ORGANIC LETTERS

XXXX
Vol. XX, No. XX
000–000

Efficient Preparation of 2, 4-Diaminopyrimidine Nucleosides: Total Synthesis of Lysidine and Agmatidine

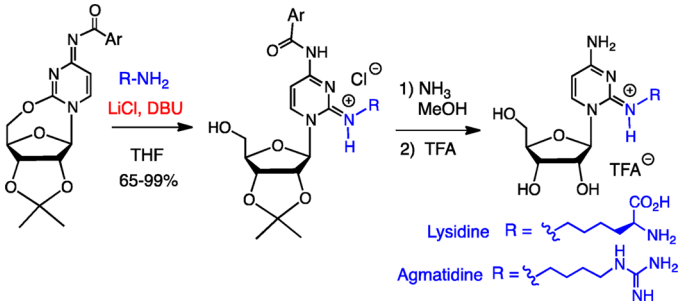
Brett J. Kopina and Charles T. Lauhon*

Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin—Madison, Madison, Wisconsin 53705, United States

clauhon@wisc.edu

Received June 27, 2012

ABSTRACT



An efficient route for the synthesis of 2,4-diaminopyrimidine ribosides from cytidine is described consisting of six steps with overall yields >50% and only one chromatographic step. The key amine addition step utilizes LiCl and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) to ensure clean conversion to a single tautomeric product. This route has been used to prepare the modified tRNA nucleosides lysidine and agmatidine in quantities suitable for structural characterization.

2,4-Diaminopyrimidine ribosides are one of the poorest studied classes of nucleosides capable of Watson–Crick pairing. Only a handful of derivatives have been described, and a general large scale method for their preparation has yet to be reported. Two members of this class of nucleosides have biological significance. Lysidine (**1**, Figure 1) is a modified nucleoside found in bacterial tRNA and is nearly universally conserved.[1] The similarly structured agmatidine (**2**) is exclusive to tRNA in archea where it serves a similar purpose.[2] These N1-alkylated 2,4-diaminopyrimidines are unique in that they are stable despite their unusually high basicity, with pK_a values close to 13.[3] In addition, they have access to a number of tautomeric forms for which direct experimental evidence is lacking.

The biosynthesis of lysidine[4] and agmatidine[5] occurs post-transcriptionally from tRNA at cytidine 34, the first position of the anticodon. This single modification changes not only the amino acid attached to the tRNA (from methionine to isoleucine) but also its codon recognition from 5′-AU**G** to 5′-AU**A**.[6] The altered pairing preference of lysidine and agmatidine for A over G is proposed to be the result of a combination of steric bulk at position 2 and protonation at N3.[1] However, there is to date no experimental evidence to support this hypothesis. Thus, it is useful to have an efficient route to prepare these nucleosides as well as analogs for both structural and biological study.

(1) Muamatsu, T.; Yokoyama, S.; Horie, N.; Matsuda, A.; Ueda, T.; Yamaizumi, Z.; Kuchino, Y.; Nishimura, S.; Miyazawa, T. *J. Biol. Chem.* **1988**, *263*, 9261.

(2) (a) Mandal, D.; Kohrer, C.; Su, D.; Russell, S. P.; Krivos, K.; Castleberry, C. M.; Blum, P.; Limbach, P. A.; Soll, D.; Raj Bhandary, U. L. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 2872. (b) Ikeuchi, Y.; Kimura, S.; Numata, T.; Nakamura, D.; Yokogawa, T.; Ogata, T.; Wada, T.; Suzuki, T.; Suzuki, T. *Nat. Chem. Biol.* **2010**, *6*, 277.

(3) Brown, D. J.; Jacobsen, N. W. *J. Chem. Soc.* **1962**, 3172.

(4) (a) Soma, A.; Ikeuchi, Y.; Kanemasa, S.; Kobayashi, K.; Ogasawara, N.; Ote, T.; Kato, J.; Watanabe, K.; Sekine, Y.; Suzuki, T. *Mol. Cell* **2003**, *12*, 689. (b) Ikeuchi, Y.; Soma, A.; Ote, T.; Kato, J.; Sekine, Y.; Suzuki, T. *Mol. Cell* **2005**, *19*, 235.

(5) (a) Terasaka, N.; Kimura, S.; Osawa, T.; Numata, T.; Suzuki, T. *Nat. Struct. Biol.* **2011**, *18*, 1268. (b) Osawa, T.; Satoshi, K.; Terasaka, N.; Inanaga, H.; Suzuki, T. *Nat. Struct. Biol.* **2011**, *18*, 1275.

(6) Muramatsu, T.; Nishikawa, K.; Nemoto, F.; Kuchino, Y.; Nishimura, S.; Miyazawa, T.; Yokoyama, S. *Nature* **1988**, *336*, 179.

10.1021/ol301769j

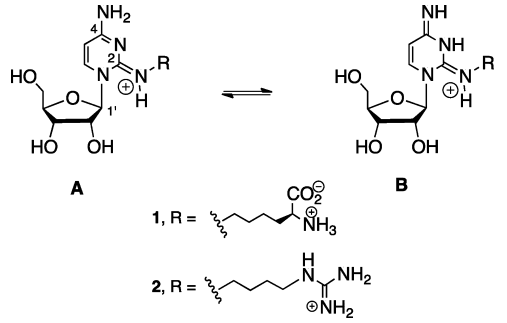


Figure 1. N4 amino (A) and imino (B) tautomeric structures for the protonated forms of lysidine **1** and agmatidine **2**.

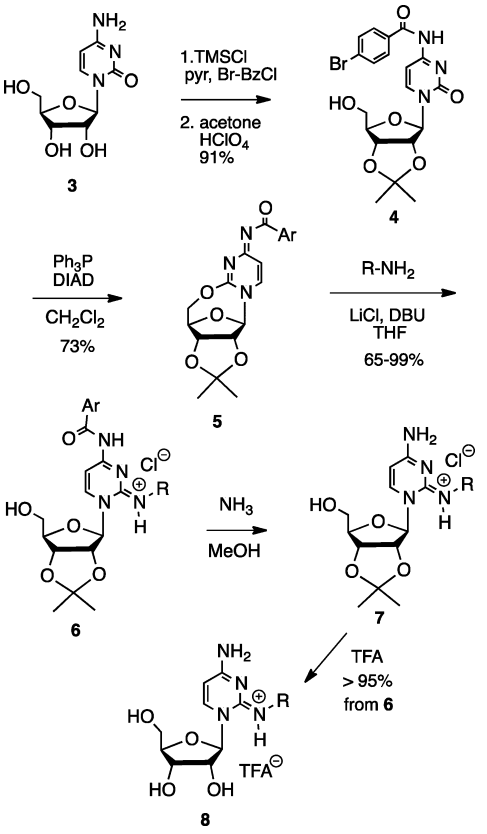
Previously, lysidine and agmatidine have been prepared in micromole amounts for initial identification of the biologically produced nucleosides.[1,2] These routes involve reacting protected lysine or free agmatine with S-alkylated 2-thiocytidine salts using the methodology of Ueda et al.[7] Since 2-thiocytidine is not readily available in large amounts, we looked for a more practical and efficient route from cytidine. We report herein a general synthetic route that can provide a number of N2-alkylated-2,4-diaminopyrimidine nucleosides in high yield with only one chromatographic step. In addition, we present the first experimental data on the structure and tautomeric state of these nucleosides.

Our synthesis (Scheme 1) begins with cytidine and utilizes 4-bromobenzoyl and 2′,3′-isopropylidene protection to give intermediate **4** in 91% overall yield. The cyclization of N4-benzoylcytidine to the 2,5′-anhydrocytidine derivative under Mitsonobu conditions has been previously reported in modest yield.[8] Our use of *p*-bromobenzoyl protection at N4 gave a more crystalline product **5** that could be cleanly isolated from the reaction byproducts in 73% yield on a multigram scale.

The next step involved addition of an amine to the 2,5′-anhydrocytidine **5** specifically at C-2. In our hands reaction of the anhydrocytidine derivatives with alkylamines alone was sluggish, providing low yields of complex product mixtures. However, addition of LiCl with THF as solvent gave good conversion to a small number of amine addition products. NMR and LC-MS analysis of the products was consistent with a mixture of tautomeric hydrochloride salts. Interestingly, these isomers appear to interconvert slowly enough to be partially separable. However, conditions in which the purported tautomers can be reliably separated or interconverted have proven elusive.

To resolve the issue of tautomeric mixtures, 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) was added to the amine addition reactions to keep the product in a neutral form. This modification resulted in the production of a

Scheme 1. Synthetic Route toward 2,4-Diaminopyrimidine Nucleosides



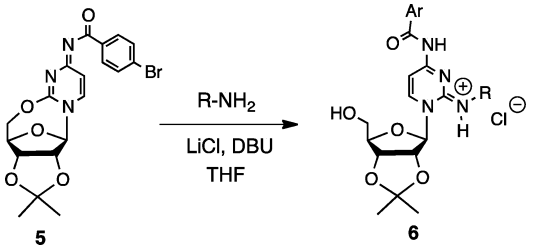
single product by TLC. After silica gel chromatography the desired products were isolated as hydrochloride salts. As shown in Table 1 (entries 1−7), the optimized reaction conditions gave high yields under mild conditions for a variety of basic primary amines. Less basic amines, such as anilines or hydroxylamines (entries 8 and 9), and secondary amines (not shown) did not react under these conditions. Unprotected agmatine appeared to react smoothly, but we were unable to isolate the resulting product.

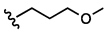
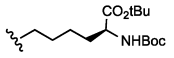
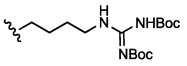
The use of LiCl in these reactions was intended to increase the reactivity of the electrophile via coordination to either the acyl group or the 5′-oxygen. We observed that LiCl increased both the solubility of the reactants and the rate of reaction. It also provided a counterion for efficient purification of the product by silica gel chromatography. Of the many Lewis acids tested, only LiCl and $LiBF_4$ gave satisfactory results. Reactions containing only DBU gave conversion to a product that could only be isolated in low yield.

To confirm the structural characteristics of these molecules, the ethanolamine adduct **6b** was crystallized from propanol and an X-ray structure was obtained. The structure (Figure 2) clearly shows the product as a hydrochloride salt in which the chloride is found in close proximity to C-2 of the pyrimidine ring. This is consistent with NMR evidence for protonation at N2. The ethanol side chain is in the *trans* configuration relative to the glycosidic bond to

(7) Ueda, T.; Ohtsuka, H. *Chem. Pharm. Bull.* **1973**, *21*, 1530.
(8) Kimura, J.; Yagi, K.; Suzuki, H.; Mitsunobu, O. *Bull. Chem. Soc. Jpn.* **1980**, *53*, 3670.

Table 1. Addition of Amines to 2,5′-Anhydrocytidine **5**



entry	R	time (min)	temp (°C)	product	yield (%)[a]
1		45	0	**6a**	96
2		30	rt	**6b**	99
3	Bn	60	0	**6c**	95
4	iPr	45	0	**6d**	94
5	cyclohexyl	60	0	**6e**	65
6		30	0	**6f**	84
7		20	rt	**6g**	85
8	Ph				nd[b]
9	-OBn				nd

[a] Yields are for the isolated products after chromatography. [b] No product detectable.

avoid interaction with the H1′ proton as previously predicted for lysidine itself[1] using computational methods.[9] NMR analysis of **6b** is consistent with the tautomer shown in Figure 2 in which the N2 and N4 nitrogens are protonated. Although the positive charge is likely greatest among the N1−C2−N2 atoms of the ring, a true representation is likely that of a highly resonant cation in which electrons are delocalized throughout the ring as well as the exocyclic nitrogens.[10]

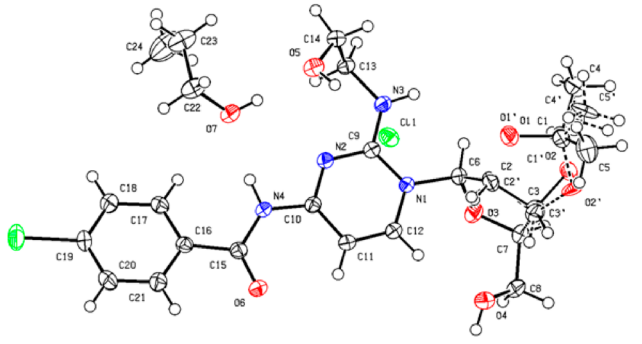


Figure 2. X-ray structure of intermediate **6b** crystallized from *n*-propanol. Two populated conformations of the isopropylidene are shown.

Deprotection of the amine addition products **6** by sequential treatment with ammonia in methanol, followed by trifluoroacetic acid at 0 °C, gave the desired target molecules **8b** (R = CH_2CH_2OH), lysidine (**1**), and agmatidine (**2**) in high yield and > 95% purity. Overall yields for the synthesis of both lysidine and agmatidine from cytidine ranged from 50 to 60%.

While the usual convention is to draw the cationic derivatives of lysidine and agmatidine as imino tautomers with the N2, N3, and N4 nitrogens each protonated (Structure **B** in Figure 1), NMR analysis is consistent with the N4 amino tautomer **A**. Figure 3 shows ^{15}N HSQC data for agmatidine in DMSO-d_6. The two NH protons at 8.1 and 8.3 ppm both correlate to a single nitrogen at 109 ppm, assigned as N4. The ^{15}N signal at 97 ppm is assigned as N2 based on gCOSY data, in which the N2 proton correlates to the adjacent methylene protons. The ^{15}N signal at 137 ppm is assigned to the guanidine of the agmatine side chain based on gCOSY data and its significant ^{15}N chemical shift in the HSQC spectrum of Boc-protected precursor **7g**.[11] Lysidine exhibits a similar HSQC spectrum under these conditions, except that the ^{15}N signal at 137 ppm is absent.[11]

Although our NMR data support the amino tautomer in DMSO, this does not preclude involvement of the imino tautomer in base pairing. We have been unable thus far to determine the tautomeric structures of lysidine and agmatidine in water at pH 7, as the heteroatom protons exchange rapidly with solvent. Thus, in aqueous media at physiological pH it is possible that the tautomeric equilibrium is shifted toward the imino form. Alternatively, the tautomers may interconvert rapidly enough to provide the imino tautomer for base pairing to adenosine.

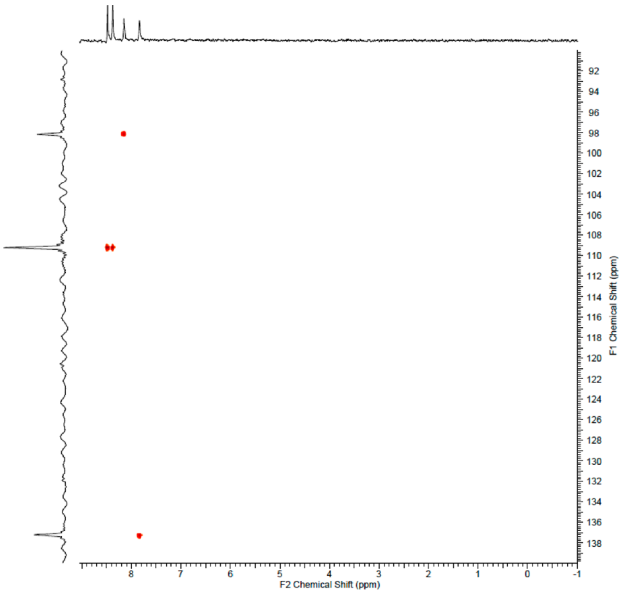


Figure 3. ^{15}N HSQC data for agmatidine in DMSO-d_6 consistent with tautomeric structure **2A** (Figure 1). The ^{15}N signal at 137 ppm is absent in the lysidine ^{15}N HSQC data.

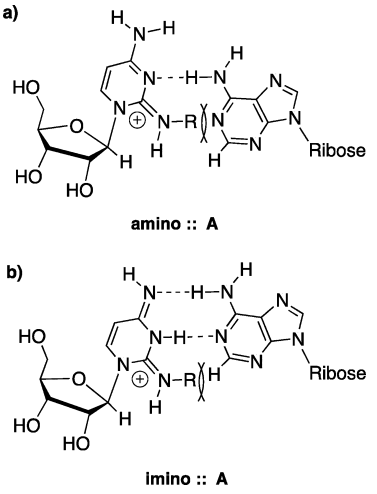


Figure 4. Potential base pairing interaction between adenosine (**A**) and (a) the amino tautomer of lysidine and agmatidine or (b) the imino tautomer.

The predominance of the N4-amino tautomer in the ^{15}N HSQC spectra of lysidine and agmatidine is consistent with previous work on *arabino* diaminopyrimidines, in which the authors favored protonation on the exocyclic nitrogens based on UV spectral analysis.[10] Interestingly, this tautomer precludes donation of a hydrogen bond from N3 to N1 of adenosine during base pairing. As shown in Figure 4, this hydrogen bond is crucial considering the steric demand imposed by the alkylamino substituent at C-2.[1,9] Thus, for a G−U wobble type of pairing[12] (Figure 4a) between the amino tautomer of lysidine/agmatidine and adenosine, two hydrogen bonds can only be obtained if the N2-alkyl group is in the unfavorable *cis* orientation. Pairing with a single hydrogen bond and the N2-alkyl group *trans* will incur steric interaction with the adenine base. The imino tautomer (Figure 4b) is able to form a Watson−Crick base pair with two hydrogen bonds and the more favorable *trans* N2-alkyl group due to protonation at N3. However, there remains a possible steric interaction with H-2 of the adenine base. Thus, it is likely that some deviation in the geometry of the base pairing accompanies the lysidine/agmatidine interaction in the anticodon. Such deviation may be more readily achieved for the amino tautomer which is less constrained. This possibility should not be discounted, as wobble pairings with altered geometry and hydrogen bonding patterns are common.[13]

In summary, we have developed an efficient synthetic route to the biologically important 2,4-diaminopyrimidine ribosides lysidine and agmatidine. We have also presented the first experimental evidence describing their conformation and tautomeric state. Current efforts are underway to expand our initial structural studies and to incorporate these nucleosides and their analogs into RNA to fully understand the mechanism of their binding specificity.

Acknowledgment. The authors acknowledge Dr. Victor Young at the University of Minnesota for determining the crystal structure of intermediate **6b**; Tom Stringfellow, Chenxi Jia, and Cameron Scarlett of the UW—Madison School of Pharmacy Analytical Instrumentation Center; and Richard Hsung for discussion and comments on the manuscript. We also gratefully acknowledge the University of Wisconsin School of Pharmacy for financial support.

Supporting Information Available. Experimental procedures, NMR and mass spectrometry data, and the crystallographic information file for compound **6b**. This material is available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interest.

(9) Sonawane, K. D.; Tewari, R. *Nucleosides, Nucleotides Nucleic Acids* **2008**, *27*, 1158.
(10) Doerr, I. L.; Fox, J. J. *J. Chem. Soc.* **1967**, 1462.
(11) See Supporting Information.
(12) Varani, G.; McClain, W. H. *EMBO Rep.* **2000**, *1*, 18.
(13) Agris, P. F.; Vendeix, A. P.; Graham, W. D. *J. Mol. Biol.* **2006**, *366*, 1.