BIOCHEMISTRY
including biophysical chemistry & molecular biology

Rapid Report

pubs.acs.org/biochemistry

LSD1/CoREST Reversible Opening−Closing Dynamics: Discovery of a Nanoscale Clamp for Chromatin and Protein Binding

Riccardo Baron* and Nadeem A. Vellore

Department of Medicinal Chemistry, College of Pharmacy, The Henry Eyring Center for Theoretical Chemistry, The University of Utah, Salt Lake City, Utah 84112, United States

Supporting Information

ABSTRACT: LSD1 associated with its corepressor protein CoREST is an exceptionally relevant target for epigenetic drugs. Hypotheses for the role of LSD1/CoREST as a multidocking site for chromatin and protein binding would require significant molecular flexibility, and LSD1/CoREST large-amplitude conformational dynamics is currently unknown. Here, molecular dynamics simulation reveals that the LSD1/CoREST complex in solution functions as a reversible nanoscale binding clamp. We show that the H3 histone tail binding pocket is a potential allosteric site for regulation of the rotation of SWIRM/SANT2 domains around the Tower domain. Thus, targeting this site and including receptor flexibility are crucial strategies for future drug discovery.

Post-translational modifications on the histone tails protruding from the nucleosome particle play fundamental roles in gene expression and chemically label the status of a gene as repressed or activated.[1] These histone labels form an epigenetic code that is recognized by transcription factors and dynamically regulated by specific histone-modifying enzymes.[2,3] Methylation of histone residues was considered irreversible until the discovery of lysine-specific demethylase 1 (LSD1 or KDM1A), the first lysine histone demethylase.[4−7] LSD1 associated with its corepressor protein CoREST catalyzes the oxidative, specific demethylation of Lys4 of histone H3 protein.[4−7] LSD1 is overexpressed in many solid tumors like breast, colon, neuroblastoma, bladder, small cell lung, blood, and prostate cancers and, thus, is an exceptionally relevant target for epigenetic drugs.[8] Hypotheses for the role of the LSD1/CoREST complex as a multidocking site are supported by the observation that its size and shape seem to be ideal for literally hugging mononucleosomes[9] and that the complex recruits the N-terminal tails of various chromatin proteins[10−13] by molecular mimicry of the histone H3 tail (Figure 1a,b).[13] However, formation of LSD1/CoREST complexes with such a diverse array of binding partners would require significant molecular flexibility, and LSD1/CoREST large-amplitude conformational dynamics is currently unknown. Here, we report a 0.5 μs explicit solvent molecular dynamics (MD) simulation revealing that the LSD1/CoREST complex in solution is remarkably flexible and also assumes configurations that are substantially more open or closed than the available X-ray crystal structures.[9] The distance between SWIRM and SANT2 domains oscillates reversibly, suggesting that LSD1/CoREST functions as a reversibly opening−closing nanoscale binding clamp. We demonstrate that the H3 tail binding pocket is a potential allosteric site for regulation of the rotation of SWIRM/SANT2 around the Tower domain and anticipate that targeting this site and including LSD1/CoREST flexibility are keys for future drug discovery.

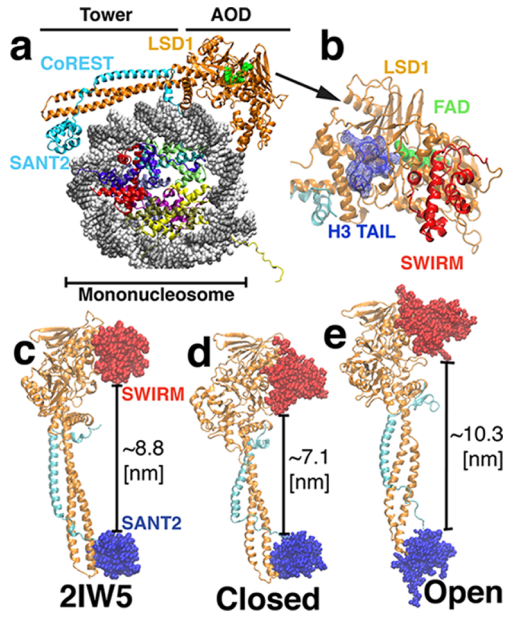

Figure 1. LSD1/CoREST complex and comparison with the nucleosome core particle shape and size. (a) LSD1/CoREST seems ideal for hugging mononucleosomes as a nanoscale clamp formed by the SWIRM domain of LSD1 (orange) and the SANT2 domain of CoREST (cyan). Only the simplest model for a 1:1 LSD1/CoREST−mononucleosome complex is displayed for graphical purposes. The LSD1 amino oxidase domain (AOD) and Tower domain are highlighted, together with the histone proteins. (b) View of LSD1 AOD ($FADH^-$ cofactor in green, H3 N-terminal tail in blue, from PDB 2V1D) after a 90° rotation for graphical purposes. (c−e) Comparison of the (c) X-ray crystal structure (PDB 2IW5) with (d) closed-clamp and (e) open-clamp MD snapshots in solution.

We initialized our MD simulation from the high-resolution X-ray structure of Yang et al.[9,14] (PDB 2IW5, 0.26 nm resolution) and generated a 0.5 μs long trajectory using

Received: January 16, 2012
Revised: March 2, 2012
Published: April 2, 2012

GROMACS.[15] Computational details are reported as Supporting Information. Figure S1 of the Supporting Information shows the values of the LSD1/CoREST backbone C^{α} atom root-mean-square deviation (RMSD) along the trajectory as a measure of the deviation from the initial X-ray reference model. RMSD values fluctuate along the trajectory consistently with the observation that the LSD1/CoREST complex is stable in solution. The LSD1/CoREST complex appears as a nanoscale binding clamp that can assume under standard conditions states clearly distinct from the X-ray crystal structure (Figure 1c–e). It is also clear that sampling of closed- and open-clamp states is reversible (Figure S2 of the Supporting Information) and that several hundred nanoseconds are needed to capture this large-amplitude motion. SWIRM–SANT2 distances between the center of mass of the domains vary from ∼11.2 nm in the X-ray structure[9] to ∼10 nm (closed-clamp) and ∼12.9 nm (open-clamp). In closed-clamp states, SWIRM and SANT2 domains move closer to each other (maximal and minimal interdomain distances of 13.0 and 7.1 nm, respectively); in open-clamp states, they move farther apart (17.3 and 10.3 nm, respectively). The distribution of RMSD values along our MD trajectory is bimodal (Figure S1b), in line with the pronounced oscillation of the distance between the SWIRM and SANT2 domains. The maximal, center of mass, and minimal distances between the C^{α} atoms of these domains fluctuate during the 0.5 μs around their average values, with striking amplitudes up to ∼4 nm and standard deviations up to 0.7 nm (Figure S1c,d). The novel observation of these fluctuations is in line with the expected adaptability of the LSD1/CoREST complex required to bind both chromatin and various chromatin proteins.

We investigated the dominant protein motion connecting these clamp states by principal component analysis (PCA) of protein fluctuations[16,17] with bio3d[18] (see Movies 1–3 of the Supporting Information). Taken as a whole, these results show that the three most dominant principal components (PC1–PC3) represent ∼88% of the overall protein backbone fluctuations (see Figure S3 of the Supporting Information). They all involve relative motion between the SWIRM/amino oxidase domain (AOD) and the SANT2 domain, confirming that the LSD1/CoREST clamp motion is the most dominant along the 0.5 μs MD trajectory. In detail, PC1 (54%) is responsible for the torsional clamp motion of the SANT2 domain with respect to the SWIRM domain around the axis defined along the LSD1/CoREST Tower domain (Movie 1). PC2 (27%) also describes the opening and closing of the clamp, yet without a pronounced torsional component (Movie 2). PC3 (7%) is a combination of both PC2 clamp opening and closing motions (Movie 3).

Figure 2 summarizes LSD1/CoREST conformational sampling in the two-dimensional space of the PC1 and PC2 lowest-eigenvalue frequency (highest-amplitude) components. From the initial X-ray structure (Figure 2a, black circle), the LSD1/CoREST clamp closes the first time after 13 ns and samples the first closed-clamp state up to ∼0.08 μs (Figure 2a, green circle 1). Suddenly, as typical of activated processes, the system leaves this state and assumes the most open configuration thus far after ∼55 ns (Figure 2a, red circle 2). The clamp resides in this open-clamp region up to ∼0.13 μs, then quickly closes back, and samples the closed-clamp state with the shortest SWIRM–SANT2 interdomain distance throughout the MD trajectory (in Figure 2a, green circle 3 highlights the closest clamp snapshot). At ∼0.29 μs, the clamp fully opens a second time and the LSD1/CoREST complex visits the most open clamp snapshot

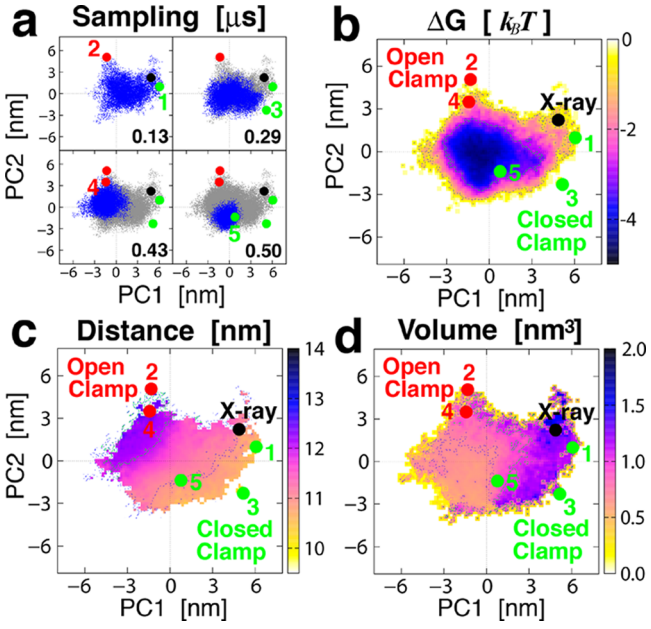

Figure 2. Reversible opening and closing of the LSD1/CoREST nanoscale binding clamp. (a) Principal component analysis of the motion of the LSD1/CoREST complex along the 0.5 μs MD trajectory in the two-dimensional space of the most dominant components, PC1 and PC2. Newly (blue) and previously (gray) visited regions are shown. Sampling starts from the X-ray structure (black circle). LSD1/CoREST visits closed-clamp states (1, 3, and 5), while states 2 and 4 are open-clamp states, numbered chronologically. (b) Corresponding relative free energy map calculated from probability distributions. (c) SWIRM–SANT2 interdomain distances (center of mass). (d) Volumes of the H3 pocket. See also Movies 1–3.

along the MD trajectory (in Figure 2a, red circle 4 highlights the most open clamp snapshot). This open-clamp region is sampled up to ∼0.43 μs. Lastly, we observe a second complete transition from the open-clamp to the closed-clamp form right before the end of the overall period of 0.5 μs (Figure 2a, green circle 5). We estimated the relative free energy surface corresponding to these dominant conformational transitions (Figure 2b,c). The free energy surface is characterized by a major well, and LSD1/CoREST oscillates across its minimum while reaching closed-clamp and open-clamp configurations that are up to ∼4k_BT higher in energy. These data are in line with the observation that the LSD1/CoREST clamp can reversibly visit open-clamp and closed-clamp states in solution following an oscillatory behavior.

A number of recent studies indicated LSD1/CoREST as a binding partner for various proteins involved in gene regulation and chromatin modification (see refs 10–13 and references therein). The (demethylated) nucleosomal particle(s) and transcription factors such as SNAIL1 have their primary N-terminal anchoring site in the LSD1 amino oxidase domain [AOD (Figure 1b and Figure S4a,b of the Supporting Information)].[13] SNAIL1, a master regulator of the epithelial–mesenchymal transition at the basis of many morphogenetic events, including the establishment of tumor invasiveness,[10] binds to LSD1/CoREST using a molecular mimicry recognition mechanism in which a transcription factor is optimally hosted in the H3 tail binding pocket (Figure S4a,b).[13] SNAIL1 is part of the large SNAIL/Scratch superfamily of transcription factors.[19] A high degree of binding sequence conservation of these SNAIL1-related proteins unravels the

dx.doi.org/10.1021/bi300068r | Biochemistry 2012, 51, 3151–3153

outstanding potential of LSD1/CoREST as a multiple-docking site for diverse chromatin proteins.[13]

The LSD1 H3 tail binding pocket is highly dynamic in our MD simulation, as summarized using the H3 pocket volume calculated with POVME[20] (Figure 2d). During the 0.5 μs period, the H3 pocket closes and opens reversibly, assuming also configurations more open or significantly more closed than the peptide-bound X-ray structures (PDB 2IW5, 1.49 nm^3; PDB 2VID, 1.42 nm^3; PDB 2Y48, 1.45 nm^3). The most dominant pocket configurations correspond to open-pocket and closed-pocket volumes of 1.1 and 0.54 nm^3, respectively. The most extreme open-pocket and closed-pocket configurations are sampled at ~0.13 μs, upon LSD/CoREST clamp closing, and at ~0.18 μs upon clamp opening (1.82 and 0.26 nm^3, respectively). We observe that the Lys triad formed by Lys355, Lys357, and Lys359 is highly flexible during our MD simulation and transiently gates the access to the H3 pocket. The three Lys residues act as fingers that reversibly expose their positive charges to the solvent or retract them to partially cover the entrance of the H3 pocket, helped by electrostatic attraction with residues Asp376 and Asp379. Figure S4 displays examples of these H3 pocket states and their comparison with X-ray structures. Overall, these data raise the fascinating hypothesis that substrate binding could trigger LSD1/CoREST nanoscale clamp opening and closing as an allosteric mechanism for chromatin and protein binding. Is the LSD1/CoREST clamp closing−opening correlated with the H3 pocket breathing dynamics?

To analyze this hypothesis, we investigated the two-dimensional space of PC1 and PC2 in two alternative ways: (1) using SWIRM−SANT2 interdomain distances determined from SWIRM and SANT2 centers of mass (Figure 2c) and (2) using the H3 pocket volumes and exploring how open-pocket and closed-pocket states are distributed with respect to the major LSD1/CoREST clamp motion (Figure 2d). A direct comparison between distance and volume maps highlights the correlation between LSD1/CoREST clamp closing and opening and H3 pocket breathing dynamics (cf. panels c and d of Figure 2). When the PC1 motion is followed along the horizontal axes, our results clearly show that a direct, inverse correlation exists between SWIRM−SANT2 interdomain distances and the volume of the H3 tail pocket. When SWIRM and SANT2 domains rotate and become far apart, the H3 pocket volume decreases, and vice versa. However, no evident correlation is found when moving along the PC2 direction (i.e., the vertical axis). This observation suggests that the H3 binding pocket is an allosteric site for regulating the rotation of SWIRM/AOD and SANT2 domains around the major Tower domain axis as captured by PC1 (see Movie 1). We expect the extent of this motion to be amplified in dimeric (and trimeric) assemblies of LSD1/CoREST complexes in which motional coupling between two (or three) clamps could regulate nucleosomal binding.

We reported the molecular dynamics of the LSD1/CoREST histone-modifying complex of great biophysical and biomedical relevance. We observed that H3 histone binding pocket breathing dynamics is anticorrelated with the rotation of SWIRM and SANT2 domains around the major Tower domain axis. This allows us to propose that the H3 binding site is a putative allosteric site regulating the opening−closing motion of the LSD1/CoREST clamp. The results presented in this report have high and immediate practical impact for the targeting of LSD1/CoREST interactions for epigenetic pharmacological goals. The discovery of a mechanism that regulates LSD1/CoREST clamp opening and closing might be a more general feature of other histone-modifying enzymes. We hope this computational study will stimulate the design of experiments for characterizing LSD1/CoREST dynamics.

■ ASSOCIATED CONTENT

Supporting Information

Figures S1−S4, supporting methodological details, and Movies 1−3 and their corresponding captions. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (801) 585-7117. E-mail: r.baron@utah.edu.

Funding

R.B. acknowledges startup funding from The University of Utah and the Department of Medicinal Chemistry.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The Center for High Performance Computing at The University of Utah is greatly acknowledged for technical support and resources.

■ REFERENCES

(1) Berger, S. L. (2007) *Nature 447*, 407−412.
(2) Kouzarides, T. (2007) *Cell 128*, 693−705.
(3) Ruthenburg, A. J., Li, H., Patel, D. J., and Allis, C. D. (2007) *Nat. Rev. Mol. Cell Biol. 8*, 983−994.
(4) Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004) *Cell 119*, 941−953.
(5) Forneris, F., Binda, C., Battaglioli, E., and Mattevi, A. (2008) *Trends Biochem. Sci. 33*, 181−189.
(6) Mosammaparast, N., and Shi, Y. (2010) *Annu. Rev. Biochem. 79*, 155−179.
(7) Shi, Y. (2007) *Nat. Rev. Genet. 8*, 829−833.
(8) Rotili, D., and Mai, A. (2011) *Genes Cancer 2*, 663−679.
(9) Yang, M., Gocke, C. B., Luo, X., Borek, D., Tomchick, D. R., Machius, M., Otwinowski, Z., and Yu, H. (2006) *Mol. Cell 23*, 377−387.
(10) Lin, Y., Wu, Y., Li, J., Dong, C., Ye, X., Chi, Y.-I., Evers, B. M., and Zhou, B. P. (2010) *EMBO J. 29*, 1803−1816.
(11) Forneris, F., Binda, C., Dall'Aglio, A., and Fraaije, M. (2006) *J. Biol. Chem. 281*, 35289−35295.
(12) Forneris, F., Binda, C., Adamo, A., Battaglioli, E., and Mattevi, A. (2007) *J. Biol. Chem. 282*, 20070−20074.
(13) Baron, R., Binda, C., Tortorici, M., McCammon, J. A., and Mattevi, A. (2011) *Structure 19*, 212−220.
(14) Yang, M., Culhane, J. C., Szewczuk, L. M., Gocke, C. B., Brautigam, C. A., Tomchick, D. R., Machius, M., Cole, P. A., and Yu, H. (2007) *Nat. Struct. Mol. Biol. 14*, 535−539.
(15) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) *J. Chem. Theory Comput. 4*, 435−447.
(16) Garcia, A. E. (1992) *Phys. Rev. Lett. 68*, 2696−2699.
(17) Amadei, A., Linssen, A. B., and Berendsen, H. J. (1993) *Proteins 17*, 412−425.
(18) Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A., and Caves, L. S. D. (2006) *Bioinformatics 22*, 2695−2696.
(19) Barrallo-Gimeno, A., and Nieto, M. A. (2009) *Trends Genet. 25*, 248−252.
(20) Durrant, J. D., de Oliveira, C. A. F., and McCammon, J. A. (2011) *J. Mol. Graphics Modell. 29*, 773−776.