethylation by LSD2. The requirement for R2 can be structurally explained as R2 is a key residue for substrate recognition, forming a cation-π interaction with W559 and an electrostatic interaction with E563 and D560. Moreover, two hydrogen bonds between the side chain of R2 and the amide nitrogens of G13 and G12 in H3K4 were identified, allowing the H3 tail peptide to bend to fit into the pocket. R2 methylation would disrupt these interactions, and thus impair the binding of the substrate H3 tail peptide. On the other hand, T3 phosphorylation would be expected to interrupt the hydrogen bond between N562 and T3 and weaken the enzyme-substrate interaction.

The LSD2 SWIRM domain

The SWIRM domain [30] is predicted to be an α-helical domain that promotes protein-protein interactions in the context of chromatin-related protein complexes. Indeed, like the previously determined SWIRM domain structures, LSD2 SWIRM consists of a bundle of helices (Figure 3A). Superimposition of the SWIRM domains of LSD1 and LSD2 reveals conserved characteristics of the overall fold (Figure 3A). However, as the SWIRM domains of LSD1 and LSD2 share only 24% sequence identity, the two differ by a relatively large RMSD of 2.0 Å, as calculated by the Dali server [31] (Figure 3A).

Similar to LSD1, there are extensive hydrophobic and hydrogen-bond interactions between the SWIRM domain and the AOD of LSD2. The interface between the SWIRM domain and the AOD buries a surface area of 1354 Å^2, mostly involving L13, L15, α12 and α14 of the SWIRM domain (Figure 3A) and L17, L30, α17 and α33 of the AOD domain (Figure 3B). The key residues in the SWIRM domain, L315, W318, K323, L326 and R358, are conserved between LSD1 and LSD2 (Supplementary information, Figure S2); however, the SWIRM-AOD interaction in LSD2 differs from that of LSD1 in several aspects. Compared with LSD1, the AOD of LSD2 lacks a C-terminal helix that interacts with the SWIRM domain through hydrophobic and Coulombic interactions. However, the LSD2 AOD gains two additional hydrogen-bond interactions via the side chains between E452 and N266, and Y268 and D571, and gains one electrostatic interaction between the positively charged side chain of K323 and negatively charged side chain of E452.

Notably, the SWIRM domain also closely packs against the N-terminal zinc-finger domain (Figure 3C). The hydrophobic interface between the zinc finger and the SWIRM domains buries a surface area of 1872 Å and involves an extensive network of van der Waals contacts derived from loops L13, L14 and L16, and helices α11, α12, α13 and α14 of the SWIRM domain, and L2, α3, L5, β3, β4, L7, α4, α6, α7, L10, α8, L11, α9 and L12 of the zinc-finger domain. Furthermore, several hydrogen bonds form between the zinc-finger and the SWIRM domains. These bonds involve the side chain of R338 and four carboxyl oxygen atoms of D133, Q134, Q135 and L136 in α3, as well as the side chain of Q158 in the zinc-finger domain; in addition, the carboxyl oxygen atom of W139 and the side chain of R342, and the amino group of L340 and side chain of E89 form hydrogen bonds. Moreover, the negatively charged side chain of Y138 attracts the positively charged side chain of R338 through electrostatic interactions. The side chain of Y202 in the zinc-finger domain inserts into the non-polar pocket in
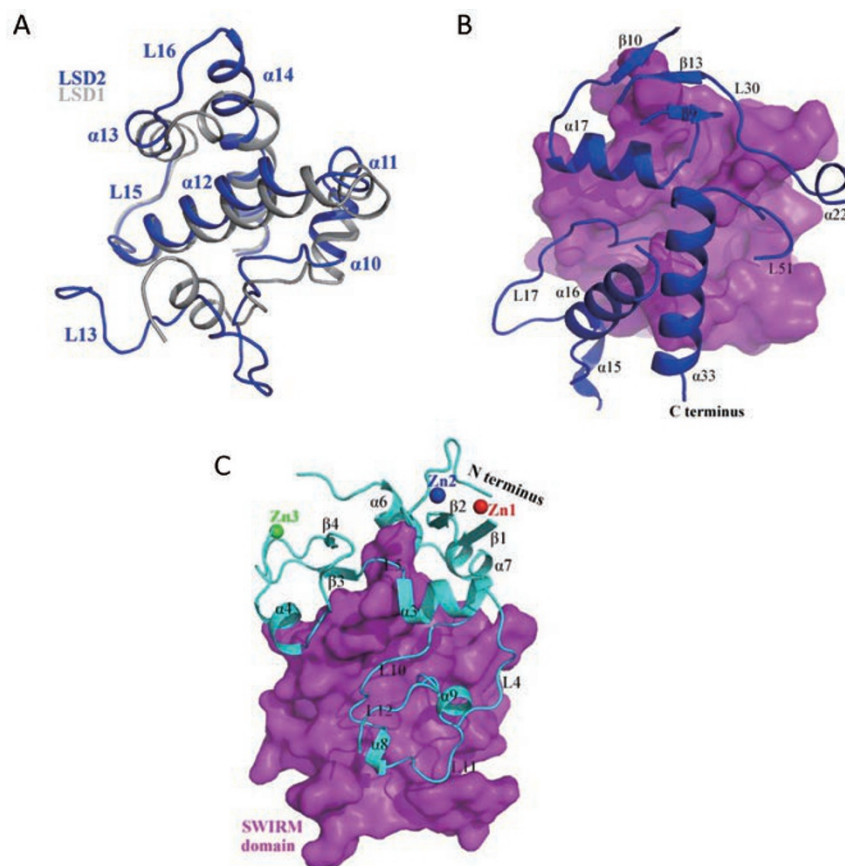
Figure 3 The SWIRM domain interacts extensively with the AOD and the N-terminal zinc-finger domain. **(A)** Superimposition of LSD1 and LSD2 SWIRM domains. **(B**, **C)** The SWIRM domain (shown in magenta surface) forms an extensive interface with the AOD **(B)** and the zinc-finger domain **(C)**, colors are the same as in **Figure 2B**.

the SWIRM domain and forms a hydrophobic network as well as a cation-π interaction with R344. It is noteworthy that the key interface residues L313, R338, R342 and R344 in the SWIRM domain are substituted by non-conserved residues in LSD1 (Supplementary information, Figure S2), indicating that the interaction of the SWIRM domain with the zinc-finger domain of LSD2 evolved independently.

The zinc-finger domain contains a novel conserved C4H2C2-type zinc finger and a specific CW-type zinc finger

The N-terminal zinc-finger domain is unique to LSD2 and is absent in LSD1. As shown in Figure 4A, the N-terminal zinc-finger domain actually consists of two individual zinc fingers and two linkers: an N-terminal C4H2C2-type zinc finger (residues 48-106, magenta), a linker region (Linker 1) composed of α2 and α3 (residues 107-136, light green), a CW-type zinc finger (residues 137-190, yellow), and the C-terminal linker region

(Linker 2) containing four helices and several loops (residues 191-237, pink). These classifications are similar to those calculated by the DIAL server [32]. The existence of three zinc ions is consistent with the 1:3 protein to zinc molar ratio determined by mass spectrometry [13].

The first zinc-finger motif has a unique $Cys2_A$-$Cys2_B$-$His2_A$-$Cys2_B$ pattern in a cross-brace topology (where applicable, A refers to residues that coordinate the first zinc ion, B to the second). In this C4H2C2 motif, residues C53, C58, H84 and H90 engage the first zinc ion (Zn1), whereas residues C65, C73, C92 and C95 bind the second zinc ion (Zn2) (Figure 4B); residues C62 and C87 are not involved in zinc binding and are not conserved (Supplementary information, Figure S4). The pattern of this zinc finger is distinct from the three zinc-finger motifs of similar size that exist widely in nuclear regulatory proteins: the $Cys2_A$-$Cys2_B$-$HisCys_A$-$Cys2_B$ PHD domain, the $Cys2_A$-$CysHis_B$-[His/Cys]Cys_A-$Cys2_B$ RING finger, and the $Cys2_A$-[His/Cys]2_A-$Cys2_B$-$Cys[His/Cys/Asp]_B$ LIM domain. A search of the protein data bank using DA-

Figure 4 Structural features of the N-terminal zinc-finger domain. **(A)** The overall structure of the zinc-finger domain. C4H2C2 zinc finger, magenta; Linker 1, light green; CW zinc finger, yellow; Linker 2, pink. **(B)** Close-up views of the C4H2C2 zinc finger and the positively charged surface. **(C)** Sequence alignment of several CW domains. Secondary structure elements of LSD2 are shown. The conserved zinc-binding residues are marked by red stars and the strictly conserved aromatic residues are marked by black circles. The residues which have potential to be involved in binding H3 peptide are indicated by blue circles. **(D)** The structure of CW zinc finger. A hydrophobic zone is shown as pink ellipse and key hydrogen bonds are depicted as black dashed lines. **(E)** Superimposition of the region around the CW zinc finger of LSD2 with that of ZCWPW1 in complex with H3 peptide. **(F)** Close-up view of CW zinc fingers of LSD2 and ZCWPW1. A number of larger residues in LSD2 will block the binding of H3 and the H3 peptide will crash the C4H2C2 zinc finger and the SWIRM domain.

LI[33] and CATH revealed no significant similarity to other domains. Thus, the C4H2C2-type zinc finger appears to be novel and found exclusively within the LSD2 family of proteins.

The second zinc finger of LSD2 is of the CW-type (Figure 4D) composed of approximately 50 residues containing four conserved cysteine and two conserved tryptophan residues. The CW zinc finger has been found in several chromatin-associated proteins in animals and plants [34]. A sequence alignment of the LSD2 CW

and other known CW domains is shown in Figure 4C. Two tryptophan residues (W139 and W150) are highly conserved in all of the subgroups of CW domains. Like the structures of the CW domains from ZCWPW1 and ASHH2 [22, 23], the LSD2 CW domain is built around two β-strands and a Zn^{2+}-binding site (Figure 4A and 4D). Residues C142, C147, C169 and C185 chelate with the third zinc ion (Zn3). An extensive array of hydrogen bonds is observed between the side chains of L6 and the main chains of L9, and these multiple hydrogen-bond interactions play a major role in the formation of a compact CW fold.

The CW zinc finger in LSD2 is not likely required for the binding of H3 substrate

Two previous studies have shown that the CW zinc fingers recognize different methylated H3K4 histone tails [22, 23], raising the possibility that the LSD2 CW domain might also play a role in binding H3 substrates. We carried out peptide pulldown assays using various H3 tail peptides and the LSD2 CW zinc-finger peptide purified from *E. coli* or expressed by *in vitro* coupled transcription and translation. We found that the CW zinc finger did not bind to the H3 tail peptide (residues 1-21), with or without dimethylation at H3K4 (data not shown). Examining the binding of LSD2 mutant polypeptides that lack the CW zinc finger was not feasible due to the tendency of the mutant proteins to form aggregates.

To understand why the CW zinc finger of LSD2 lacks H3-binding activity, the structure of the region around the LSD2 CW zinc finger was aligned with that of the human ZCWPW1 (pdb:2RR4) using the Dali server (Figure 4E and 4F). Helices α13, α14, and loop L16 in the SWIRM domain, and the β1 and β2 strands of the C4H2C2 zinc finger occupy the position of the putative histone tail and would block the binding of the methylated histone peptide to the LSD2 CW zinc finger (Figure 4E). Compared with the small G265 residue of ZCWPW1, the corresponding positively charged R148 residue would be expected to repulse the positively charged R2 in the putative H3K4 peptide (Figure 4C and 4F). Likewise, R192, which corresponds to a hydrophobic or acidic residue in the other CW proteins, would repulse the positive charge of the histone tail. In addition, when compared with ZCWPW1, the bulkier side chains Y138, T143, L159 and K165 of LSD2 are expected to significantly reduce the size of the binding pocket (Figure 4C and 4F). Indeed, inspection of the molecular surface revealed no conspicuous cleft running across the surface of LSD2 CW zinc finger. In addition, the hydrogen bond networks between the side chain of R192, and the side chains of Q152, D190, the carboxyl oxygen of G88 cover the surface on the reverse side of the putative tryptophan cage [22, 23] and would be expected to block the binding of the substrate. These properties strongly suggest that the CW zinc finger of LSD2 is not suitable for the binding of histone tails.

The zinc-finger domain is required for LSD2 demethylase activity

Our previous study [15] demonstrated that the zinc-finger domain is required for LSD2 demethylase activity. However, the structure of LSD2 revealed no direct contact between the zinc-finger domain and the AOD, raising the question as to how the zinc-finger domain regulates demethylase activity. As the previous study was based on the analyses of LSD2 deletion mutants, we tested the role of the zinc-finger domain in demethylation by introducing single-site mutations that would specifically disrupt the structures of the C4H2C2 zinc finger (C53A, H84A and H90A) or the CW zinc finger (C185A) (Figure 4B and 4D). We found that all of these mutants failed to demethylate H3K4me2 in HeLa cells, as demonstrated by immunofluorescent staining of H3K4me2 (Figure 5A). Consistently, recombinant C53A, H84A and H90A mutant LSD2 proteins prepared from *E. coli* (Figure 5B) could not demethylate H3K4me2 from the core histone substrate (Figure 5C). These experiments indicate that the zinc-finger domain is absolutely required for LSD2 demethylase activity both *in vitro* and *in vivo*.

Notably the surface of the C4H2C2 zinc finger has a high density of basic residues, including four arginine and six lysine residues (Figure 4B). The sequence alignment indicates that six of these basic residues are conserved in the LSD2 homologs among different species (Supplementary information, Figure S4). This positively charged surface may potentially interact with phosphorylated RNA polymerase II [14] and/or DNA and contribute to H3K4 demethylation by LSD2. We tested the function of these positively charged residues by generating LSD2 mutants with reduced positive charge. Unlike the mutants with disruption of the zinc finger structure, these mutants generally maintained demethylase activity both in HeLa cells (Figure 5A) and *in vitro* (Figure 5C), suggesting that these positively charged residues are not essential for LSD2 demethylase activity toward H3K4me2.

The interaction between the zinc finger and SWIRM domains is required for LSD2 enzymatic activity

The above results have demonstrated that the function of the zinc finger is unlikely to be mediated by the binding of the H3 substrate. The structure of LSD2 suggests that the zinc-finger domain packs against the middle SWIRM domain, through extensive hydrophobic and
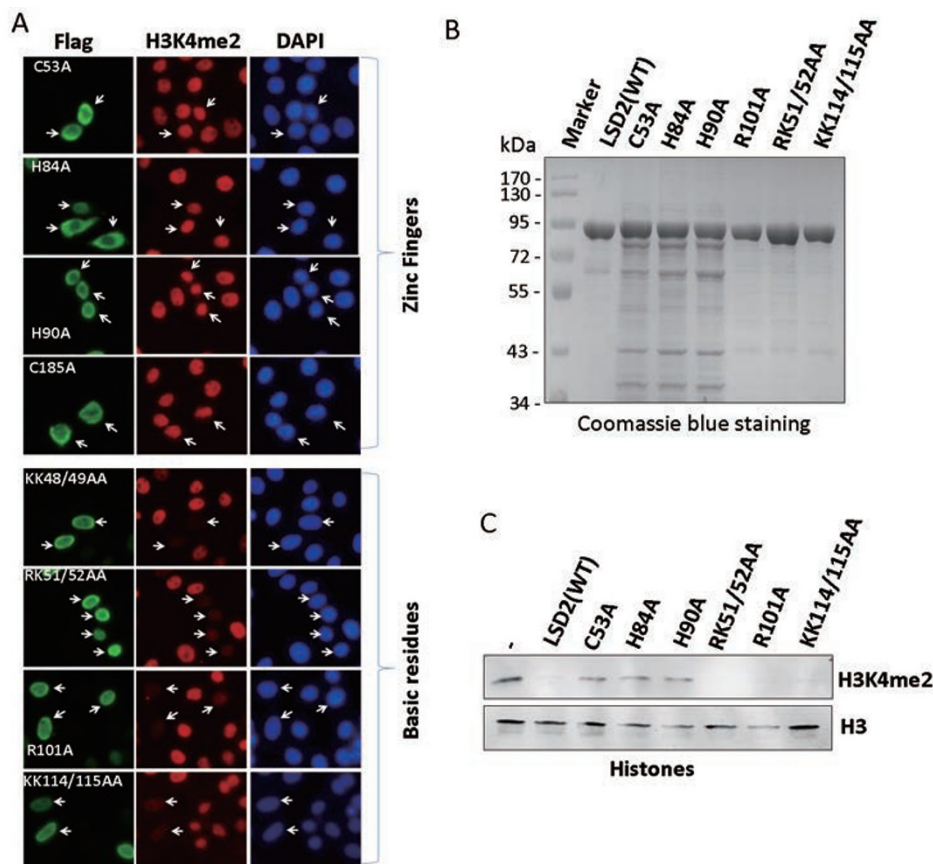
Figure 5 Both C2H4C2 and CW zinc fingers are essential for LSD2 demethylase activity. **(A)** Flag-tagged LSD2 mutants were transfected into HeLa cells and the demethylase activity was detected by immunofluorescent staining using H3K4me2 antibody. Arrows mark the cells in which the proteins of interest were overexpressed. All the mutants with disruption of zinc finger structure are enzymatically inactive. Note that the positively charged residues on the surface of C2H4C2 zinc finger are not essential for LSD2 demethylase activity. **(B)** The LSD2 and mutants expressed and purified from *E. coli* were resolved by SDS-PAGE and visualized by Coomassie blue staining. **(C)** The purified LSD2 and mutants were tested for demethylase activity using core histones as substrates. The levels of H3K4me2 were analyzed by western blot.

hydrophilic interactions, which in turn packs against the AOD domain. We reasoned that the observed interaction network through the zinc finger-SWIRM-AOD might be required to maintain LSD2 in its active conformation. To test this hypothesis, we first analyzed the interaction between the zinc finger and SWIRM domains. Among the residues involved in the extensive interaction between the zinc finger and SWIRM domains (Figure 6A), we mutated W82, W139, W150 and R151 in the zinc-finger domain and LV340/341 in the SWIRM domain. The W82A and R151A mutations were expected to affect the C4H2C2 and the CW zinc finger structure, respectively, and therefore indirectly affect the zinc finger-SWIRM interaction (Figure 4B and 4D), whereas W139A, W150A and LV340/341AA were expected to directly disrupt the interaction between zinc finger and SWIRM domains (Figure 6A). When tested in HeLa cells by immunofluo-

rescent staining against H3K4me2, all mutants were enzymatically inactive (Figure 6C), thus demonstrating the functional importance of the interaction between the zinc finger and SWIRM domains.

We next examined the functional importance of the interaction between the SWIRM domain and the AOD. Among the residues involved in this interaction (Figure 6B), we chose to mutate WY318/319, LI361/362 and IN443/444, as these residues are not the structural determinants of the SWIRM domain or the AOD. Again, when tested in HeLa cells by immunofluorescent staining against H3K4me2, these mutants were also defective in H3K4me2 demethylase activity (Figure 6D).

The relay of zinc finger-SWIRM-AOD interaction controls LSD2 binding of FAD
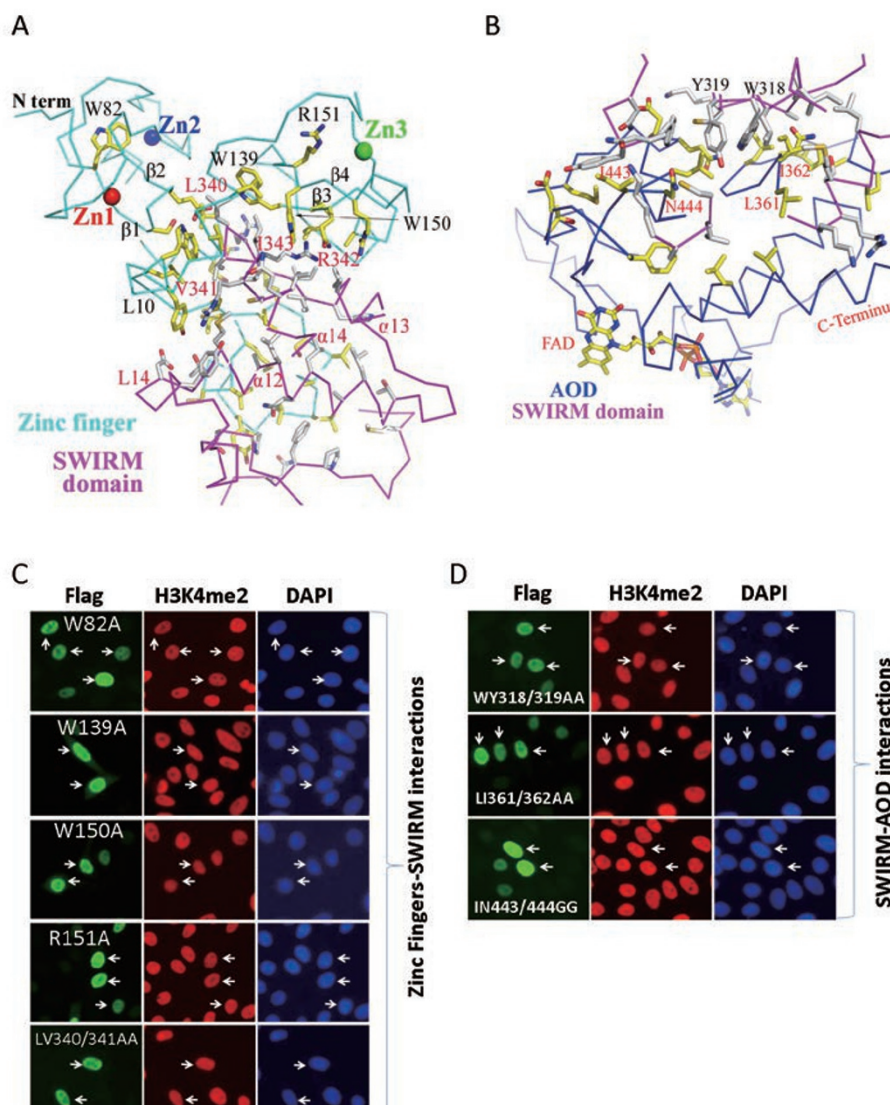
As the interaction between the zinc finger and SWIRM

Figure 6 The relay of the interaction through zinc finger-SWIRM-AOD is required for H3K4me2 demethylation by LSD2. **(A)** Interaction between the zinc-finger and SWIRM domain. **(B)** Interaction between the SWIRM domain and the AOD. **(C)** The mutants with disruption of the zinc finger-SWIRM interaction are enzymatically inactive. The mutants were overexpressed in HeLa cells by transient transfection and the demethylase activities were detected by immunostaining using anti-H3K4me2 antibody. **(D)** The mutants with disruption of the SWIRM-AOD interaction are also enzymatically inactive.

domains involves an extensive network of hydrophobic van der Waals contacts, disruption of this interaction might affect protein folding, leading to loss of enzymatic activity. However, the results of circular dichroism (CD) assays revealed that C53A, H84A (Figure 7A) and H90A (Supplementary information, Figure S5) mutant proteins exhibited a similar spectrum as the wild-type LSD2, indicating that these mutations did not induce gross conformational changes of LSD2.

As LSD2 demethylase activity depends on the non-covalent binding of FAD, we next analyzed whether the

mutant proteins could bind FAD. Spectrometry analysis revealed the presence of two characteristic absorbance peaks for FAD, one at 374 nm and one at 459 nm, for LSD2 purified from *E. coli* (Figure 7B). However, neither of the FAD-specific spectra were observed for any of the mutant proteins tested (Figure 7B). Furthermore, addition of FAD could not rescue the demethylase activity of these mutant proteins (data not shown), indicating that these mutants are defective in the binding of FAD. To further confirm that mutants with mutations in the zinc-finger domain are defective in the binding of FAD,
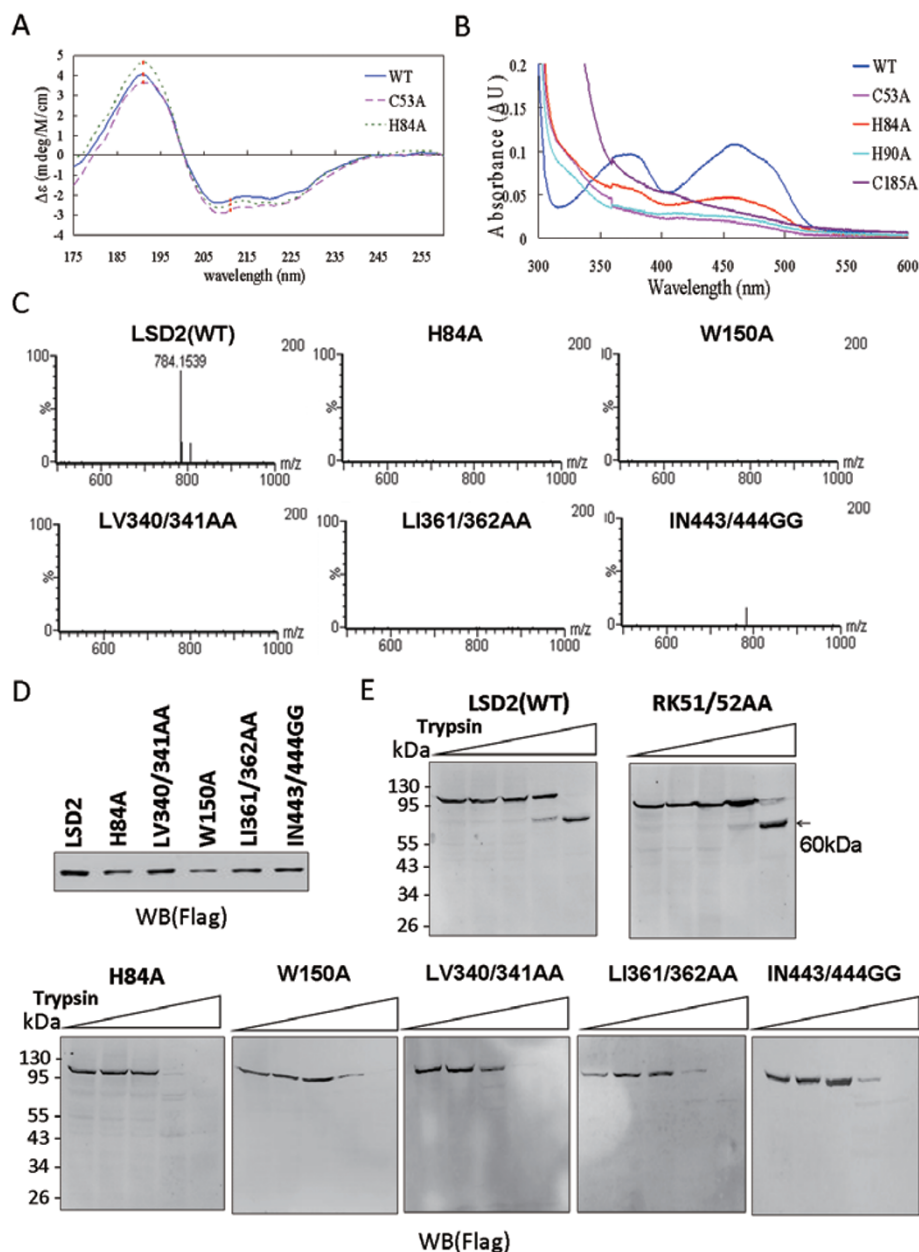
Figure 7 The zinc-finger domain and the relay of interaction among zinc finger-SWIRM-AOD are required for the incorporation of cofactor FAD in AOD. **(A)** Wild-type LSD2 and mutants with mutations in the zinc-finger domain have similar secondary structures. Superimposed Far-UV CD spectra of wild-type and mutant LSD2 are shown. The error bar is shown in red color. The spectral shapes of wild-type and mutant LSD2 are almost identical, and the values recorded are the same within the error of concentration measurements (~10%), demonstrating that the two mutants are well folded and their secondary structures are similar to wild-type LSD2. **(B)** Mutant LSD2 lacks FAD-associated spectrum. Superimposed absorption spectra of wild-type and mutant LSD2 are shown. The absorbance peaks at 374 and 459 nm are diagnostic of a flavoprotein. The experiments were repeated three times. **(C)** The wild-type and mutant LSD2 proteins as indicated were expressed and purified from 293T cells and subjected to the identification of FAD by mass spectrometry. The 784.1539 dalton peak represents FAD. H84A, zinc-finger mutant; W150A and LV340/341AA, mutants impaired in the interaction between zinc finger and SWIRM; LI361/362AA and IN443/444GG, mutants impaired in the interaction between SWIRM and AOD. Note a very small FAD peak was detected for IN443/444GG mutant. **(D)** The relative amounts of wild-type and mutant proteins used for mass spectrometry were determined by western blot analysis. **(E)** The C-terminal Flag-tagged wild-type and mutant proteins expressed and purified from 293T cells were partially hydrolyzed by an increasing concentrations of trypsin. A ~60 kDa fragment containing the SWIRM domain and the AOD was detected by anti-Flag western blot for the wild-type and basic residue RK51/52AA mutant, but not the other LSD2 mutants that are also defective in the binding of FAD and enzymatic activity.

we expressed wild-type and H84A mutant LSD2 in 293T cells and purified the proteins from whole-cell extracts via immune-affinity purification using anti-Flag M2 agarose beads. The purified proteins were boiled to release the FAD cofactor, which was then analyzed by a more sensitive technique, mass spectrometry. FAD with a molecular weight of 784.1539 dalton was detected for the wild-type but not the H84A mutant protein (Figure 7C). A similar result was observed for the H90A mutant (data not shown), indicating that the all the zinc-finger mutants are defective in the binding of FAD.

To test whether the mutants defective in the interaction either between the zinc finger and SWIRM domains or between the SWIRM domain and the AOD are able to bind FAD, we expressed and purified W150A, LV340/341AA, LI361/362AA and IN443/444GG mutants from 293T cells, quantified the relative amounts of the purified proteins by western blot analysis (Figure 7D), and measured their associations with FAD by mass spectrometry, as described above. The results in Figure 7C show that all of these mutants contained little, if any, FAD. Thus, like the zinc-finger mutants, the mutants with a disruption of the relay of interaction among zinc finger-SWIRM-AOD are also defective in binding FAD. These data explain the loss of enzymatic activity for all the mutants tested.

All the residues involved in FAD binding are located within the AOD (Supplementary information, Figure S6). The inability of the zinc-finger mutants to incorporate FAD suggests a structural alteration in the AOD of these mutant proteins. Indeed, the zinc-finger mutants were less stable and more prone to degradation than the wild-type LSD2 in *E. coli*, preventing us from obtaining crystal structures of any of the mutant proteins. We investigated the potential structural changes of the mutant proteins expressed and purified from 293T cells using limited trypsin digestion. Upon digestion with an increasing concentration of trypsin, the wild-type LSD2 protein produced a relatively stable fragment of ~60 kDa that included the SWIRM domain and the AOD (Figure 7E). In contrast, with the exception of the RK51/52AA mutant that maintained the demethylase activity, other mutants failed to generate the stable ~60 kDa fragment. We conclude that these mutations, which disrupt the zinc-finger domains or the interfaces of the zinc finger-SWIRM-AOD, lead to subtle conformational alterations in the AOD, which in turn impair the incorporation of FAD and result in a loss of demethylase activity.

Discussion

As the only other histone demethylase identified in the LSD1 amine oxidase family of proteins, LSD2 is emerging as an important epigenetic regulator [13-15, 18]. Here, we present the high resolution crystal structures of LSD2 in its apo form and in complex with a mimic peptide of H3K4me2, together with a comprehensive structural and functional analysis of human LSD2. Structural comparisons show both significant similarities and striking differences between LSD1 and LSD2 [19, 20]. Within the AOD, both proteins adopt a highly similar core structure, despite a low level of sequence identity (33%). The catalytic cavity is large enough to accommodate a stretch of 13 residues of the H3 tail peptide and H3K4me3, indicating that the LSD2 substrate specificity for H3K4me1/2 also lies in the nature of the chemical reaction rather than the substrate binding [8]. Comparison of the apo- and complex structures revealed that the binding of the H3K4M substrate does not significantly affect the structure of the AOD (Supplementary information, Figure S3). The major structural shift was actually observed at the L13 linker in the SWIRM domain and the zinc-finger region, most likely due to a series of allosteric conformational changes induced by substrate binding. Based on the structure, the SWIRM domain interacts extensively with the AOD and forms an elaborate interface maintained primarily through hydrophobic interactions. In addition, the SWIRM domain extensively interacts with both zinc-finger subdomains at the N-terminus. Given that the relay of the interaction among zinc finger-SWIRM-AOD is required for LSD2 enzymatic activity and binding of FAD, the binding of the H3 substrate would be expected to induce conformational changes in the SWIRM and zinc-finger domains to facilitate FAD binding and LSD2 activity.

From the structure, it is clear that the core structures of the SWIRM domain and the AOD are conserved between LSD1 and LSD2. Thus, it is not surprising that both LSD1 and LSD2 exhibit the same H3K4me1/2 demethylase specificity and that the SWIRM domain is required for demethylase activity for both proteins [8, 15]. Notably, in the LSD2 structure, the K4M residue is positioned in "front" of the cofactor FAD, with the side chain pointing toward the flavin ring of FAD (Figure 2A and 2E), while the N-terminal three residues of H3K4M are inserted deeply into the enzyme pocket (Figure 2D). Thus, the structure suggests that LSD2 is unlikely to demethylate H3K9me2 because the cavity is not deep enough to simultaneously allow the insertion of the first eight residues of the H3 peptide and position the K9 residue in front of FAD. Indeed, our *in vitro* demethylase assay revealed that both full-length LSD2 (Supplementary information, Figure S1) and LSD2 (residues 30-822) (Figure 1B) demethylated H3K4me2, but not H3K9me2.

However, our data do not rule out the possibility that LSD2 may demethylate H3K9me2 under different experimental conditions. Of note, the structural similarity with LSD1 in the AOD explains why PCPA (tranylcypromine), a clinically used antidepressant and an inhibitor of LSD1, also efficiently inhibits LSD2 [13]. Nevertheless, the surface structure of LSD2 AOD differs from that of LSD1 considerably (Figure 2C). These features may be explored in the future to design LSD2-specific inhibitors, thus avoiding potential side effects associated with inhibiting structurally similar enzymes, such as LSD1 and monoamine oxidases (MAOs).

Our study reveals that the predicted LSD2 N-terminal zinc-finger domain contains two distinct zinc fingers: C4H2C2-type and CW-type. The zinc fingers are small protein structural motifs that can coordinate one or more zinc ions and can be classified into several families by the type and order of the zinc-coordinating residues. They are structurally and functionally diverse and are generally involved in protein-protein and protein-DNA interactions. In this regard, the zinc-finger domain of LSD2 is likely to interact with a transcriptional repressor(s) as this region alone is able to repress transcription when tethered to a transcriptional reporter [15]. Different from other CW zinc finger-containing proteins, the LSD2 CW zinc finger is unlikely to bind a methylated H3 tail, as the reported H3-binding pocket is not observed in the LSD2 CW zinc finger due to steric hindrance. Interestingly, our structural study reveals a highly conserved and positively charged region on the surface of the C2H4C2 zinc finger. Unlike LSD1, LSD2 is able to demethylate both core histone and nucleosomal substrates [14, 15]. Although single- or double-amino acid mutations of the positively charged surface do not appear to affect the LSD2 demethylase activity towards histones, it remains possible that this region is involved in DNA interactions, thus conferring LSD2 with the ability to demethylate nucleosomal substrate [15]. In addition, this positively charged region may be involved in the interaction between LSD2 and the phosphorylated CTD of RNA polymerase II [14].

Our structural study reveals that the N-terminal zinc-finger domain of LSD2 is involved in an extensive intramolecular interaction with the SWIRM domain (Figure 4). In fact, both zinc-finger subdomains exhibit extensive interactions with the SWIRM domain. Significantly, we show that both zinc-finger subdomains are critically important for LSD2 demethylase activity and FAD binding. In fact, mutations that disrupt the relay of interactions through zinc-finger-SWIRM-AOD inactivate LSD2 and abrogate the binding of FAD. However, the effects of these mutations are unlikely due to a radical conformational change, as similar spectrums were observed for the wild-type and mutant proteins in CD analysis. Given that the mutant proteins are more sensitive to trypsin digestion and are incapable of binding FAD, we suggest that the zinc fingers and the relay of interactions among zinc finger-SWIRM-AOD are required for maintaining LSD2 in a proper active conformation for the binding of cofactor FAD and, consequently, enzymatic activity (Figure 8). It is noteworthy that the corresponding mutations in the LSD1 SWIRM domain abolish demethylase activity but have no effect on FAD binding [28]. Thus, the interaction between the SWIRM domain and the AOD may affect the AOD conformation to various extents in LSD1 and LSD2.

The zinc fingers are commonly observed in both histone methyltransferases and demethylases. For example, the post-SET domain in histone methyltransferases is also a zinc finger that forms part of the catalytic center and is essential for enzymatic activity [3-4]. In addition, a non-PHD zinc-finger domain at the N-terminus of JHDM2A/JMJD1A/KDM3A has been shown to be required for its demethylase activity [35]. Indeed, any mutation disrupting this zinc-finger structure abolishes the JMJD1A H3K9 demethylase activity (Supplementary information, Figure S7). Thus, the zinc-finger domain is not only a module used in protein-protein and protein-DNA interactions, but can also have a critical structural scaffolding function through intramolecular interactions (Figure 8).

Materials and Methods

Plasmids and antibodies

The pcDNA5-LSD2/AOF1/KDM1b-2xFlag was constructed as previously described [15]. All LSD2 point mutations were generated by site-directed mutagenic PCR reactions, according to the manufacturer's instruction.

Antibodies used in the study are: anti-H3 (Diagenode), anti-H3K4me2 (Millipore), anti-H3K9me2 (Abcam) and anti-Flag (Sigma-Aldrich).

Protein expression and purification

DNA fragments encoding various fragments of wild-type human or mouse LSD2 and their mutants were amplified by PCR and ligated into pGEX-4T-2 (GE Healthcare) and pET-28b (Novagen). The final clones were verified by DNA sequencing. The recombinant plasmids were transformed into BL21 (DE3) cells. Cells were grown at 310 K till OD_{600} reached 0.8-1.0 and protein production was induced at 291 K by the addition of 0.25 mM of IPTG and 100 μM $ZnSO_4$ for another 12 h. Cells were harvested by centrifuging at 4 000 rpm for 10 min, and then resuspended in buffer A (20 mM Tris pH 8.0, 500 mM NaCl) supplemented with PMSF followed by sonication. The cells were centrifuged at 20 000 rpm for 10 min and the supernatant was filtrated by 0.45-μm filter membrane to remove cell debris and other impurities. Soluble LSD2 was

Figure 8 A model for controlling the FAD binding and enzymatic activity of LSD2 through the relay of zinc finger-SWIRM-AOD interaction. The extensive interactions between zinc finger-SWIRM and SWIRM-AOD allow AOD in correct conformation for binding FAD, which is essential for LSD2 demethylase activity. Disruption of this chain of interaction leads to inappropriate conformation and failure of binding FAD.

purified with His or GST affinity chromatography according to the protocols [36], respectively (GE Healthcare). After eluted from the beads, fusion proteins were cleaved with the corresponding proteases at 277 K overnight. Then proteins were further purified by ion exchange (Q sepharose, GE Healthcare) and size exclusion chromatography (Superdex™-200, GE Healthcare). The protein was concentrated to 8 mg/ml in 20 mM Tris pH 8.5, 200 mM NaCl, 1mM TCEP. Sample purity was checked by SDS-PAGE.

Crystallization and data collection

Native monomer crystals were grown in well solution containing 21% PEG3350, 200 mM diammonium citrate, pH 7.0. Anomalous data were obtained from crystals grown at 289 K in a reservoir solution consisting of 20% PEG3350, 200 mM potassium phosphate pH 7.0. The LSD2-H3K4M complex crystals were obtained by soaking the LSD2 native crystals with a well solution containing 5 mM H3K4M peptide (residues 1-21). All the crystals were transferred into the well solution supplemented with 20% ethylene glycol and flashly frozen with liquid nitrogen before data collection.

The anomalous adsorption of the zinc atom in the crystal was collected on beamline 1W2B at the Beijing Synchrotron Radiation Facility. The other data were collected on beamline BL17U1 at the Shanghai Synchrotron Radiation Facility or beam line NE3A at the Photon Factory. All data were processed using HKL2000 [37]. Data collection statistics are summarized in Table 1.

Structure determination and refinement

The anomalous adsorption data (peak) of LSD2 were collected at 1.2800 Å, with two molecules in the asymmetric unit of orthorhombic space group I222 (Table 1). The number and position of the zinc atoms were confirmed using SAD method by shelex C/D program [26]. The structure was further solved by oasis [24] via molecular replacement methods using the atomic coordinates of

LSD1 (PDB code 2HKO [25]). The structures were further refined using REFMAC5 [38] and CNS [39]. The structures were fitted into electron-density maps using the graphics program COOT [27].

Measurement of FAD absorption spectra

FAD absorption spectra were recorded with TU-1900 double-beam UV spectrophotometers (Beijing Purkinje General Instrument) at room temperature. Wild-type or mutant proteins at approximately 1.5 mg/ml in buffer containing 20 mM Tris, pH 8.5 and 500 mM NaCl were analyzed in the wavelength ranged from 190 to 600 nm.

Circular Dichrosim (CD) analysis

The CD spectroscopy analysis was measured at 4B8 beamline in Beijing synchrotron radiation facility. Spectra were collected at 1 nm intervals over the wavelength ranging from 260 to 175 nm in a 0.0007-cm optical path length at temperature 299 K. A pure solvent baseline collected with the same cell was subtracted, all spectra data were processed using the CD tool software package [40]. The machine unit (mdeg) was converted into the per residue molar absorption unit, delta epsilon ($\Delta\epsilon$) in M/cm, by normalization with respect to polypeptide concentration and path length.

Cell culture and transient transfection

Hela and 293T cells were maintained in DMEM medium supplemented with 10% fetal bovine serum and 100 U penicillin-streptomycin (all from Invitrogen) in a 5% CO_2 atmosphere at 37 °C. Transient transfection of cells was carried out using Lipofectamine™ 2000 (Invitrogen) essentially according to manufacturer's instruction.

Immunostaining

For detecting LSD2 demethylase activity by immunostaining, Hela cells were transfected with the wild-type or mutant LSD2.

Twenty-four hours after transfection, the cells were fixed in 4% paraformaldehyde and permeabilized in cold PBS containing 1% Triton X-100. Permeabilized cells were blocked with 5% BSA for 30 min before incubation with primary antibodies (mouse anti-Flag, 1:1 000 dilution, and rabbit anti-H3K4me2, 1:200 dilution) for 2 h and secondary antibodies for 1 h in a 37 °C incubator. Then cells were incubated with 4, 6-diamidino-2-phenylindole dihydrochloride for 20 min and visualized under a fluorescent microscope.

In vitro demethylase assay

Bulk histones were incubated with purified wild-type or mutant LSD2 in the reaction buffer containing 50 mM Tris pH 8.5, 50 mM KCl, 5 mM $MgCl_2$, 0.5% BSA and 5% glycerol at 37 °C for 3 h. A typical 50 μl reaction contained 3 μg histones and 2.5 μg LSD2. Histone demethylation was then analyzed by western blot using antibodies as indicated.

Detection of FAD by mass spectrometry

Purified wild-type or mutant LSD2 were denatured at 100 °C for 10 min. Denatured protein was removed by centrifugation using a Centriprep YM-10 centrifugal filter (Amicon) according to the manufacturer's instructions. The samples were analyzed for the presence of FAD by an ACQUITY TQD mass spectrometer (Waters).

Limited trypsin digestion

293T cells transfected with the wild-type or mutant LSD2 were lysed in buffer (50 mM Tris pH 7.6, 1% Triton, 150 mM NaCl, 1 mM EDTA, 0.1% sodium deoxycholate). Total cell lysates were incubated with different concentration gradients of trypsin (Sangon) for 3 h on ice. The reaction was stopped by the addition of SDS-loading buffer and analyzed by western blot.

Isothermal titration calorimeter

The binding affinities between the LSD2 (residues 30-822) and the H3K4M were measured using TA instruments Nano ITC. The sample for ITC was prepared in a buffer containing 20 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP. Calorimetric titration was performed by injection of 2 mM H3K4M peptides into a cell containing 0.058 mM LSD2 at 4 °C .The ITC data were processed with NanoAnalyze software.

Accession numbers

The atomic coordinates have been deposited in the Protein Data Bank (PDB) as 4FWE, 4FWF and 4FWJ.

Acknowledgments

References

1 Barski A, Cuddapah S, Cui K, *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* 2007; **129**:823-837.

2 Zhang Y, Reinberg D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev* 2001; **15**:2343-2360.

3 Martin C, Zhang Y. The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* 2005; **6**:838-849.

4 Lan F, Shi Y. Epigenetic regulation: methylation of histone and non-histone proteins. *Sci China C Life Sci* 2009; **52**:311-322.

5 Allis CD, Berger SL, Cote J, *et al.* New nomenclature for chromatin-modifying enzymes. *Cell* 2007; **131**:633-636.

6 Jenuwein T, Allis CD. Translating the histone code. *Science* 2001; **293**:1074-1080.

7 Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000; **403**:41-45.

8 Shi Y, Lan F, Matson C, *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 2004; **119**:941-953.

9 Tsukada Y, Fang J, Erdjument-Bromage H. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 2006; **439**:811-816.

10 Whetstine JR, Nottke A, Lan F, *et al.* Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* 2006; **125**:467-481.

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

12 Shi YJ, Matson C, Lan F, *et al.* Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* 2005; **19**:857-864.

13 Karytinos A, Forneris F, Profumo A, *et al.* A novel mammalian flavin-dependent histone demethylase. *J Biol Chem* 2009; **284**:17775-17782.

14 Fang R, Barbera AJ, Xu Y, *et al.* Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic H3K4me2 methylation. *Mol Cell* 2010; **39**:222-233.

15 Yang Z, Jiang J, Stewart DM, *et al.* AOF1 is a histone H3K4 demethylase possessing demethylase activity-independent repression function. *Cell Res* 2010; **20**:276-287.

16 van Essen D, Zhu Y, Saccani S. A feed-forward circuit controlling inducible NF-kappaB target gene activation by promoter histone demethylation. *Mol Cell* 2010; **39**:750-760.

17 Core LJ, Lis JT. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* 2008; **319**:1791-1792.

18 Ciccone DN, Su H, Hevi S, *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 2009; **461**:415-418.

19 Yang M, Gocke CB, Luo X, *et al.* Structural Basis for CoR-EST-dependent demethylation of Nucleosomes by the Human LSD1 Histone Demethylase. *Molecular Cell* 2006; **23**:377-388.

20 Yang M, Culhane JC, Szewczuk LM, *et al.* Structural basis of histone demethylation by LSD1 revealed by suicide inactivation. *Nat Struct Mol Biol* 2007; **14**:535-539.

21 Lee MG, Wynder C, Cooch N, Shiekhattar R. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* 2005; **437**:432-435.

22 He F, Umehara T, Saito K, *et al.* Structural insight into the zinc finger CW domain as a histone modification reader. *Structure* 2010; **18**:1127-1139.

23 Hoppmann V, Thorstensen T, Kristiansen PE, *et al.* The CW domain, a new histone recognition module in chromatin proteins. *EMBO J* 2011; **30**:1939-1952.

24 Wu LJ, Zhang T, Gu YX, Zheng CD, Fan HF. Direct-method SAD phasing of proteins enhanced by the use of intrinsic bimodal phase distributions in the subsequent phase-improvement process. *Acta Crystallogr D Biol Crystallogr* 2009; **65**:1213-1216.

25 Chen Y, Yang Y, Wang F, *et al.* Crystal structure of human histone lysine-specific demethylase 1 (LSD1). *Proc Natl Acad Sci USA* 2006; **103**:13956-13961.

26 Sheldrick GM. A short history of SHELX. *Acta Crystallogr A* 2008; **64**:112-122.

27 Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 2010; **66**:486-501.

28 Stavropoulos P, Blobel G, Hoelz A. Crystal structure and mechanism of human lysine-specific demethylase-1. *Nat Struct Mol Biol* 2006; **13**:626-632.

29 Forneris F, Binda C, Adamo A, Battaglioli E, Mattevi A. Structural basis of LSD1-CoREST selectivity in histone H3 recognition. *J Biol Chem* 2007; **282**:20070-20074.

30 Aravind L, Iyer L. The SWIRM domain: a conserved module found in chromosomal proteins points to novel chromatin-modifying activities. *Genome Biology* 2002; **3**:research0039.

31 Holm L, Kaariainen S, Rosenstrom P, Schenkel A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* 2008; **24**:2780-2781.

32 Pugalenthi G, Archunan G, Sowdhamini R. DIAL: a web-based server for the automatic identification of structural domains in proteins. *Nucleic Acids Res* 2005; **33**:W130-W132.

33 Holm L, Rosenström P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* 2010; **38**:W545-W549.

34 Perry J, Zhao Y. The CW domain, a structural module shared amongst vertebrates, vertebrate-infecting parasites and higher plants. *Trends Biochem Sci* 2003; **28**:576-580.

35 Yamane K, Toumazou C, Tsukada Y, *et al.* JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* 2006; **125**:483-495.

36 Yu L, Wang Y, Huang S, *et al.* Structural insights into a novel histone demethylase PHF8. *Cell Res* 2010; **20**:166-173.

37 Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. In: Charles W. Carter J, Sweet RM, eds. Methods in Enzymology. Academic Press, 1997:307-326.

38 Murshudov GN, Skubak P, Lebedev AA, *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 2011; **67**:355-367.

39 Brunger AT. Version 1.2 of the Crystallography and NMR system. *Nat Protoc* 2007; **2**:2728-2733.

40 Lees JG, Smith BR, Wien F, Miles AJ, Wallace BA. CDtool-an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Anal Biochem* 2004; **332**:285-289.

41 Schrödinger L. The PyMOL Molecular Graphics System, Version 1.3r1. 2010.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)