5706–5717 Nucleic Acids Research, 2012, Vol. 40, No. 12 Published online 13 March 2012
doi:10.1093/nar/gks212

Analysis of lysine recognition and specificity of the *Bacillus subtilis* L box riboswitch

Sharnise N. Wilson-Mitchell[1,2], Frank J. Grundy[1,3] and Tina M. Henkin[1,2,3,*]

[1]Center for RNA Biology, [2]The Ohio State Biochemistry Program and [3]Department of Microbiology, the Ohio State University, Columbus, OH 43210, USA

Received December 2, 2011; Revised February 15, 2012; Accepted February 16, 2012

ABSTRACT

The ever-changing environment of a bacterial cell requires sophisticated mechanisms to adjust gene expression in response to changes in nutrient availability. L box riboswitch RNAs regulate gene expression in response to cellular lysine (lys) concentrations in the absence of additional regulatory factors. In *Bacillus subtilis*, binding of lysine (lys) to the L box RNA causes premature transcription termination in the leader region upstream of the *lysC* coding sequence. To date, little is known about the specific RNA–lys interactions required for transcription termination. In this study, we characterize features of the *B. subtilis lysC* leader RNA responsible for lys specificity, and structural elements of the lys molecule required for recognition. The wild-type *lysC* leader RNA can recognize and discriminate between lys and lys analogs. We identified leader RNA variants with mutations in the lys-binding pocket that exhibit changes in the specificity of ligand recognition. These data demonstrate that *lysC* leader RNA specificity is the result of recognition of ligand features through a series of distinct interactions between lys and nucleotides that comprise the lys-binding pocket, and provide insight into the molecular mechanisms employed by L box riboswitch RNAs to bind and recognize lys.

INTRODUCTION

Recent work has uncovered a set of bacterial regulatory RNAs, designated riboswitches, in which conserved *cis*–acting RNA elements located in the 'leader region' between the promoter and the regulated coding sequence respond to a variety of physiological signals without additional proteins or co-factors (1). Typical riboswitches are composed of two domains: the aptamer or ligand–sensing domain and the expression platform (2). The aptamer binds the cognate ligand and initiates a series of structural rearrangements that affect the fate of the downstream gene through structural modulation of the expression platform. Regulation of gene expression by these conserved leader RNAs can occur by transcription attenuation, modulation of translation initiation and effects on RNA stability (2,3).

A large number of genes that control essential processes are under riboswitch regulation in *Bacillus subtilis* (1). One of the earliest identified riboswitches, designated the L box, regulates transcription of the *lysC* gene in response to lysine (lys) (4–6). L box leader RNAs have been found upstream of a variety of genes involved in lys biosynthesis, transport and utilization (6–8). Lys biosynthesis in most bacteria and some plants occurs through the diaminopimelate pathway. The first step of lys biosynthesis is the phosphorylation of aspartate by aspartokinase. There are three differentially regulated aspartokinase isozymes in *Escherichia coli* and *B. subtilis*. However, only aspartokinase II, the product of the *lysC* gene, responds to lys (9). Formation of aspartate-4-phosphate is critical, as it is also a precursor for methionine and threonine biosynthesis (9). Other L box-regulated genes synthesize metabolites in the lys pathway that are required for other essential physiological processes, such as synthesis of cell wall peptidoglycan (10) and formation of bacterial endospores (11). Additional phylogenetic analyses suggest that L box leader RNAs are also involved in regulation of genes responsible for lys transport and the regulation of cellular pH (6,8).

When lys concentrations are high, lys binds to the *B. subtilis lysC* leader RNA and promotes the formation of an intrinsic terminator that prevents transcription of the *lysC* coding region (4,5). Low lys concentrations result in the formation of an alternate antiterminator structure that allows transcription to continue (Figure 1). In contrast, the *E. coli lysC* leader RNA is predicted to regulate expression at the translational level through the formation of a Shine–Dalgarno sequestering helix when lys concentrations are high (4,6,9). The lys analog aminoethylcysteine (AEC) is 10-fold less effective

*To whom correspondence should be addressed. Tel: +1 614 688 3831; Fax: +1 614 292 8120; Email: henkin.3@osu.edu

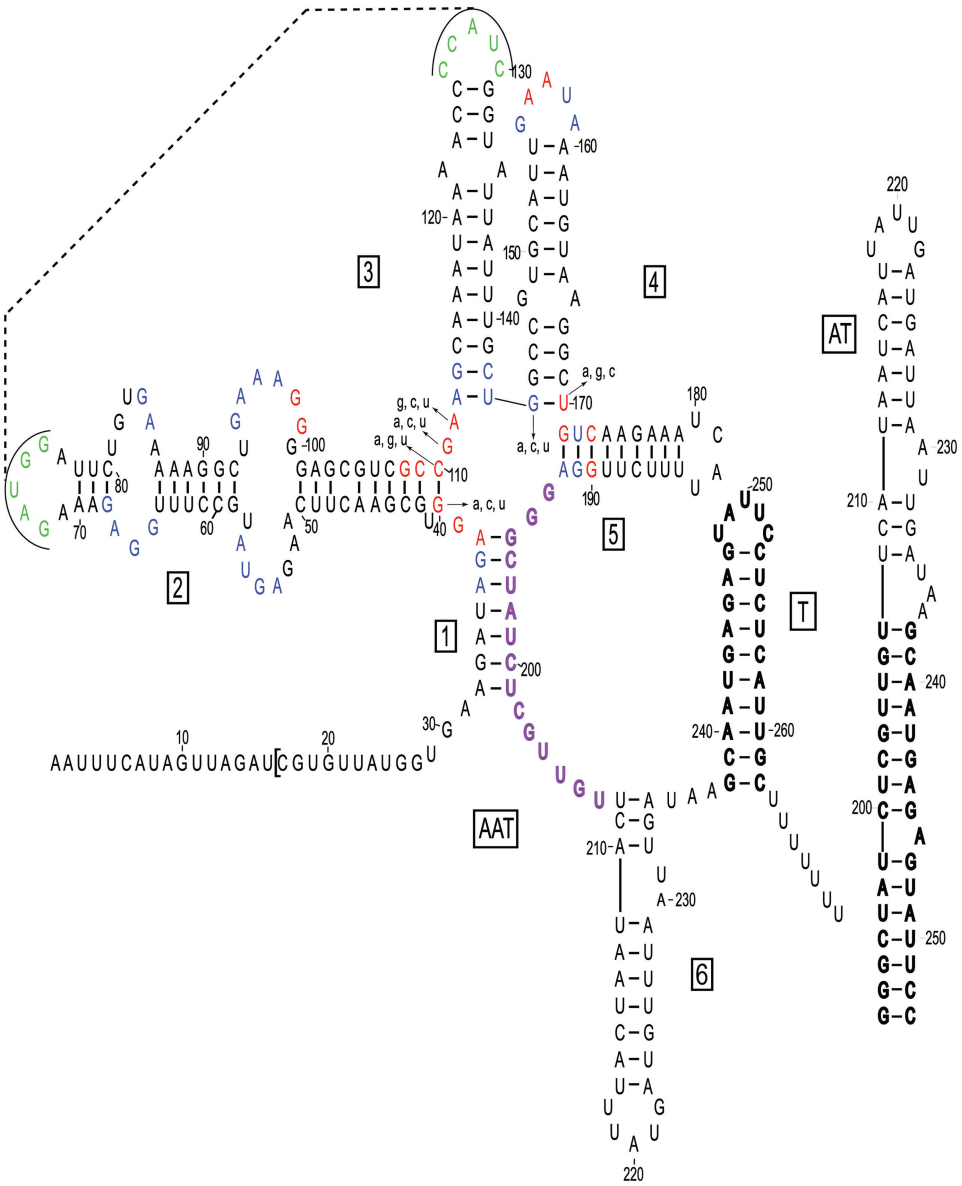

Figure 1. Sequence and predicted secondary structural model of the *B. subtilis lysC* leader RNA. Boxed letters and numbers indicate the terminator (T), antiterminator (AT), anti-antiterminator (AAT) and helices 1–6. Residues shown in red are found in 100% of the sequences analyzed and blue residues are found in >80% (4). Green residues connected by dashed lines indicate sequences that show extensive covariation (4) and form a tertiary interaction in the crystal structure (23). Arrows identify the *lysC* variants constructed for this study. Bold nucleotides indicate residues shared between the terminator and antiterminator elements. Purple nucleotides are complementary to the DNA oligonucleotide probe used in RNase H cleavage assays.

than lys in promotion of *B. subtilis lysC* termination *in vitro* (4). AEC differs from lys by a single substitution of carbon with sulfur (Figure 3); the observed differential sensitivity demonstrates specific lys recognition by the *lysC* leader RNA (4). Lys analog-binding studies using the aptamer region of the *lysC* leader RNA further suggested that leader RNA binding is specific for lys (12).

Previous studies of riboswitch specificity identified several unique features adopted by riboswitch RNAs to recognize their cognate ligands. For example, the purine responsive riboswitches use canonical Watson–Crick base pairing for recognition (13). A single mutation of a conserved base in the guanine responsive riboswitch results in a specificity switch that allows recognition of adenine (14). Natural variants of the guanine and adenine riboswitches were identified that selectively bind 2′-deoxyguanosine (dG) and discriminate against guanine (15). The discovery of these variants demonstrates that riboswitch RNAs can use similar RNA architectures to recognize different molecules.

Several riboswitches use similar mechanisms to recognize nucleotide-like molecules. In each example, the RNA uses common features such as the face of the nucleotide to assist in recognition. Recognition of *S*-adenosylmethionine

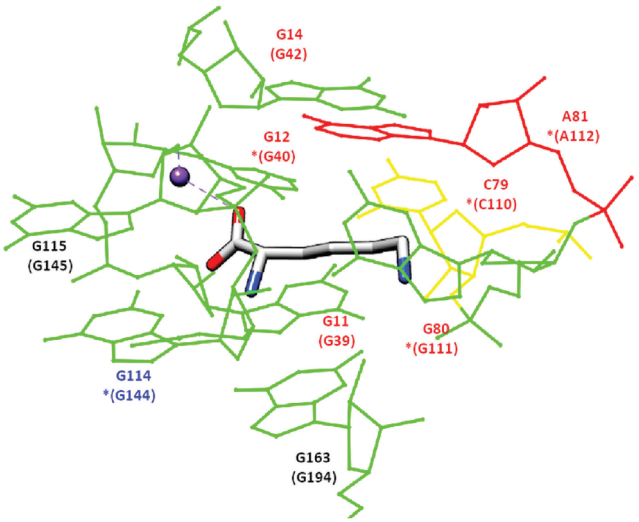

Figure 2. Tertiary structure of the *Thermatoga maritima* L box lys binding pocket. Lys-binding pocket adapted from Serganov *et al.* (23). Nucleotides shown are invariant among all known L box leaders with the exception of G114 (G144 in the *B. subtilis* sequence) which is >80% conserved. Corresponding nucleotides in *B. subtilis lysC* are shown in parentheses. Positions mutated for this study are identified with an asterisk. The purple sphere represents a potassium ion.

(SAM) by the S box (SAM-I) riboswitch occurs through a base triple interaction that includes the adenine ring of the SAM molecule (16). Similarly, the aminopyrimidine ring of the thiamine pyrophosphate (TPP) molecule is recognized by the THI-box RNA through molecular stacking with conserved residues (17). The nucleotide face of the cyclic di-GMP (c-di-GMP) molecule is also recognized using asymmetric RNA interactions (18). Although the RNAs recognize similar features of different ligands, each riboswitch utilizes a different mechanism for recognition.

Another common mechanism for ligand recognition incorporates divalent metal ions for both binding and recognition of negatively charged groups. RNAs that have been shown to use this method include the S box (16), THI-box (17), glycine riboswitch (19) and the *glmS* ribozyme (20). Analysis of glycine recognition by the *Vibrio cholerae* riboswitch suggests that Mg^{2+} assisted recognition of the carboxylate group contributes to glycine specificity (20). In addition, glycine analog binding and crystal structure analyses indicate that shape complementarity between glycine and the binding pocket is essential for glycine selectivity (20). The binding pocket of the leader RNA is precisely structured to only accommodate glycine, whereas the RNA did not bind other small amino acids (21). Recognition of amino acids by riboswitch RNAs presents an interesting challenge, as the RNA must discriminate against similar molecules found in the cell.

Crystal structure analysis of an L box leader RNA from *Thermatoga maritima* revealed a complex tertiary structure (22,23). The lys molecule is completely encapsulated by the RNA and is buried within a lattice formed from conserved core residues (22,23). Recent analysis of the *lysC* leader RNA has highlighted the importance of peripheral elements and long-range interactions in the organization of the lys–binding pocket; the significance of peripheral elements has also been demonstrated for the S-box riboswitch (24,25). Residues in direct contact with lys or in the binding pocket are the most highly conserved. Additionally, a K^+ ion was buried between lys and the RNA (Figure 2) (23). This arrangement is similar to the Mg^{2+} ions found in the glycine-binding pocket of the glycine riboswitch (20). Currently, the role of K^+ in the binding and function of L box leader RNAs has not been established, but it was suggested that it might assist in lys recognition (23).

With the variety of genes under L box control, it is essential that the RNA selectively binds lys and efficiently regulates expression under appropriate conditions. The current structural model of the lys-binding site does not clearly demonstrate how the RNA can specifically recognize lys and discriminate against substitutions in the ligand. Previous studies of lys analog recognition focused on a subset of available compounds (12). In this study, we test a broad range of lys analogs and identify mutations that affect ligand recognition; these mutations reveal molecular interactions responsible for recognition and specificity. We also investigate the role of K^+ in the binding site and the potential effects on leader RNA function.

MATERIALS AND METHODS

Construction of DNA templates

Templates for *in vitro* transcription by *B. subtilis* RNA polymerase (RNAP) were generated by fusion of the *B. subtilis glyQS* promoter sequence to the *lysC* leader RNA sequence. PCR was carried out using DNA oligonucleotide primers that included the *glyQS* promoter and hybridized within the *lysC* leader region. A transcriptional fusion plasmid, pFG328, with a wild-type *lysC* insertion (4), was the DNA source for PCR amplification using *Taq* DNA polymerase as per the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). The resulting PCR fragment contained the *lysC* leader template downstream of the *glyQS* promoter such that transcription would start at +17 relative to the *lysC* transcription start-site and included *lysC* sequence extending to 14 nt downstream of the leader region termination site. The transcription start site was designed to allow transcription initiation with the dinucleotide ApC and a halt at position G42 (relative to the native *lysC* transcript) during transcription in the absence of cytidine 5′-triphosphate (CTP). The 3′-end of the construct included an additional 106 bp of random sequence to allow resolution between terminated (252 nt) and readthrough (372 nt) transcripts. PCR products were purified using a QIAquick PCR clean up kit (Qiagen, Chatsworth, CA, USA) and sequenced by Genewiz (South Plainfield, NJ, USA).

DNA templates used for T7 RNAP transcription were generated using complementary pairs of overlapping DNA oligonucleotides as previously described (25,26).

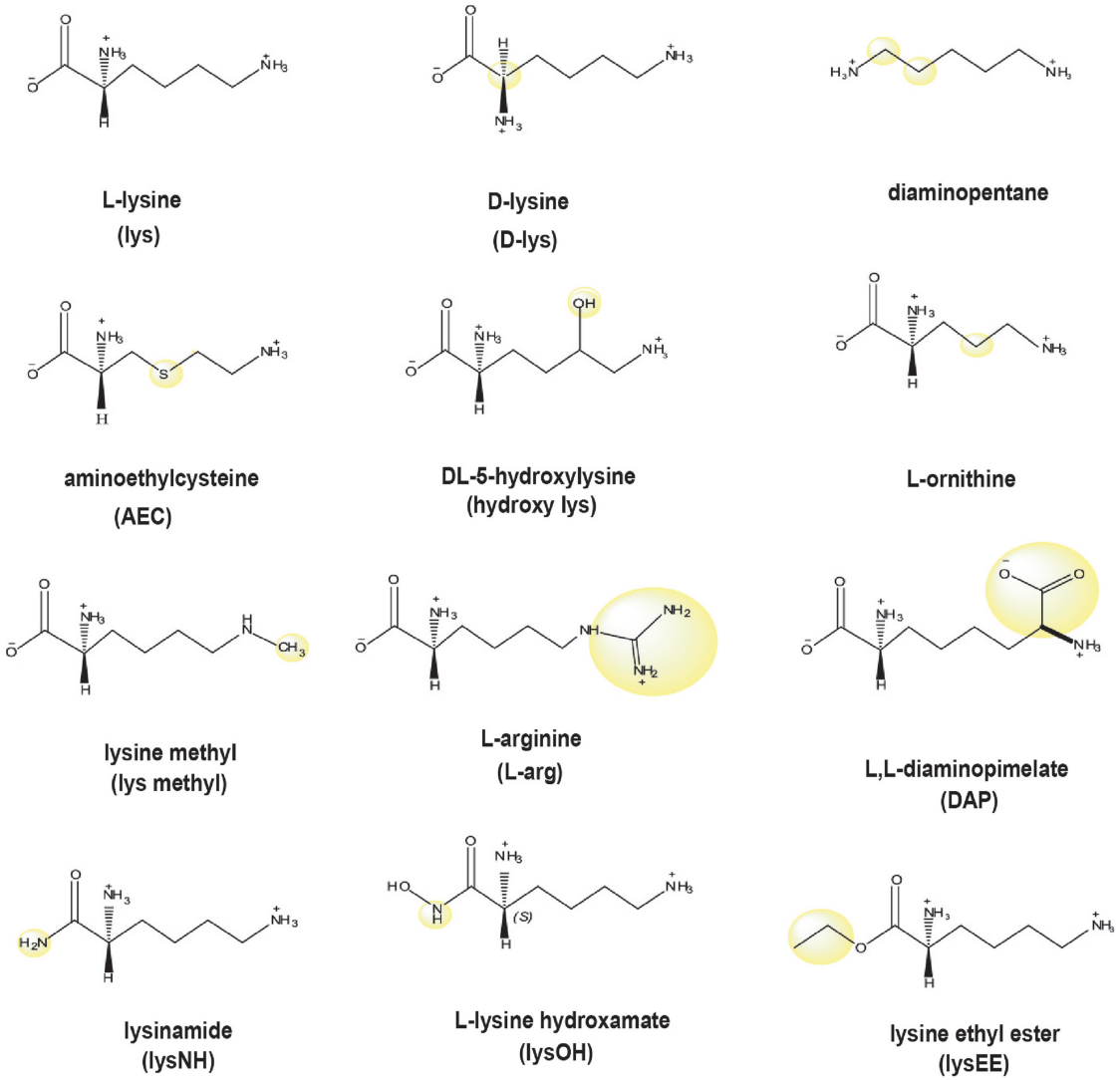

Figure 3. Chemical structures of lys and closely related lys analogs. Representations of each lys analog used in this study are depicted with molecular changes highlighted in yellow. Abbreviated names referred to in the text are in parentheses.

Briefly, the 5′ pair contained the phage T7 RNAP promoter sequence fused to the +18 position of the *lysC* leader region. The remaining complementary pairs contained the wild-type *lysC* sequence, each with a 5-nt 3′ overhang complementary to the 5′ region of the adjacent pair. The terminal pair contained the 3′ leader RNA sequence that ended at +234, 3 nt prior to the 5′ position of the transcription terminator. This position was selected so that helix 1 formation could be monitored without the competing antiterminator structure. Each internal oligonucleotide pair was phosphorylated using T4 polynucleotide kinase (New England Biolabs, Beverly, MA, USA). The pairs were then mixed, incubated at 95°C and slow cooled to room temperature for annealing. The paired oligonucleotides were ligated using T4 DNA ligase per the manufacturer's instructions (New England Biolabs, Beverly, MA, USA). The resulting DNA template was amplified using the flanking 5′ and 3′DNA oligonucleotides as primers for PCR using *Taq* DNA polymerase (Invitrogen, Carlsbad, CA, USA). The final template DNA was purified and sequenced as described above.

Site-directed mutagenesis

The transcriptional fusion vector pFG328 containing the wild-type *lysC* DNA template (4) was used as a template for oligonucleotide-directed mutagenesis. DNA oligonucleotides containing the desired mutations were designed as primers for PCR amplification of pFG328-*lysC* DNA using *Pfu* DNA polymerase (Stratagene, La Jolla, CA, USA). The resulting products were subjected to digestion with DpnI (New England Biolabs, Beverly, MA, USA) to remove the starting wild-type template and introduced into XL-2 blue ultracompetent cells by transformation as per the manufacturer's instructions (Stratagene, La Jolla, CA, USA). The plasmid DNA was isolated using a Promega Wizard prep kit and sequenced to confirm the mutations (Genewiz, South

Plainfield, NJ, USA). The final constructs contained the *glyQS* promoter sequence upstream of the *lysC* leader RNA with the desired mutations, and were used as the DNA template for *in vitro* transcription.

The DNA templates for transcription by the T7 phage RNAP contained the wild-type *lysC* leader RNA sequence downstream of the T7 RNAP promoter sequence. Mutation of the *lysC* construct was performed by the substitution of complementary oligonucleotide pairs containing the desired mutations. The wild-type and mutant pairs were phosphorylated, ligated and amplified by PCR using *Taq* DNA polymerase (Invitrogen, Carlsbad, CA, USA). The resulting DNA templates with the appropriate mutations were used in T7 RNAP transcription.

In vitro transcription termination assays

Single round transcription of the wild-type *lysC* leader RNA template was performed as previously described by Grundy *et al*. (27). A reaction mixture of 20 mM Tris–HCl, pH 8, 20 mM NaCl, 10 mM $MgCl_2$, 100 μM EDTA, 150 μM ApC (Sigma), 2.5 μM GTP and ATP, 0.75 μM UTP, [α-^{32}P]-UTP (GE Healthcare; 0.25 μM, 800 Ci/mmol [30 TBq/mmol]), DNA template (10 nM) and His-tagged purified *B. subtilis* RNAP (6 nM) (28) was incubated for 15 min at 37°C. CTP was excluded from the reaction mixture to generate a transcriptional halt at +42 (relative to the native *lysC* transcription start site). After incubation, Heparin was added to block reinitiation and synchronize transcript synthesis. KCl (27 mM) was added where indicated. Lys or lys analogs were added, and elongation was resumed by the addition of 10 μM rNTPs and 40 mM $MgCl_2$ followed by incubation for 15 min at 37°C. The transcription reactions were stopped by phenol/chloroform extraction, and the products were resolved by denaturing PAGE and visualized by PhosphorImager analysis (Molecular Dynamics). Percent termination represents terminated transcripts relative to the total amount of RNA transcribed; this value was plotted as a function of ligand concentration. Non-linear regression analysis was used to determine the concentration required for half-maximal termination ($Trm_{1/2}$) (SigmaPlot, Systat Software). Reactions were performed in duplicate and reproducibility was ± 5%.

RNase H cleavage

RNase H cleavage was carried out as described by Grundy *et al*. (4). The DNA template for T7 RNAP transcription was generated as described above. Transcription was carried out using a MEGAshortscript T7 RNAP transcription kit (Ambion, Austin, TX, USA) in the presence of 0.5 μM [α-^{32}P]-UTP (GE Healthcare; 800 Ci/mmol [30 TBq/mmol]) for radiolabeling. RNAs were transcribed in the presence or absence of 7.5 mM lys or lys analog, as indicated, for 30 min at 37°C. After transcription, a DNA oligonucleotide (47 μM) complementary to positions 193–204 of the *B. subtilis lysC* leader RNA was added and hybridized for 5 min at 37°C. RNase H (0.45 U; Ambion, Austin, TX, USA) was then added to the reaction and incubated for 10 min at 37°C. The reactions were stopped by phenol–chloroform extraction. The resulting RNA products were resolved by denaturing PAGE and visualized by PhosphorImager analysis (Molecular Dynamics). The percentage of RNAs protected from cleavage was calculated as the amount of full-length RNAs relative to the total amount of RNA transcribed in each reaction. The reactions were performed in duplicate and reproducibility was ± 10%.

Table 1. *In vitro* transcription termination of wild-type *lysC* leader in the presence of lys or lys analogs

In vitro termination				
Bacillus subtilis lysC leader RNA				
Lysine modification	Ligand	Wild-type		
		Term.[a]	Ratio[b]	$Trm_{½}$ (mM)[c]
	None	31		
	Lys	96	3.2	0.17 ± 0.02
Global	D-lys	46	1.5	>13
	Diaminopentane	43	1.4	>13
Alkyl chain	AEC	84	2.7	0.80 ± 0.11
	Hydroxy lys	35	1.1	>13
	L-ornithine	32	1.0	ND[d]
ε-amino group	Lys methyl	73	2.4	1.3 ± 0.40
	L-arg	35	1.1	ND
	DAP	30	0.97	>13
Carboxyl group	lysNH	90	2.9	0.30 ± 0.02
	lysOH	87	2.8	1.9 ± 0.34
	lysEE	76	2.5	1.0 ± 0.25

[a]Termination (term.) is the percent of RNAs terminated in the presence of 10 mM ligand relative to the total amount of RNA transcribed.
[b]Ratio is the ratio of percent terminated transcripts in the presence of 10 mM ligand to the percent terminated transcripts in the absence of ligand.
[c]$Trm_{1/2}$ is the concentration of ligand required to promote 50% termination.
[d]ND, not determined.

RESULTS

Wild-type *lysC* transcription termination in response to lys analogs

In vitro transcription of the wild-type *lysC* leader by *B. subtilis* RNAP in the presence of lys or lys analogs (Table 1) was used to identify features of the lys molecule required for ligand-dependent transcription termination. Lys and lys analogs with the ability to promote termination of the *lysC* leader RNA were further tested *in vitro* using a broad range of ligand concentrations (0–12.8 mM) (Table 1). Transcription of the *lysC* leader region in the presence of lys (10 mM) resulted in 96% termination and the concentration required for $Trm_{1/2}$ was 0.17 mM. *In vitro* transcription in the presence of 11 different lys analogs (added at 10 mM final concentration) showed that alterations to the main chain functional groups, as seen in D-lys and diaminopentane, resulted in a reduction of ligand-dependent termination (Table 1). The $Trm_{1/2}$ for both D-lys and diaminopentane was >13 mM. These results suggest that changes to functional groups that participate in RNA–ligand interactions are

detrimental to ligand-dependent termination. Reduced termination in the presence of AEC (84%) suggests that the sulfur substitution in the side chain of the molecule causes reduced ligand recognition. The $Trm_{1/2}$ in the presence of AEC (0.8 mM) increased 5-fold as compared to lys. The increase in $Trm_{1/2}$ in the presence of AEC as compared to lys demonstrates sensitivity to the presence of the sulfur molecule in AEC. Transcription of the leader RNA in the presence of ligands with more complex functional groups, as seen in L-arginine (L-arg) and L,L-diaminopimelate (DAP), resulted in a loss of recognition, possibly because the changes result in a molecule too large to fit in the lys-binding pocket.

Alteration to the global structure of the molecule, such as the shape of the molecule (hydroxy lys), and the length of the side chain (L-ornithine) caused reduced termination efficiency (Table 1). Modification of the ε-amino group was tolerated only in the context of the addition of a methyl group, as in lysine methyl (lys methyl). Transcription in the presence of lys methyl resulted in a $Trm_{1/2}$ of 1.3 mM, an 8-fold increase as compared to lys. The increase in $Trm_{1/2}$ in the presence of lys methyl suggests that the methyl addition to the ε-amino group is deleterious to ligand recognition. The $Trm_{1/2}$ in the presence of lys methyl (1.3 mM) as compared to AEC (0.8 mM) suggests that the sulfur substitution found in AEC is less detrimental to ligand-dependent termination than the methyl addition to the ε-amino group (Table 1). These results may reflect similarities between the *B. subtilis* and *T. maritima* lys-binding pockets. In the *T. maritima* structure, the RNA participated in fewer contacts with the side chain as compared to the ε-amino group (22,23). Additionally, analogs with modifications to the carboxyl group resulted in >70% termination. This is consistent with the current structural model of the binding pocket, which indicates the presence of small channels near both the carboxyl and ε-amino groups of lys (23).

The carboxy-modified ligands were more effective than analogs with ε-amino or side chain modifications and exhibited lower $Trm_{1/2}$ concentrations (Table 1). Transcription in the presence of lysinamide (lysNH), which contains a nitrogen substitution of the hydroxyl carbon, resulted in a $Trm_{1/2}$ of 0.3 mM, a 2-fold increase as compared to lys. In the presence of lysNH, the leader RNA exhibited a smaller increase in $Trm_{1/2}$ as compared to lys methyl and AEC; this may be indicative of a larger or more stable network of RNA–ligand interactions near the carboxyl group. Substitution of the amide (lysNH) with an ethyl ester (lysEE) caused a 3-fold increase in $Trm_{1/2}$ (1.0 mM) as compared to lysNH, and a 6-fold increase relative to lys. The ethyl ester extension of the oxygen may disrupt other essential interactions required for lys-dependent termination. The observed decrease in $Trm_{1/2}$ in the presence of lysNH as compared to lysEE may reflect a direct interaction between the leader RNA and the substituted amide. Transcription in the presence of lysine hydroxamate (lysOH) resulted in a 10-fold increase in $Trm_{1/2}$ (1.9 mM) relative to lys. Similar to the other carboxy-modified analogs, lysOH contains a hydroxy amide substitution of the hydroxyl carbon (Figure 3). The increase in $Trm_{1/2}$ of lysOH as compared to lysNH suggests that the addition of a hydroxyl group causes a reduction in ligand-dependent termination. The observed increase in $Trm_{1/2}$ of lysOH as compared to lysEE emphasizes the importance of functional group recognition as compared to minor alterations in ligand length. This indicates that both the identity of the molecular substitution and additional interactions within the binding pocket affect ligand recognition.

Effects of substitutions at G111 and A112

In an effort to identify sequence elements within the *lysC* leader that are important for ligand recognition, we used site-directed mutagenesis to probe RNA requirements for lys-dependent termination. The nucleotides chosen for mutation are highly conserved (Figure 1) and predicted to interact with lys (Figure 2). Universally conserved G111 and A112 are positioned nearest to the ε-amino group in the binding pocket (Figure 2). Substitution of G111 with C resulted in reduced termination (12%) in the absence of ligand as compared to wild-type (Table 1; 31%) and caused a reduction in sensitivity to lys and AEC (Table 2). Each G111 variant exhibited a reduction in termination in the presence of lys and AEC. However, the G111A substitution resulted in a 2-fold increase in termination in response to DAP (Table 2), which is inactive in the context of the wild-type *lysC* sequence (Table 1). Similarly, the G111U substitution resulted in a 5-fold increase in termination in response to lys methyl, as compared to the no ligand control (Table 2). The reduction in lys-dependent termination suggests that G111 plays a pivotal role in lys recognition and/or binding. However, the termination response in the presence of DAP (G111A) and lys methyl (G111U) suggests that G111 is not required for termination. The sensitivity of G111A and G111U to ε-amino-modified analogs is in good agreement with the published crystal structure, as G111 is positioned nearest the ε-amino group of the lys molecule (23).

Unlike G111, in which no substitution allowed normal function, A112 was amenable to substitution with a guanine residue, exhibiting a ligand response similar to that of wild-type, although termination in the presence of ε-amino and carboxyl modified analogs was reduced (Tables 1 and 2). When A112 was substituted with either pyrimidine base (Table 2), there was a decrease in termination in both the presence and absence of ligand (~7% for A112C and ~19% for A112U). Both A112 and G111 were intolerant to pyrimidine substitutions, which suggests that purines are required for ligand recognition and lys-dependent termination. In contrast to A112G, the purine–purine substitution of G111 with A was deleterious to the function of the RNA. This suggests that there are specific interactions between G111 and the lys molecule, which directly contribute to lys-dependent termination.

Mutations near the lys carboxyl group result in altered ligand specificity

Lys-binding pocket variants were generated to assess the interaction between the leader RNA and the carboxyl

Table 2. *In vitro* transcription termination of G111 and A112 *lysC* leader variants in the presence of lys or lys analogs

In vitro termination													
Bacillus subtilis lysC G111 and A112 variants													
Lysine modification	Ligand	Variant											
		G111C		G111A		G111U		A112G		A112C		A112U	
		Term.[a]	Ratio[b]	Term.	Ratio	Term.	Ratio	Term.	Ratio	Term.	Ratio	Term.	Ratio
	None	12		14		10		42		6.8		19	
	Lys	13	1.1	14	1.0	9.1	0.89	98	2.3	5.0	0.74	20	1.1
Global	D-lys	16	1.3	14	1.0	14	1.4	51	1.2	8.8	1.3	23	1.3
	Diaminopentane	15	1.2	13	0.92	7.7	0.75	38	0.92	6.7	1.0	21	1.2
Alkyl chain	AEC	12	1.0	13	0.92	10	1.0	87	2.1	3.4	0.50	21	1.2
	Hydroxy lys	13	1.1	12	0.88	11	1.1	37	0.88	5.5	0.81	19	1.1
	L-ornithine	13	1.1	13	0.92	8.5	0.85	34	0.81	4.5	0.66	19	1.1
ε-amino group	Lys methyl	13	1.1	13	0.90	51	5.1	69	1.6	5.9	0.87	20	1.1
	L-arg	14	1.2	13	0.92	11	1.1	38	0.90	7.0	1.0	20	1.1
	DAP	14	1.2	26	1.9	8.8	0.88	42	1.0	2.3	0.34	21	1.2
Carboxyl group	lysNH	13	1.1	12	0.87	11	1.1	69	1.6	4.4	0.65	18	0.9
	lysOH	13	1.1	13	0.90	10	1.0	65	1.5	4.3	0.63	18	0.9
	lysEE	16	1.3	12	0.85	11	1.1	51	1.2	6.7	1.0	23	1.2

[a]Termination (term.) is the percent of RNAs terminated relative to the total transcription.
[b]Ratio is the ratio of percent terminated transcripts in the presence of 10 mM ligand to the percent terminated transcripts in the absence of ligand.

region of the lys molecule. Crystal structure analysis revealed base pairing between G40 and C110 of the leader RNA, which is in agreement with previous studies that identified mutations at this position that result in deregulation of the *lysC* leader and cause AEC resistance (29). The conserved G40–C110 and G144–U170 base pairs were substituted with other canonical base pairs and tested for termination in the presence of lys or lys analogs. The G40U–C110A substitution inhibited lys-dependent termination (Table 3). Lys or lys analogs with substitutions in the side chain (AEC) or ε-amino group were not recognized by the G40A–C110U variant (Table 3). However, termination of the G40A–C110U variant was observed in the presence of carboxyl-modified analogs. Transcription of the G40A–C110U variant in the presence of lysNH, which replaces a carboxyl oxygen with an amide, resulted in a reduction in termination (Table 3) as compared to wild-type (Table 1). Termination of the G40A–C110U variant increased 4-fold in the presence of lysOH (75%) as compared to the no ligand control (Table 3). The $Trm_{1/2}$ of the G40A–C110U variant in the presence of lysOH was 2.1 mM as compared to >13 mM for lys (Table 4). The $Trm_{1/2}$ exhibited by the G40A–C110U variant in the presence of lysOH was similar to wild-type; however, the $Trm_{1/2}$ was increased, as compared to wild-type, in the presence of lys (Table 4). The response to lysOH and lack of response to lys suggests that the G40A–C110U mutations cause a change in ligand recognition specificity. Termination of the G40A–C110U variant in the presence of lysNH, as compared to lysOH, indicates that the hydroxyl substitution at the carbonyl more effectively promotes termination than an amide at the same position. It is plausible that the change in specificity is due to the formation of new molecular interactions between G40A and the amide or hydroxyl group of the lys analog.

Table 3. *In vitro* transcription termination of G40–C110 *lysC* leader variants in the presence of 10 mM lys or lys analogs

In vitro termination					
Bacillus subtilis lysC G40–C110 variants					
Lysine modification	Ligand	Variant			
		U–A		A–U	
		Term.[a]	Ratio[b]	Term.	Ratio
	None	26		17	
	Lys	24	0.92	17	1.0
Global	D-lys	34	1.3	21	1.2
	Diaminopentane	25	0.96	16	0.94
Alkyl chain	AEC	22	0.85	14	0.82
	Hydroxy lys	24	0.92	13	0.76
	L-ornithine	22	0.85	12	0.71
ε-amino group	Lys methyl	29	1.1	21	1.2
	L-arg	34	1.3	13	0.8
	DAP	26	1.0	15	0.9
Carboxyl group	lysNH	30	1.2	49	2.9
	lysOH	33	1.3	75	4.4
	lysEE	24	0.92	15	0.9

[a]Termination (term.) is the percent of RNAs terminated relative to the total transcription.
[b]Ratio is the ratio of percent terminated transcripts in the presence of 10 mM ligand to the percent terminated transcripts in the absence of ligand.

The G144–U170 leader variants also exhibited specificity for carboxyl-modified analogs. Substitution of the wild-type G–U pair with U–G caused a loss of lys recognition and a reduction in termination in the presence of lys analogs (Table 5). The U170G mutation, which replaces the G144–U170 pair with a G–G pair, resulted in complete loss of ligand-dependent termination *in vitro*, suggesting

that base pairing is required at this position (Table 5). Variants that maintained base pairing between positions 144 and 170 showed an increase in termination in the presence of lys methyl, lysNH and lysEE. With the exception of the A–U and C–G constructs, all other variants resulted in termination efficiencies lower than that of wild-type in the presence of lys methyl, lysNH and lysEE (Tables 1 and 5). The C–G variant (G144C–U170G) showed a 14% increase in termination in the absence of ligand (Table 5) as compared to wild-type (Table 1), which suggests that the nucleotide identity at position 144 affects the unliganded RNA structure (C–G). In contrast to the C–G variant, the G–C variant (G144–U170C) resulted in a reduction of ligand-dependent termination as compared to wild-type in the absence of ligand (G–C) (Tables 1 and 5). Termination of the G–C variant in the presence of lys and lys analogs was reduced as compared to wild-type. These results indicate that the U170C mutation does not abolish function of the RNA but perturbs ligand recognition. Transcription of the G144A–U170 variant in the presence of lysEE resulted in increased termination (94%) (Table 5; A–U) as compared to wild-type termination in the presence of lysEE (76%) (Table 1). The $Trm_{1/2}$ of the G144A variant for lysEE was 0.85 mM, similar to that of wild-type (Table 4). However, $Trm_{1/2}$ of the G144A variant in the presence of lys was >13 mM, suggesting a loss in lys recognition (Table 4). The G144A substitution, a G–U to A–U base pair exchange, caused an alternate ligand specificity where the variant leader RNA responds to lysEE and discriminates against lys. Substitution of the G144A–U170 pair with U–A resulted in a reduction of termination efficiency in the presence of the carboxyl-modified analogs; however, ligand specificity was similar to that of the A–U variant. Recognition of lys methyl by the G144–U170 variants is surprising as these positions were not predicted to reside near the ε-amino group. One explanation is that the G144–U170 base pair is not primarily responsible for ε-amino group recognition. Termination in the presence of lys analogs as opposed to lys and AEC suggests that substitution of the G40–C110 and G144–U170 base pairs resulted in a loss of lys recognition. However, G144–U170 variants retained specificity for ε-amino and carboxyl-modified analogs while G40A–C110U exhibited specificity for analogs with carboxyl modifications. Together, these results indicate that specific mutations of the RNA sequence can result in alternate ligand specificity whereby the leader RNA responds to lys analogs while discriminating against lys.

Table 4. Termination efficiency of *lysC* leader variants in the presence of lys or lys analogs

	$Trm_{1/2}$ (mM)[a]		
	Bacillus subtilis lysC leader variants		
Ligand	Variant		
	A112G[b]	G40A–C110U[b]	G144A–U170[b]
lys	0.38 ± 0.07	>13	>13
AEC	1.2 ± 0.20	ND	ND
lysOH	ND	2.1 ± 0.20	ND
lysEE	ND	ND	0.85 ± 0.10

[a]$Trm_{1/2}$ is the concentration of ligand required to promote 50% termination.
[b]ND, not determined.

Table 5. *In vitro* transcription termination of G144–U170 *lysC* leader variants in the presence of 10 mM lys or lys analogs

		In vitro termination											
		Bacillus subtilis lysC G144–U170 variants											
Lysine modification	Ligand	Variant											
		U–G		G–G		C–G		G–C		A–U		U-A	
		Term.[a]	Ratio[b]	Term.	Rratio	Term.	Ratio	Term.	Ratio	Term.	Ratio	Term.	Ratio
	None	33		22		45		26		22		21	
	Lys	33	1.0	21	1.0	43	1.0	29	1.1	22	1.0	19	0.9
Global	D-lys	41	1.2	28	1.3	48	1.1	29	1.1	26	1.2	23	1.1
	Diaminopentane	43	1.3	23	1.0	69	1.5	35	1.3	38	1.7	25	1.2
Alkyl chain	AEC	30	0.9	19	0.9	41	0.9	24	0.9	23	1.0	18	0.9
	Hydroxy lys	38	1.2	23	1.0	51	1.1	24	0.9	30	1.4	18	0.9
	L-ornithine	28	0.8	21	1.0	40	0.9	24	0.9	22	1.0	17	0.8
ε-amino group	Lys methyl	50	1.5	24	1.1	92	2.0	70	2.7	68	3.1	66	3.1
	L-arg	30	0.9	23	1.0	40	0.9	21	0.8	22	1.0	18	0.9
	DAP	32	1.0	21	1.0	43	1.0	25	1.0	23	1.0	16	0.8
Carboxyl group	lysNH	49	1.5	24	1.1	87	1.9	57	2.2	73	3.3	63	3.0
	lysOH	47	1.4	20	0.9	67	1.5	34	1.3	41	1.9	36	1.7
	lysEE	72	2.1	24	1.1	91	2.0	49	1.9	94	4.2	72	3.4

[a]Termination (term.) is the percent of RNAs terminated relative to the total transcription.
[b]Ratio is the ratio of percent terminated transcripts in the presence of 10 mM ligand to the percent terminated transcripts in the absence of ligand.

K^+ decreases the concentration of lys required for transcription termination

The *B. subtilis lysC* leader RNA does not require K^+ for lys-dependent termination *in vitro* (4,5). However, recent analyses suggest that K^+ is important for lys recognition and affinity (7,30). We, therefore, tested the effect of addition of K^+ to our purified *in vitro* transcription system. In the presence of K^+, the wild-type leader RNA showed a 5-fold decrease in $Trm_{1/2}$ for lys (Fig. 4A). Other monovalent (NaCl) and divalent salts ($CaCl_2$) were also assayed to test whether the observed effect was specific to K^+ and no observable change in $Trm_{1/2}$ was detected (our unpublished results). The 5-fold decrease in $Trm_{1/2}$ for lys in the presence of K^+ suggests that K^+ plays a key role in lys-dependent termination by increasing the affinity of the RNA for lys.

Variant leader RNAs and lys analogs were further tested to determine if the K^+ is involved in ligand recognition. The A112G variant, which does not show altered ligand recognition but exhibited a 2-fold decrease in lys-dependent termination relative to the wild-type (Table 2), exhibited no observable change in lys sensitivity in response to K^+ (Figure 4A). The lack of response to the presence of K^+ may be the result of changes in the structure of the binding pocket imposed by the A112G mutation. Based on the position of K^+ in the structure, it is plausible that the G40–C110 or G144–U170 base pairs of the *B. subtilis lysC* binding pocket interact directly with K^+ (Figure 2). The G40A–C110U variant, which exhibited increased termination only in the presence of lysOH (Table 3), exhibited no change in $Trm_{1/2}$ for lysOH in the presence of K^+ (Figure 4B). Similarly, the G144A variant, which responds to lysEE, showed no reduction in $Trm_{1/2}$ for lysEE in the presence of K^+ (Figure 4B). Transcription of the G144A variant in the presence of lys demonstrated that K^+ was unable to enhance lys recognition (our unpublished results). These results indicate that K^+ does not play a role in the specificity of ligand recognition but rather affects affinity of the wild-type leader RNA for lys.

lysC leader RNA variants undergo ligand-dependent structural changes

Structural rearrangement of the leader RNA when the regulatory ligand is present is the hallmark of riboswitch regulation. This structural change is crucial as it directly controls gene expression. The current model of the *lysC* leader RNA structural transition indicates that formation of the transcription terminator requires the stabilization of the anti-antiterminator (AAT; Figure 1) through the formation of helix 1 (4). To monitor the structural change of the variant leader RNAs in response to lys analogs, we used RNase H probing of helix 1. Addition of a DNA oligonucleotide complementary to the 3′ side of helix 1 will result in cleavage only when helix 1 is unpaired; stabilization of the helix prevents binding of the oligonucleotide, which results in protection of the RNA from cleavage by RNase H, an endonuclease that is specific for RNA–DNA hybrids. In the absence of ligand, 31% of the wild-type RNAs were protected from RNase H cleavage as compared to 89% protection in the presence of lys (Figure 5, lanes 1 and 2). These results indicate that the presence of lys promotes the formation of helix 1. When wild-type RNAs were transcribed in the presence of high concentrations of AEC, 84% protection was observed (Figure 5, lane 3). Similarly, wild-type RNAs exhibited 82% protection in the presence of lysOH and 60% protection in the presence of lysEE (Figure 5, lanes 4 and 5). The reduced protection of wild-type helix 1 in the presence of lys analogs, as compared to lys, is consistent with lower wild-type termination *in vitro* in the presence of lys analogs. For the A112G variant, 48% of the transcribed RNAs were protected from cleavage in the no ligand control as compared to 65% in the presence of lys (Figure 5, lanes 6 and 7). A reduction in protection of the A112G variant, as compared to wild-type, was observed in the presence of AEC, lysOH, and lysEE (Figure 5, lanes 8–10). This demonstrates that protection

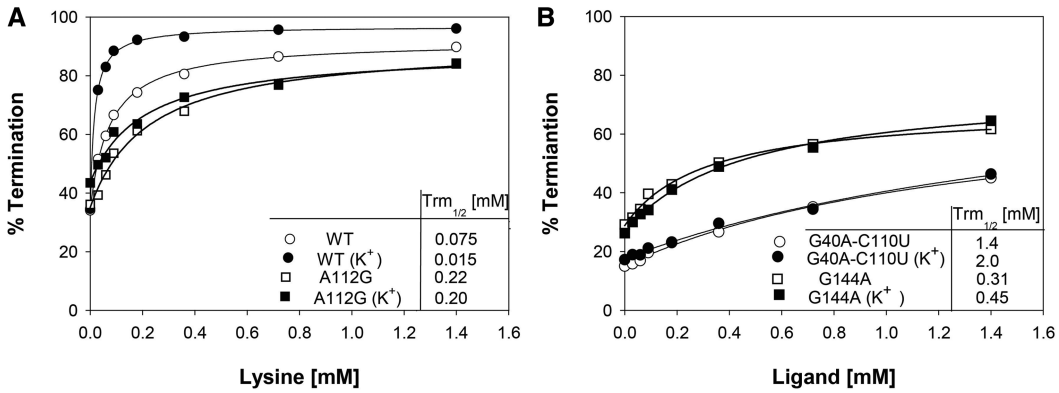

Figure 4. *In vitro* transcription of wild-type and variant *lysC* leader RNAs in the presence and absence of potassium. Wild-type or variant leader RNAs were transcribed in the presence or absence of 27 mM KCl. Open symbols represent transcription in the absence of KCl, filled symbols represent transcription in the presence KCl. (**A**) Wild-type (circles) and A112G (squares) templates transcribed in the presence of lys. (**B**) Transcription of the G40A-C110U variant in the presence of lysOH (circles) and the G144A-U170 variant transcribed in the presence of lysEE (squares).

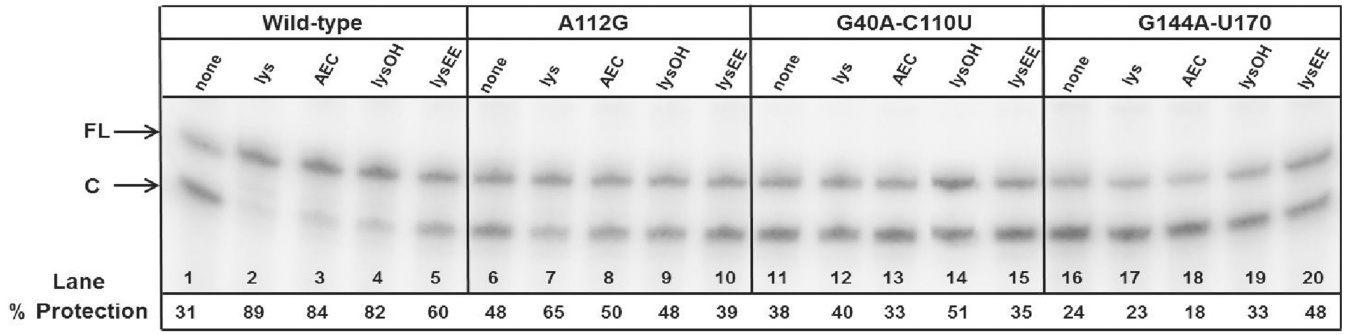

Figure 5. RNase H cleavage of wild-type and variant *lysC* leader RNAs. Radiolabeled RNAs were transcribed in the presence of 7.5 mM lys, AEC, lysOH and lysEE and incubated with a DNA oligonucleotide complimentary to the 3' side of helix 1 to allow hybridization followed by RNase H digestion. Helix 1 protection of wild-type in the presence of lys and lys analogs is shown in lanes 1–5; lanes 6–10, A112G variant; lanes 11–15, G40A–C110U variant; and lanes 16–20, G144A variant. FL, full-length transcripts protected from RNase H cleavage. C, transcripts cleaved by RNase H. Percent protection is amount of full-length RNAs relative to total amount of RNA transcribed.

of helix 1 in the A112G construct was more effective in the presence of lys than protection in the presence of lys analogs. RNase H probing of the G40A–C110U variant (Figure 5, lanes 11–15) resulted in 38% protection in the absence of ligand and 51% protection in the presence of lysOH. Protection of the G40A–C110U variant in presence of lys (40%) and AEC (33%) decreased 2-fold as compared to protection of wild-type in presence of lys or AEC (Figure 5). The lack of helix 1 formation by the G40A–C110U variant is in agreement with *in vitro* termination results that suggest that the G40A–C110U variant is specific for lysOH and discriminates against lys and AEC. Similarly, protection of the G144A variant increased 2-fold in the presence of lysEE as compared to the no ligand control (Figure 5, lane 20). Protection of the G144A and wild-type constructs in the presence of lysEE was 48% and 60%, respectively. For each *lysC* leader variant, stabilization of helix 1 was specific for the lys analogs to which they responded in the *in vitro* transcription assay and was reduced in the presence of lys. These results confirm that the ligand-dependent termination effects described above correlate with stabilization of helix 1.

DISCUSSION

Previous studies of riboswitch specificity have identified RNA elements involved in ligand recognition (25,30,31). However, these features are variable and often specific to the RNA and its cognate ligand. Structural analysis of ligand-bound riboswitches has revealed mechanistic themes employed for ligand recognition. Ligand features such as negatively charged groups and nucleotide-like structures are targets for recognition. Shape complementarity, base pairing, nucleotide stacking and metal ion coordination are known mechanisms of ligand recognition by riboswitches. The molecular structure of lys makes a number of these mechanisms impossible.

Recent analysis of the aptamer region of the *T. maritima* L box leader RNA has identified the nucleotide composition of the lys-binding pocket. Although the nucleotides that directly contact lys were identified, the molecular interactions and their effect on function are poorly understood. In this study, we use the lys-responsive *B. subtilis lysC* leader RNA to analyse recognition determinants and their effects on ligand-dependent structural changes and transcription termination.

Analysis of transcription *in vitro* in the presence of lys analogs correlates with the crystal structure, which suggested that lys recognition relies heavily on shape complementarity between the ligand and the RNA-binding pocket (23). This feature seems to be a common characteristic of riboswitch RNAs and is also observed in the c-di-GMP and TPP riboswitches (17,18). For the *lysC* leader RNA, the effective concentration for each analog is dependent on its structural similarity to lys. The primary determinant for ligand recognition by the *lysC* leader RNA appears to be the ability of the ligand to fit into the lys-binding pocket. This suggests that shape recognition is likely to be the first step in discrimination between lys and similar molecules. Analogs that maintain the overall shape of the molecule, with minor molecular substitutions such as those found in AEC and lysNH, were more effective at promoting both the ligand-dependent structural changes and transcription termination. Changes to the chemistry and charge of the molecule at either the carboxyl or ε-amino termini were tolerated to some degree, and groups at these positions could be replaced with other functional groups that make similar interactions. It is possible that molecular extensions such as a methyl group (lys methyl) or ethyl ester group (lysEE) can fit into small channels in the RNA pocket found near the carboxyl and ε-amino groups. Although these analogs fit into the binding pocket, they are unable to make the necessary molecular interactions required to promote termination at concentrations similar to those required for lys. This suggests that the lys-binding pocket has evolved to bind the lys molecule with high affinity and specificity. These results demonstrate that the *B. subtilis lysC* leader RNA is optimized for lys-dependent regulation.

Mutational analysis uncovered features of the RNA that are essential for lys recognition. G111 is critical for lys-dependent regulation. Previous studies have indicated that this residue is involved in structurally significant base

triple interactions that may be essential for formation of the lys-binding pocket (23). The data presented here provide evidence that G111 is required for formation of the lys-binding pocket as well as lys recognition. It is possible that G111 must be precisely positioned to interact with other nucleotides within the RNA and the bound lys molecule. Additionally, our data suggest that A112 is not essential for ligand recognition but is primarily involved in lys affinity. This effect on affinity may be achieved through extensive contacts with A112 that stabilize the structure of the lys-binding pocket. Variants of the lys-binding pocket near the carbonyl region of the lys-molecule (G40A–C110U and G144A–U170) exhibited preferential specificity for certain lys analogs and discriminated against lys. The alternate specificity of leader RNA variants suggests that each base pair is responsible for the recognition of molecular features of the lys carboxyl group. Additionally, mutations in the carboxyl region of the binding pocket did not inhibit recognition of the ε-amino region, as G144–U170 variants retained the ability to recognize lys methyl. The leader RNA is able to use alternate molecular interactions with lys analogs to promote termination. However, the increased concentration requirement for lys analogs suggests that the new RNA–analog interactions do not perfectly mimic the interactions with lys. This analysis confirms that the sequence of *lysC* leader RNA is optimized for lys recognition as all of the variants in this study, with the exception of A112G, exhibited a loss of lys-dependent termination. In total, these data indicate that recognition of lys by the L box riboswitch is a modular process in which individual nucleotides recognize independent features of the lys molecule.

The presence of a K^+ ion in the L box crystal structure (23) raised questions about the role of this ligand in binding. Analysis of the effects of the addition of K^+ to the *in vitro* transcription termination assay revealed that the potassium ion may increase the affinity of the leader RNA for lys. The increased affinity was observed only in the presence of the wild-type RNA and lys. This suggests that both the sequence and structure of the binding pocket are necessary for stimulation of tighter lys binding by K^+. Additional studies are required to determine the direct effects of K^+ on lys binding.

The discovery of new riboswitches and characterization of RNA–ligand complexes continue to diversify our understanding of the mechanisms used by RNAs to recognize and bind small molecules. Small molecule recognition by riboswitches is essential to regulation, as gene expression must occur only in response to the appropriate stimulus. Recognition of the appropriate ligand by riboswitch RNAs can occur in a variety of ways with some common themes distributed amongst the different classes. The *lysC* leader RNA not only utilizes some common mechanisms, such as shape complementarity and metal ion assistance for lys binding and recognition, but also makes nucleotide-specific interactions for ligand identification. These interactions differ from the mechanism used by the glycine riboswitch to recognize similar amino acid features (19). The alternate specificities exhibited by *lysC* variants demonstrate that the leader RNA uses a unique arrangement of contacts to recognize the ligand and regulate gene expression. This characterization has provided a more detailed view of how the *lysC* leader RNA specifically recognizes the lys molecule and contributes to our current understanding of small molecule recognition by RNAs. However, the specific features responsible for lys binding affinity are still unclear and require further analysis.

ACKNOWLEDGEMENTS

We would like the thank members of the Henkin lab for helpful discussions and Dr Mike Ibba for providing the lys analogs.

FUNDING

National Institutes of Health (R01GM063615 to T.M.H. and F31GM076894 to S.M.). Funding for open access charge: NIH.

Conflict of interest statement. None declared.

REFERENCES

1. Henkin,T.M. (2008) Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes. Dev.*, **22**, 3383–3390.
2. Smith,A.M., Fuchs,R.T., Grundy,F.J. and Henkin,T.M. (2010) Riboswitch RNAs: regulation of gene expression by direct monitoring of a physiological signal. *RNA Biol.*, **7**, 104–110.
3. Cheah,M.T., Watcher,A., Sudarsan,N. and Breaker,R.R. (2007) Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature*, **447**, 497–501.
4. Grundy,F.G., Lehman,S.C. and Henkin,T.M. (2003) The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc. Natl Acad. Sci.*, **100**, 12057–12062.
5. Sudarsan,N., Wickiser,J.K., Nakamura,S., Ebert,M.S. and Breaker,R.R. (2003) An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes. Dev.*, **17**, 2688–2697.
6. Rodionov,D.A., Vitreschak,A.G., Mironov,A.A. and Gelfand,M.S. (2003) Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch? *Nucleic Acids Res.*, **31**, 6748–6757.
7. Serganov,A. and Patel,D.J. (2009) Amino acid recognition and gene regulation by riboswitches. *Biochim. Biophys. Acta.*, **1789**, 592–611.
8. Jorth,P. and Whiteley,M. (2010) Characterization of a novel riboswitch-regulated lysine transporter in *Aggregatibacter actinomycetemcomitans*. *J. Bacteriol.*, **192**, 6240–6250.
9. Patte,J.C., Akrim,M. and Mejean,V. (1998) The leader sequence of the *Escherichia coli lysC* gene is involved in the regulation of LysC synthesis. *FEMS Microbiol. Lett.*, **169**, 165–170.
10. Pavelka,M.S. and Jacobs,W.R. (1996) Biosynthesis of diaminopimelate, the precursor of lysine and a component of the peptidoglycan, is an essential function of *Mycobacterium smegmatis*. *J. Bacteriol.*, **178**, 6496–6507.
11. Forman,M. and Aronson,A. (1972) Regulation of dipicolinic acid biosynthesis in sporulating *Bacillus cereus* characterization of enzymatic changes and analysis of mutants. *Biochem. J.*, **126**, 503–513.
12. Blount,K.F., Wang,J.X., Lim,J., Sudarsan,N. and Breaker,R.R. (2007) Antibacterial lysine analogs that target lysine riboswitches. *Nat. Chem. Biol.*, **3**, 44–49.
13. Noeske,J., Richter,C., Grundl,M.A., Nasiri,H.R., Schwalbe,H. and Wöhnert,J. (2005) An intermolecular base triple as the basis of ligand specificity and affinity in the guanine- and adenine-sensing riboswitch RNAs. *Proc. Natl Acad. Sci.*, **102**, 1372–1377.

14. Mandal,M.M. and Breaker,R.R. (2004) Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat. Struct. Mol. Biol.*, **11**, 29–35.
15. Kim,J., Roth,A. and Breaker,R.R. (2007) Guanine riboswitch variants from *Mesoplasma florum* selectively recognize 2'-deoxyguanosine. *Proc. Natl Acad. Sci.*, **104**, 16092–16097.
16. Lu,C., Ding,F., Chowdhury,A., Pradhan,V., Tomsic,J., Holmes,W.M., Henkin,T.M. and Ke,A. (2010) SAM recognition and conformational switching mechanism in the *Bacillus subtilis yitJ* S box/SAM-I riboswitch. *J. Mol. Biol.*, **404**, 803–818.
17. Edwards,T.E. and Ferré-D'Amaré,A.R. (2006) Crystal structures of the *Thi*-Box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. *Structure*, **14**, 1459–1468.
18. Smith,K.D., Lipchock,S.V., Livingston,A.L., Shanahan,C.A. and Strobel,S.A. (2010) Structural and biochemical determinants of ligand binding by the c-di-GMP Riboswitch. *Biochemistry*, **49**, 7351–7359.
19. Huang,L., Serganov,A. and Patel,D.J. (2010) Structural insights into ligand recognition by a sensing domain of the cooperative glycine riboswitch. *Mol. Cell.*, **40**, 774–786.
20. Edwards,T.E., Klein,D.J. and Ferré-D'Amaré,A.R. (2007) Riboswitches: small-molecule recognition by gene regulatory RNAs. *Curr. Opin. Struct. Biol.*, **17**, 273–279.
21. Mandal,M., Lee,M., Barrick,J.E., Weinberg,Z., Emilsson,G.M., Ruzzo,W.L. and Breaker,R.R. (2004) A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science*, **306**, 275–279.
22. Garst,A.D., Héroux,A., Rambo,R.P. and Batey,R.T. (2008) Crystal structure of the lysine riboswitch regulatory mRNA element. *J. Biol. Chem.*, **283**, 22347–22351.
23. Serganov,A., Huang,L. and Patel,D.J. (2008) Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature*, **455**, 1263–1268.
24. Blouin,S., Chinnappan,R. and Lafontaine,D.A. (2011) Folding of the lysine riboswitch: importance of peripheral elements for transcriptional regulation. *Nucleic Acids Res.*, **39**, 3373–3387.
25. McDaniel,B.A., Grundy,F.J. and Henkin,T.M. (2005) A tertiary structural element in S box leader RNAs is required for *S*-adenosylmethionine-directed transcription termination. *Mol. Microbiol.*, **57**, 1008–1021.
26. Yousef,M.R., Grundy,F.J. and Henkin,T.M. (2005) Structural transitions induced by the interaction between $tRNA^{Gly}$ and the *Bacillus subtilis glyQS* T box leader RNA. *J. Mol. Biol.*, **349**, 273–287.
27. Grundy,F.J., Winkler,W.C. and Henkin,T.M. (2002) tRNA-mediated transcription antitermination *in vitro*: codon-anticodon pairing independent of the ribosome. *Proc. Natl Acad. Sci.*, **99**, 11121–11126.
28. Qi,Y. and Hulett,F.M. (1998) Pho∼P and RNA polymerase σ^A holoenzyme are sufficient for transcription of Pho regulon promoters in *Bacillus subtilis*: Pho∼P activator sites within the coding region stimulate transcription *in vitro*. *Mol. Microbiol.*, **28**, 1187–1197.
29. Lu,Y., Chen,N. and Paulus,H. (1991) Identification of *aecA* mutations in *Bacillus subtilis* as nucleotide substitutions in the untranslated leader region of the aspartokinase II operon. *J. Gen. Microbiol.*, **137**, 1135–1143.
30. Serganov,A., Polonskaia,A., Phan,A.T., Breaker,R.R. and Patel,D.J. (2006) Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature*, **441**, 1167–1171.
31. Kim,J. and Breaker,R.R. (2008) Purine sensing by riboswitches. *Biol. Cell*, **100**, 1–11.