G-quadruplex RNA binding and recognition by the lysine-specific histone demethylase-1 enzyme

ALEXANDER HIRSCHI,1,2 WILLIAM J. MARTIN,1,3 ZIGMUND LUKA,1 LIoudmila V. LOUKACHEVITCH,2 and NICHOLAS J. REITER1

1Department of Biochemistry, Vanderbilt University Medical Center, Nashville, Tennessee 37232-0146, USA
2Department of Pharmacology, Vanderbilt University Medical Center, Nashville, Tennessee 37232-6600, USA

ABSTRACT

Lysine-specific histone demethylase 1 (LSD1) is an essential epigenetic regulator in metazoans and requires the co-repressor element-1 silencing transcription factor (CoREST) to efficiently catalyze the removal of mono- and dimethyl functional groups from histone 3 at lysine positions 4 and 9 (H3K4/K9). LSD1 interacts with over 60 regulatory proteins and also associates with IncRNAs (TERRA, HOTAIR), suggesting a regulatory role for RNA in LSD1 function. We report that a stacked, intramolecular G-quadruplex (GQ) forming TERRA RNA (GG[UUAGGG]8UUA) binds tightly to the functional LSD1–CoREST complex (Kd ≈ 96 nM), in contrast to a single GQ RNA unit ([UUAGGG]4U), a GQ DNA ([TTAGGG]4T), or an unstructured single-stranded RNA. Stabilization of a parallel-stranded GQ RNA structure by monovalent potassium ions (K+) is required for high affinity binding to the LSD1–CoREST complex. These data indicate that LSD1 can distinguish between RNA and DNA as well as structured versus unstructured nucleotide motifs. Further, cross-linking mass spectrometry identified the primary location of GQ RNA binding within the SWIRM/amine oxidase domain (AOD) of LSD1. An ssRNA binding region adjacent to this GQ binding site was also identified via X-ray crystallography. This RNA binding interface is consistent with kinetic assays, demonstrating that a GQ-forming RNA can serve as a noncompetitive inhibitor of LSD1-catalyzed demethylation. The identification of a GQ RNA binding site coupled with kinetic data suggests that structured RNAs can function as regulatory molecules in LSD1-mediated mechanisms.

Keywords: LSD1; TERRA; IncRNA; ncRNA; G-quadruplex; chromatin; enzyme; kinetics; binding; mass spectrometry; structure; RNA–protein interactions

INTRODUCTION

Long noncoding RNAs (lncRNAs) are proposed to assist in a myriad of roles in the cell, acting as guides, scaffolds, decoys, or signaling molecules (Wang and Chang 2011). Although it is well established that distinct ncRNAs can act as gene regulators, recognize defined targets, and even function as catalysts (ribozymes) (Reiter et al. 2011), key mechanistic questions remain regarding how lncRNAs interact with and recruit chromatin-associated protein complexes to specific regions in the genome.

The lysine specific histone demethylase 1A (LSD1 or KDM1A) is an essential chromatin-remodeling enzyme conserved from yeast to humans and is also known to interact with lncRNAs (Shi et al. 2004; Stavropoulos et al. 2006; Khalil et al. 2009; Amente et al. 2013; Porro et al. 2014b). A primary function of LSD1 is to influence gene expression and chromatin structure by catalyzing the removal of mono- and dimethyl functional groups from Histone 3 proteins at lysine positions 4 and 9 (H3K4/K9) (Shi et al. 2004; Forneris et al. 2007; Laurent et al. 2015). LSD1 interacts with over 60 regulatory proteins, including transcription factors (CoREST, REST, p53, E2F1) and key enzymes (DNMT1, MRE11, HDAC1/2), as well as essential nutrients (tetrahydrofolate [THF]) (Luka et al. 2011; Kooistra and Helin 2012; Chatr-Aryamontri et al. 2013). Of these, CoREST is the primary interacting partner required for post-transla-

© 2016 Hirschi et al. This article is distributed exclusively by the RNA Society.
tional LSD1 stabilization and is required for H3K4 demethylation during development, hematopoiesis, and stem cell maintenance (Shi et al. 2004; Forneris et al. 2008; Hwang et al. 2011).

While distinct LSD1-containing protein complexes are established to repress or activate gene transcription (Shi et al. 2004; Laurent et al. 2015), it is unknown how lncRNAs bind and modulate these complexes. The telomeric repeat containing RNA (TERRA) is an integral component of telomeric heterochromatin, acts as a negative regulator of telomere length in human cells, and interacts with critical epigenetic regulators that include LSD1 demethylase and SUV39H1 methyltransferase enzymes (Porro et al. 2014a,b; Azzalin and Lingner 2015; Cusanelli and Chartrand 2015; Rippe and Luke 2015). In addition, TERRA has a strong propensity to form intramolecular, parallel-stranded G-quadruplex (GQ) RNA structures due to its repeating UUAGGG sequence (Patel et al. 2007).

Like many lncRNAs, TERRA remains associated with its parental chromatin, allowing cotranscriptional modulation of gene expression by epigenetic regulators. Telomeric gene silencing appears to correlate with methylation and demethylation patterns of H3K4/K9 histone modifications and it has been established that the TERRA–LSD1 interaction enhances the telomeric DNA damage response pathway (Porro et al. 2014b; Azzalin and Lingner 2015). Upon depletion of a shelterin component, the telomeric repeat factor 2 (TRF2), global TERRA levels increase in the cell and TERRA interacts directly with LSD1. Through an unknown mechanism, this RNA–LSD1 interaction subsequently stimulates the nuclease activity of the double strand break repair protein MRE11 (Fig. 1A; Porro et al. 2014b). Validation of these interactions in vivo and in vitro (Porro et al. 2014b) suggest that: (i) RNA binding to LSD1 is important at dysfunctional telomeres, and that (ii) TERRA may scaffold chromatin modifying enzyme complexes in a manner that is similar to other chromatin-associated lncRNAs such as HOTAIR and Xist (Tsai et al. 2010; Engreitz et al. 2014; Porro et al. 2014b; Somarowthu et al. 2015). Taken together, these data support a role for TERRA in recruiting proteins to modulate heterochromatin formation at chromosome ends.

Here we identify and characterize GQ nucleic acid binding on the surface of a functional LSD1–CoREST complex (Fig. 1B). We demonstrate that LSD1–CoREST has a strong preference to bind a stacked GQ-forming RNA and reveal that the primary binding site of the GQ RNA exists within the SWIRM/amine-oxidase domain of LSD1 (Fig. 1C). Cross-linking mass spectrometry and X-ray crystallographic data, analyzed RNA binding and the influence of distinct nucleic acid structures on the kinetics of LSD1-catalyzed demethylation, demonstrating that a GQ-forming RNA acts as a potential noncompetitive inhibitor of LSD1-catalyzed histone demethylation. These data indicate that structured RNAs can function as important regulators in LSD1-mediated pathways and supports an emerging theme that structure-specific RNA binding can influence the function of chromatin-associated proteins.

**RESULTS**

**Monovalent ions dramatically influence TERRA topology**

Quadruplex-forming nucleic acid sequences require specific monovalent ions for structural stability. In particular, potassium ions stabilize GQs, while lithium ions destabilize GQs (Balaratnam and Basu 2015). We used circular dichroism (CD) spectroscopy to monitor the effect of monovalent ions on GQ formation. CD spectra with a peak at 263 nm and a trough at 240 nm are characteristic of a parallel, propeller-type GQ conformation (Martadinata et al. 2011;
In both four and eight UUAGGG repeats (in the presence of K\(^+\) and Na\(^+\)), we observe the formation of stable parallel-stranded GQ structures (Fig. 2A). In contrast, Li\(^+\) destabilizes GQ structure and likely promotes a heterogeneous RNA architecture.

Sedimentation velocity analytical ultracentrifugation (SV-AUC) was also performed to assess RNA tertiary folding in solution, GQ oligomerization, and global compaction in the presence of different monovalent ions. Analysis of SV-AUC data reveal that a four repeat UUAGGG RNA has completely different tertiary shapes in the presence of K\(^+\), Na\(^+\), or Li\(^+\) (Fig. 2B; Supplemental Fig. 2). Whereas a four repeat UUAGGG RNA forms a monomeric GQ structure in Na\(^+\), the identical RNA forms a heterogeneous profile in the presence of K\(^+\). Further, SV-AUC analysis of the four repeat RNA with K\(^+\) reveals a sedimentation profile that is positioned between the four repeat RNA (Na\(^+\)) and an eight repeat UUAGGG RNA (K\(^+\)). Bayesian analysis suggests that the four repeat UUAGGG RNA with K\(^+\) exists in equilibrium between monomeric and stacked GQ structures. This result is consistent with previous nuclease digestion studies of TERRA (Martadinata et al. 2011; Martadinata and Phan 2013). Thus, choice of cation and RNA construct enables us to manipulate the tertiary RNA structure and examine the nucleic acid binding properties of LSD1.

Stacked GQ RNA preferentially binds to the LSD1–CoREST complex

In an effort to understand the nucleotide sequence and structural preferences associated with LSD1 and a stable LSD1–CoREST complex (Supplemental Fig. 1), we measured the LSD1–nucleic acid affinities from a panel of single and stacked GQ-forming oligonucleotides under different monovalent ions (Fig. 3; Supplemental Fig. 3). A TERRA RNA that contains a stacked (GG[UUAGGG]_{8}UUA) RNA, a single G-quadruplex-forming repeat element ([UUAGGG]_{4}U), a cognate DNA ([TTAGGG]_{4}T), and a sequence-unrelated 25-nt RNA (ssRNA) for which there is no predicted structure were incubated with enzymatically active LSD1. Because both LSD1 and CoREST localize at telomeres (Zhang et al. 2011) and LSD1 is optimally stabilized in the presence of CoREST (Shi et al. 2004; Forneris et al. 2005, 2008), we have primarily focused on how TERRA interacts with the LSD1–CoREST system.

**FIGURE 2.** Monovalent ions dramatically influence the structure of GQ-forming RNAs. (A) Parallel-stranded GQ RNAs are known to have an \(\theta_{\text{max}}\) of \(\sim 263\) nm (Balaratnam and Basu 2015). Circular dichroism spectroscopy demonstrates that parallel-stranded GQ structures form in the presence of potassium (black, solid triangle, and circles) and sodium (gray diamond and box), consistent with previous studies. In contrast, lithium (outlined triangles) destabilizes GQ formation. (B) The analysis of analytical ultracentrifugation (AUC) data of (UUAGGG)_{4}U and (UUAGGG)_{8}U RNAs in the presence of potassium (K\(^+\)), sodium (Na\(^+\)), and lithium (Li\(^+\)). Figure symbols as in A. The plot shows an overlay of the continuous distribution [C(s)] versus the sedimentation distribution coefficient (S).

**FIGURE 3.** Affinity and specificity of LSD1–CoREST binding to distinct nucleic acid structures is dependent on monovalent ions. (A) Analysis of gel-mobility shift assay binding curves of (UUAGGG)_{4}U, (UUAGGG)_{8}U, and 25-nt ssRNA. Assays were performed in potassium (K\(^+\)), sodium (Na\(^+\)), and lithium (Li\(^+\)) (symbols same as in Fig. 2) using LSD1–CoREST (amino acid residues 171–852 and 286–482, respectively) with exogenous protein purification tags removed. LSD1 strongly prefers to bind stacked GQ-forming RNA structures. The plot shows the fraction of RNA bound at various LSD1–CoREST concentrations (log scale). Error bars for each data point represent the range of three independent experiments. The dissociation constant (\(K_d\)) and Hill coefficient (\(h\)) from this analysis are reported in Table 1. (B) Representative gels showing that RNA binding activity of LSD1–CoREST is dependent upon the ability to form a GQ RNA conformation. Complexes and free oligonucleotides were resolved on a 0.6\% native agarose gel. The concentration of the LSD1–CoREST complex is noted for each lane (nM).
TABLE 1. Affinity of LSD1–CoREST binding to GQ-forming RNA

<table>
<thead>
<tr>
<th>RNA (monovalent)</th>
<th>Dissociation constant ($K_d$) (nM)</th>
<th>Hill coefficient ($h$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-GG(UUAGGG)₈U-3′ (K⁺)</td>
<td>96.4 ± 3.0</td>
<td>1.70 ± 0.08</td>
</tr>
<tr>
<td>5′-GG(UUAGGG)₈U-3′ (Na⁺)</td>
<td>515.6 ± 14.6</td>
<td>2.24 ± 0.10</td>
</tr>
<tr>
<td>5′-UUAGGG(UUAGGG)₄U-3′ (K⁺)</td>
<td>835.0 ± 34.7</td>
<td>1.07 ± 0.04</td>
</tr>
<tr>
<td>5′-UUAGGG(UUAGGG)₄U-3′ (Na⁺)</td>
<td>5200 ± 460</td>
<td>1.25 ± 0.08</td>
</tr>
<tr>
<td>5′-UUAGGG(UUAGGG)₄U-3′ (Li⁺)</td>
<td>10,500 ± 1900</td>
<td>1.3 ± 0.14</td>
</tr>
<tr>
<td>25-nt ssRNA (Li⁺)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The $K_d$ (in nM) and Hill coefficient ($h$) of the RNA–protein complexes are given. Experimental standard errors of the mean (±) incorporate a range of three independent electrophoretic mobility shift assays (EMSA).

GQ RNA can act as an inhibitor of LSD1-catalyzed demethylation

A G-quadruplex RNA preferentially inhibits LSD1-catalyzed demethylation activity over other nucleic acid structures and this inhibition trend qualitatively corresponds to the apparent equilibrium dissociation constants derived from EMSA data in Figure 3. To examine this inhibition preference, (UUAGGG)₄U, the cognate TERRA DNA (TTAGGG)₄T, an unstructured 25-nt RNA, and a short 6-nt RNA (UUAGGG) were incubated with the full length LSD1 without CoREST in the presence of a 21 amino acid K4 dimethylated peptide that mimics the H3K4 substrate. Initial velocity measurements were performed using an LSD1-peroxidase-coupled assay (Culhane et al. 2010) and the inhibition profiles were measured for each nucleic acid, enabling the extraction of apparent half maximal inhibitory concentration (IC₅₀) values (Supplemental Fig. 4A).

Kinetic measurements of the full-length LSD1-catalyzed demethylation reaction with and without RNA reveal that the (UUAGGG)₄U RNA is a reversible, noncompetitive inhibitor of LSD1 activity (Supplemental Fig. 4B). Increasing amounts of the (UUAGGG)₄U RNA (0, 1.5, 2.5 µM) show an overall decrease in maximal rate of the chemical reaction ($V_{max}$) but with no or nominal changes in the apparent binding affinity of the catalyst for the substrate ($K_M = 8.0 ± 0.9, 6.7 ± 0.9, 8.8 ± 1.5 µM$), respectively. These data suggest that RNA does not compete for substrate binding at the LSD1 active site.

Cross-linking mass spectrometry (XL-MS) identifies LSD1–GQ RNA binding interactions

Previous studies concluded that a GQ RNA likely associates with the SWIRM/AOD of LSD1 (Porro et al. 2014b). To more precisely identify the RNA binding regions of LSD1, a biotinylated GQ RNA (GG[UUAGGG]₈UU) was covalently cross-linked to purified LSD1–CoREST complex and subjected to high-resolution LC–MS/MS mass spectrometry. Two separate and independent XL-MS experiments were performed using purified LSD1–CoREST with and without a 6x-His tag at the N terminus of CoREST. Consistent with previous studies, analyses of both cross-link MS data sets indicate a strong GQ RNA cross-link to residues 227–251 of the SWIRM domain (Fig. 4).

As a specificity control, we observed that LSD1–CoREST covalently cross-links with GQ RNA but not a size-matched control RNA (Supplemental Fig. 6). Here, LSD1–CoREST was irradiated with 254 nm UV light either alone (control), in the presence of a 5²³P GQ RNA, or in the presence of a size matched 5²³P non-GQ-forming RNA. Protein and RNA concentrations were identical for each experiment. These results demonstrate that UV light specifically cross-links LSD1–CoREST to a stacked GQ RNA, consistent with EMSA data in Figure 3.

For XL-MS studies, two samples, the LSD1–CoREST complex alone (control) and in the presence of a 5² biotinylated GQ RNA, were subjected to 120 mJ of 254 nm UV light (Stratalinker 1800). The cross-linked RNP complex was then isolated using streptavidin beads and both the complex and the irradiated LSD1–CoREST (control) samples were
and likely represent a primary GQ RNA binding interface. Residues 527–550 are located adjacent to the N terminus of CoREST and near the LSD1 active site and may represent a secondary nucleic acid binding site (Fig. 4). The XL-MS experiment was replicated using modified conditions that include a stacked GQ RNA and a CoREST construct with a cleaved N-terminal 6×His tag. In both XL-MS experiments, depletion of the identical peptide regions spanning residues 227–251 and 527–550 were observed. A third cross-linked region (residues 689–726) was identified on the tower domain in one experiment; however, this region had very low coverage in the control and was not reproducible.

The structure of LSD1–CoREST bound to UUAGG identifies an ssRNA binding site adjacent to the primary GQ RNA–LSD1 cross-link location

To examine how RNA may inhibit LSD1 activity and identify an RNA binding location on LSD1, cocrystallization screening and extensive soaking experiments were performed with the LSD1–CoREST complex and different RNA molecules. The 53- and 25-nt GQ RNA and a series of short oligonucleotides (UUAGG, UUUUU, CCCUAA, UUAGGAG, AAAAAAGAAA, and GGUUUUUUUUUU) were soaked into crystals at various concentrations (0.1–5.0 mM) and incubation periods (4–24 h). Cocrystallization and crystal soaking trials of the GQ RNAs with LSD1–CoREST were unsuccessful and only crystal soaking experiments with UUAGG and UUAGGAG ssRNA fragments showed difference maps ([|Fo|−|Fc|] corresponding to oligonucleotide electron density (Fig. 5). From the UUAGG ssRNA soaked LSD1–CoREST crystals, the difference Fourier map ([|Fo|−|Fc|]) identified an RNA binding location on LSD1 and the directionality of the RNA. A 2.8 Å resolution map allowed us to accurately model build and define the ssRNA structure using RCrane and Coot (Emsley and Cowtan 2004; Keating and Pyle 2012) (Rwork = 20.1% Rfree = 22.7%, PDB 4XBF, Supplemental Table 1). In addition to the location of an ssRNA binding site on LSD1, the quality of the difference map ([|Fo|−|Fc|]) revealed well-ordered regions that correspond to FAD, sulfate ions, and a previously identified glycerol molecule (Fig. 5; Supplemental Fig. 7; Luka et al. 2011). The structure of LSD1–CoREST bound to UUAGG does not overlap with the GQ RNA binding location in the XL-MS experiments; rather, it is positioned adjacent to the GQ RNA binding region.

Nucleobase RNA recognition by the conserved amine oxidase domain (AOD) of LSD1

The structure of the LSD1–CoREST complex with ssRNA reveals that UUAGG RNA binds to the AOD of LSD1, which is located opposite of the histone 3 peptide binding cleft (Fig. 5). This RNA binding site is positioned at a conserved and “intertwined” region of the AOD, containing both N- and C-terminal regions (residues 278–280 and 615–619).
Interestingly, the UUAGG RNA does not electrostatically bind to the various electropositive grooves that exist across LSD1 which are surface accessible within the crystal lattice environment. Rather, the RNA nucleobases are oriented toward a cleft that is comprised of the β-sheet interface and three loop regions (Lys276-Lys280, Ala597-Asn599, Pro782-Arg795). Unfortunately, point mutations of Lys280, Tyr616, Lys617, and Asp619 to alanine resulted in a loss of recombining protein expression and protein stability, making it difficult to assess the effect of RNA binding activity within this region of LSD1. Taken together, the RNA soaking studies identify how LSD1 may interact with the single-stranded linker regions of TERRA, supporting an extended TERRA binding interface that spans the SWIRM and AOD of LSD1.

The identification of an ssRNA binding region and the orientation of the RNA also demonstrate that LSD1 can interact with nucleic acids through the formation of sequence specific interactions. The most extensive RNA–protein contacts occur between the AG dinucleotide region and LSD1 residues that are part of a preformed binding pocket (Fig. 5C). This single-stranded oligonucleotide binding region of LSD1 may correlate with a genome-wide analysis of nucleotide binding elements of the LSD1 complex, which originally proposed that an AG dinucleotide region is part of an enriched motif in LSD1 binding (Tsai et al. 2010). Together with our kinetic assays (Supplemental Fig. 4), the structure provides additional support that RNA can serve as a noncompetitive allosteric inhibitor of LSD1 function.

**DISCUSSION**

TERRA is an integral component of telomeric heterochromatin, forms distinct G-quadruplex structures in vitro and in vivo, and has been demonstrated to interact with critical epigenetic regulators that include LSD1 demethylase and SUV39H1 methyltransferase enzymes (Luke et al. 2008; Schöffner and Blasco 2008; Deng et al. 2009; Xu et al. 2010; Iglesias et al. 2011; Lopez de Silanes et al. 2014; Porro et al. 2014a,b; Azzalin and Lingner 2015). Despite a correlation between histone H3K4/K9 methylation patterns and gene silencing at telomeres (Krogan et al. 2002; Vaquero-Sedas et al. 2012; Porro et al. 2014b), it is unclear how these enzymes are recruited to telomeres and how they might function to mediate telomere structure and composition. TERRA is thought to act as a molecular decoy to sequester telomerase in a cell cycle-dependent manner and also likely serves as a scaffold with epigenetic regulators during the early stages of DNA damage response activation (Redon et al. 2010; Porro et al. 2014a,b). The unique topology of TERRA (Fig. 1C) suggests that its structure may enable the organization of higher-order RNP complexes at telomeres. To better understand how lncRNAs may recruit protein enzymes and how TERRA might regulate protein–protein interactions at the ends of chromosomes, we determined the mode of GQ RNA binding to LSD1.

Monovalent ions dramatically influence the topology of GQ RNAs and knowledge of specific TERRA sub-structures enabled us to manipulate the GQ RNA architecture (Fig. 2; Supplemental Fig. 2). The formation of these diverse RNA structures correlates well with observed nucleic acid–LSD1 binding affinities (Fig. 3; Supplemental Fig. 3), revealing that LSD1 strongly prefers a stacked GQ (GG-[UUAGGG]₈ UUA) RNA. Previous TERRA RNA studies combining RNase T1 digestion and molecular dynamics (MD) simulations demonstrate that TERRA (containing up to 96 UUAGGG repeats) primarily assembles into four and eight UUAGGG repeat units (Martadinata et al. 2011; Martadinata and Phan 2013). It appears that a single ([UUAGGG]₈ UU) GQ unit, a stacked ([UUAGGG]₈ U) GQ unit, and a ssRNA spacer region between GQ units represent the three unique topologies adopted by TERRA (Martadinata et al. 2011; Martadinata and Phan 2013). In addition, biophysical studies of TERRA demonstrate that the 2′–OH functional groups in the RNA G-quadruplex participate in organizing water hydration and in the hydrogen-bonding network (Haider et al. 2011; Martadinata et al. 2011). This may contribute additional stability to the parallel-stranded quadruplex conformation and account for why LSD1 prefers to bind GQ RNA over a cognate GQ DNA (Supplemental Fig. 3), which is known to contain a mixture of heterogeneous
G-quadruplex topologies (Patel et al. 2007). Our data confirm that single and stacked GQ RNA structural units serve as building blocks of the extended TERRA and that the propensity for stacking in GQ RNA molecules may be a key feature in the recruitment of protein–protein complexes at telomeres.

For the recognition of TERRA by LSD1, the RNA structure and shape play a central role in protein–RNA binding specificity. Although GQ RNA shape-based recognition is not well established, it is known that other protein domains can also preferentially recognize GQ-forming RNAs including the arginine–glycine–glycine repeat (RGG) domain of Fragile-X mental retardation protein (FMRP), the glycine-arginine-rich (GAR) domain of TRF2, and the N-terminal domain of the DEAH-box ATP-dependent helicase 36 (DHX36) (Deng et al. 2009; Meier et al. 2013; Chen et al. 2015; Vasilyev et al. 2015). With the development of GQ-specific antibodies that have unambiguously identified GQ TERRA RNA structures in living cells (Xu et al. 2010; Di Antonio et al. 2012; Biffi et al. 2013, 2014), more biophysical studies are needed that probe how protein domains and multidomain protein complexes recognize the molecular architecture of GQ RNAs. Our data have defined how TERRA’s secondary and tertiary structural motifs can serve as recognition elements for LSD1 interactions by coupling biophysical studies (CD/AUC), EMSA, enzyme kinetics, XL-MS, and X-ray crystallographic methods. Results from our studies suggest that TERRA could serve as an allosteric effector of LSD1 function at telomeres, as previously proposed (Porro et al. 2014b).

Functional implications for allostery in LSD1–RNA interactions

LSD1 is associated with CoREST and together they act as an allosteric clamp on nucleosomes to catalyze specific histone H3K4/K9 demethylation (Shi et al. 2004; Chen et al. 2006; Stavropoulos et al. 2006; Yang et al. 2006; Forneris et al. 2007; Baron and Vellore 2012). Motions of the SWIRM domain and rotation of the AOD of LSD1 must occur when the substrate enters the active site pocket (Baron and Vellore 2012) and it has been demonstrated that both substrate binding and protein–protein interactions modulate LSD1 conformational dynamics and activity (Shi et al. 2005; Forneris et al. 2007).

Our data suggest that GQ-forming RNAs can provide an additional layer of allosteric regulation in LSD1 function. Using XL-MS, we identified a GQ RNA binding site distal to the H3 binding cleft. X-ray crystallographic soaking experiments of LSD1–CoREST with various ssRNAs suggest that the RNA binding interface spans the amine–oxidase domain (Fig. 6). While it is possible that there may be multiple regions of LSD1–CoREST that contact nucleic acids, the identified primary RNA binding interface likely extends from the SWIRM domain to the AOD of LSD1. Interestingly, the structural integrity and flexibility of the SWIRM-AOD is essential for recognition of nucleosomal DNA (Stavropoulos et al. 2006; Yang et al. 2006; Forneris et al. 2007). In fact, our identified GQ RNA binding region within the SWIRM domain (amino acids 227–251) overlaps with the putative nucleosomal DNA binding surface (Yang et al. 2006; Pioletto et al. 2015), suggesting that both nucleosomal DNA and TERRA RNA bind a similar SWIRM/AOD recognition interface. A putative mechanism whereby TERRA alters LSD1 demethylation activity and modulates nucleosome binding or DNA damage response activation at telomeres is consistent with our RNA–LSD1 structural and enzyme inhibition results.

Whereas the biological consequences of LSD1 activity are known, mechanistic examples of RNA-mediated LSD1 recruitment (Khalil et al. 2009; Tsai et al. 2010; Porro et al. 2014b; Hendrickson et al. 2016) and the modulation of LSD1 activity by noncoding RNAs remain unclear. The regulation of LSD1 complexes is likely to be multilayered and may be differentially influenced by distinct classes of RNA molecules. For example, RNA helicases are known to actively remodel TERRA RNA at telomeres (Cusanelli and Charlard 2015; Flynn et al. 2015; Rippe and Luke 2015). It is possible that RNA helicases provide an added level of epigenetic regulation by unwinding GQ RNAs, resulting in a loss of LSD1 binding and effectively relieving RNA-mediated LSD1 inhibition. Such a molecular mechanism has yet to be demonstrated though it should be noted that several groups have proposed a link between helicases and the regulation of chromatin modifying enzymes by ncRNAs (Cifuentes-Rojas et al. 2014; Kaneko et al. 2014; Sarma et al. 2014; Davidovich et al. 2015; Cloutier et al. 2016).

In conclusion, we have identified a primary GQ RNA binding site within the conserved SWIRM/AOD interface of LSD1 and demonstrated that a GQ-forming TERRA can function as a noncompetitive inhibitor of LSD1-catalyzed demethylation. Future structural studies will be required to demonstrate how a GQ-forming RNA alters LSD1 structure and how TERRA influences the function of LSD1 at telomeres. Defining the structural interactions of TERRA with LSD1 will provide insight into the diversity of GQ RNA–protein recognition and serves as an important model system to
explore lncRNA–protein recruitment mechanisms at the atomic level.

MATERIALS AND METHODS

LSD1, CoREST, and RNA preparation

The plasmids for the N-terminal truncated LSD1 (aa 171–852) and CoREST (aa 286–482 plus His-tag sequence) were a generous gift of Dr. Cole (Johns Hopkins University) and the full-size LSD1 plasmid (aa 1–852) was a generous gift of Dr. Shi (Harvard University) (Supplemental Fig. 1A; Shi et al. 2004; Culhane et al. 2006). The full size and truncated LSD1 were expressed in Escherichia coli BL21(DE3) as previously reported (Shi et al. 2004; Culhane et al. 2006, 2010). The full-size LSD1 was purified by using ammonium sulfate fractionation and anion-exchange chromatography (Luka et al. 2011). The truncated LSD1 was purified using GSH-agarose chromatography, with the glutathione S transferase tag removed through PreScission protease (GE Healthcare) digestion and anion-exchange chromatography (Luka et al. 2011). CoREST is a known stabilizing element of LSD1 (Shi et al. 2005) and was expressed in E. coli BL21(DE3) in Luria Broth (LB) (kanamycin) and purified using Ni-NTA agarose as previously reported (Luka et al. 2011). For EMSA assays and crystallization trials, truncated LSD1 (171–852) was coexpressed with CoREST (286–482) in Rosetta(DE3) pLysS competent cells. Complexes were purified via glutathione and nickel affinity and size-exclusion (Superdex 200) chromatography. The concentration of protein samples was determined by the BCA method (BCA Protein Assay kit, Pierce) with bovine serum albumin as a standard and by UV-VIS-spectroscopy for LSD1 preparation with extinction coefficient for FAD at 450 nm as 11300 M$^{-1}$cm$^{-1}$. Protein spectra were recorded on Shimadzu 2401-PC Spectrophotometer. Protein purity was determined by SDS electrophoresis with Coomassie staining (Supplemental Fig. 1B–D).

Stacked GQ RNAs (GG[UUAAGG]$_4$UUA) were transcribed with T7 RNA polymerase and purified on 6% TBE-polyacrylamide gels supplemented with 8 M urea according to standard methods. Dried RNA pellets were dissolved in RNase-free water and diluted to 5 μM in 10 mM Tris–HCl pH 7.4, 1 mM TCEP and 100 mM KCl, NaCl, or LiCl as indicated in the main text. RNAs were folded using a standard protocol (2 min at 95°C, 5 min at 85°C, 5 min at 75°C, 5 min at 55°C, 15 min at 37°C, and then placed on ice). All CD spectra were recorded at room temperature on a Jasco J-810 spectropolarimeter with a 1 mm cell, scan speed of 50 nm/min, and a response time of 1 sec. Spectra from 300–220 nm were averaged over three scans, and background from a matched buffer-only sample was subtracted.

Analytical ultracentrifugation (AUC)

RNA samples were run for 16 h in an Optima XLI ultracentrifuge equipped with a four-hole An-60 Ti rotor at 48,000 rpm at 4°C. Samples and buffer-matched blanks were loaded into double-sector cells (path length of 1.2 cm) with charcoal-filled Epon centerpieces and sapphire windows. Data were fit to a continuous c(s) distribution model using SedFit, with a partial specific volume of 0.73, buffer density of 1.005, buffer viscosity of 0.0102, and a frictional ratio of 1.4. Standard Bayesian modeling operations included in the Sedfit software enabled us to deconvolute asymmetric peaks.

Electrophoretic mobility shift assay (EMSA)

All oligonucleotides were 5′-end labeled with T4 Polynucleotide Kinase (NEB) and [$\gamma$-32P] ATP (6000 Ci/mmol, 10 μCi/μL, PerkinElmer). Unincorporated radiolabel was removed by application to Micro Bio-Spin columns packed with Bio-Gel P6 in 10 mM Tris–HCl pH 7.4, 0.02% sodium azide (Bio-Rad) according to manufacturer’s instructions. Immediately prior to binding, radiolabeled oligonucleotide stocks (typically 1–2 μM) were diluted to 20 nM in EMSA buffer (25 mM HEPES pH 7.4, 1 mM TCEP, 10% glycerol, 0.02% bromophenol blue, 1 U/μL RNasin (Promega) and 100 mM KCl, NaCl, or LiCl, depending on reaction conditions) and folded as described (CD spectroscopy, Materials and Methods section). To initiate binding reactions (10 μL final volume) threefold serial dilutions of 50 μM LSD1–CoREST in EMSA buffer were mixed 1:1 with 20 nM oligonucleotide stocks and incubated at room temperature for 20 min. Reactions were placed on ice and chilled for 5 min before loading into 0.75% THE (34 mM Tris base, 66 mM HEPES free acid, 0.1 mM EDTA, pH 7.4) agarose gels supplemented with 10 mM potassium acetate, sodium acetate, or lithium sulfate depending on reaction conditions. Gels were run for 45 min at 6 V/cm in THE running buffer supplemented with appropriate salt (10 mM), with constant buffer recirculation at 4°C. Gels were exposed to an Imaging-Screen K (Kodak) and images were collected with a Pharus FX Plus Molecular Imager (Bio-Rad). All binding reaction profiles were quantified using the Quantity One 4.6.9 (Bio-Rad) software package. Only signal corresponding to fully bound or unbound positions was analyzed; smears due to complex dissociation were not included. The integrated volume for each signal was determined by measuring the identical area that encompasses the probe–only control with minimal background. Results of oligonucleotide binding assays were expressed as the fraction of oligonucleotide bound and plotted as a function of protein concentration using Prism 6.0 (GraphPad Software, Inc., http://www.graphpad.com). Data were fit to a one-site hyperbolic...
binding function including a Hill coefficient \(Y = B_{\text{max}} \times X^h / (K^h_x + X^h)\), where \(Y\) is the fraction bound and \(X\) is the protein concentration (nM), and \(h\) is the Hill coefficient. An average of the Hill coefficient \((h)\) was determined by finding the slope of a straight line fitted to points from a plot of log \(\theta/(1-\theta)\) versus log of the protein concentration, where \(\theta\) is the fraction of bound oligonucleotide.

**Activity assay**

LSD1 activity assays were performed using a peroxidase-coupled assay under aerobic conditions as described previously (Forneris et al. 2005). A 150 µL reaction mixture contained 50 mM HEPES(Na), pH 7.5, 0.2–0.3 µM LSD1, 1 µg of HRP, 0.1 mM 4-aminonitroantipyrine, 1.0 mM 3,5-dichloro-2-hydroxybenzenesulfonic acid and substrate concentrations as indicated in the main text. Enzyme reactions were initiated by the addition of substrate (dimethyl-Lys4 Histone H3 peptide aa 1–21 [diMeK4H31–21] \(\Sigma\gamma\)) and the time course of the reaction was monitored using a Shimadzu 2401-PC spectrophotometer (\(\lambda = 515\) nm) at 25°C in a thermostatized quartz cell chamber. Initial velocity values were measured using an extinction coefficient of 26,000 M\(^{-1}\) cm\(^{-1}\) (Forneris et al. 2005). Kinetic data were fitted to a Michaelis–Menten equation using GraphPad Prism 6.0 (GraphPad Software, Inc., http://www.graphpad.com).

**Inhibition studies**

The IC\(_{50}\) values for different oligonucleotides were obtained via the LSD1 activity assay in the presence of different concentrations of RNA/DNA as detailed (Supplemental Fig. 4). Control experiments showed that the peroxidase components of this coupled assay were not influenced or altered by the addition of nucleic acids. The mechanism of inhibition was examined with velocity/substrate concentration curves at different concentrations of RNA in the reaction mixture as detailed in the figure legend (Supplemental Fig. 4).

**Cross-linking mass spectrometry**

To verify that LSD1–CoREST cross-links to GQ RNA in a specific manner, an (GG)[UUAGGG][\(\Sigma\gamma\)UUA] RNA and a control non-GQ-forming, size-matched RNA were 5′ S2P labeled. The RNAs were folded as described (CD spectroscopy, Materials and Methods section). The RNA was diluted to 1.4 µM in folding buffer plus 0.3 U/µL RNase Inhibitor (Promega, N2111) and incubated with or without 1.6 µM LSD1–CoREST at room temperature for 10 min. The reaction was then placed on ice and exposed to 240 ml/cm\(^2\) of 254 nm UV light with a Stratalinker 1800 and separated on a 4%–20% Mini-Protein TGX Precast Gel (Bio-Rad, 4561094). The gel was exposed using an Imaging Screen-K (Bio-Rad, 1707841) and visualized on a PharoSFX imager using the Quantity One software system (Supplemental Fig. 6).

The same general protocol was followed to identify the cross-linked peptides, with a few important distinctions. 5′ biotinylated RNA was purchased from IDT and folded as described above. 20 µM GQ RNA and 40 µM LSD1–CoREST were incubated together for 10 min at room temperature before being placed on ice and UV-cross-linked alongside a negative-control sample of LSD1–CoREST without RNA. The covalently cross-linked LSD1–RNA complex was enriched using Sera-Mag magnetic streptavidin-coated beads, medium binding (Genesee Scientific, 85–92) and a biotinylated RNA pull-down kit according to the manufacturer’s instructions (Pierce, 20164). 0.5% SDS and 1% SDS were added to the RNA capture buffer and RNA wash buffers, respectively, as these conditions removed non-cross-linked LSD1 without disrupting the biotin–streptavidin interaction (data not shown). The complex was eluted by boiling in 1xSDS loading buffer and gel purified via SDS-PAGE. The cross-linked LSD1 (control) and LSD1–RNA (sample) were visualized using colloidal Coomassie blue stain, cut out, and subjected to in-gel trypsin digestion overnight. Two separate cross-linking LC–MS/MS experiments were performed on two freshly purified LSD1–CoREST samples with a single GQ RNA and with a stacked GQ RNA. All samples were injected onto an LTQ Orbitrap high-resolution LC–MS/MS system (Vanderbilt Proteomcs Core Facility). The isotopic distribution was used to confirm the identity of peptide peaks covering ~75%–80% of LSD1 residues using SkyLine 3.5 (Schilling et al. 2012).

**Crystallization of LSD1–CoREST complex**

The LSD1–CoREST complex was crystallized as previously described (Yang et al. 2006; Luka et al. 2011). Briefly, the LSD1–CoREST complex was prepared by mixing the LSD1 and CoREST stock solutions in a 1:1.5 molar ratio of LSD to CoREST and removal of excess CoREST by using an Amicon Ultra 50K Centrifugal Filters after 1-h incubation. The LSD1–CoREST complex was crystallized by the hanging drop or sitting drop method at room temperature as previously described (Luka et al. 2011). The LSD1–CoREST complex (10–12 mg/mL concentration in 25 mM HEPES-Na, pH 7.4, 100 mM NaCl, 5 mM DTT, 1 mM PMSF) was mixed with the reservoir solution [0.60 M Li2SO4, 0.63 M (NH4)2SO4, 0.25 M NaCl, 100 mM Na-citrate, pH 5.6, 10 mM DTT]. The crystals belong to the orthorhombic I222 space group (\(a = 123.86\) Å, \(b = 179.37\) Å, \(c = 235.05\) Å).

A 5′-UUAGG-3′ RNA ligand was introduced into the crystals by the soaking method (Hassell et al. 2007; Reiter et al. 2010). Other oligonucleotides were tested in crystal soaking trials and include: the 53- and 25-nt GQ RNA, as well as UUAGG, UUUUU, CCCUUAA, UUAGGAG, AAAAAAGCAA, and GGUUUUUCUUUU RNAs. The 53- and 25-nt GQ RNAs proved to be unsuccessful in crystal soaking trials due to the size of the RNAs and because the reservoir solution contained high salt or high concentrations of Li\(^+\) ions that inhibit GQ formation (Hardin et al. 2000). Aside from UUAGG and UUAGGAG soaked crystals, no other single-stranded RNAs yielded any noticeable difference electron density upon careful analysis of the Fourier map ([\(F_o\]−[\(F_c\)]). The LSD1–CoREST complex crystals were incubated for 3 h or overnight in the cryoprotectant containing UUAGG RNA [0.76 M Li2SO4, 0.74 M (NH4)2SO4, 0.35 M NaCl, 100 mM Na-citrate, pH 5.6, 23% (v/v) glycerol and 2–5 mM RNA]. After soaking, crystals were harvested and flash-cooled in liquid nitrogen.

**X-ray data collection**

Diffraction data were collected at 100 K at LS-CAT beamline 21 (F Hutch), Advanced Photon Source, Argonne National Laboratory, using a MarMosaic 225 CCD detector. Data were processed and scaled using the HKL2000 package (Otwinowski and Minor 1997). Data collection and data processing statistics are summarized in Supplemental Table 1.
Crystal structure determination

Molecular replacement was applied to locate a solution using PHENIX and a previously determined structure of the LSD1–CoREST complex (PDB 2IW5) (Afonine et al. 2012). The difference Fourier map (|Fo|−|Fc|) revealed the presence of RNA and clear difference density enabled us to unambiguously determine the directionality of four nucleotides of the RNA (Fig. 5). While a 3’ phosphate of G5 is observed, the 3’ G nucleobase is not observed in the density. Coot was used for model building throughout the refinement and the RNA was built using RCrane (Emsley and Cowtan 2004; Keating and Pyle 2012). The final model consists of residues 171–836 of LSD1, residues 308–440 of CoREST, one FAD molecule, one RNA molecule, 5 sulfate ions, 1 glycerol molecule, and 23 water molecules. Refinement statistics are listed in Supplemental Table 1 (PDB ID 4XBF).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Conrad Wagner for enthusiastic discussions, M. Ascano, M. Egli, F.P. Guengerich, and C. Sanders for critical reading of the manuscript, J. Harp for crystallographic assistance, H. MacDonald for Structural Biology, and the Vanderbilt Department of Biochemistry. N.J.R. acknowledges ACS-IRG (58-0009-54) and AHA for Structural Biology, and the Vanderbilt Department of Biochemistry. N.J.R. performed the X-ray crystallographic studies. All authors contributed to writing the manuscript.

Received May 2, 2016; accepted May 5, 2016.

REFERENCES


www.majournal.org
Hirsch et al.


G-quadruplex RNA binding and recognition by the lysine-specific histone demethylase-1 enzyme


RNA 2016 22: 1250-1260 originally published online June 8, 2016
Access the most recent version at doi:10.1261/rna.057265.116

Supplemental Material
http://rnajournal.cshlp.org/content/suppl/2016/06/07/rna.057265.116.DC1.html

References
This article cites 65 articles, 21 of which can be accessed free at:
http://rnajournal.cshlp.org/content/22/8/1250.full.html#ref-list-1

Creative Commons License
This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see http://rnajournal.cshlp.org/site/misc/terms.xhtml). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.