The practical acetylation of nucleosides using acetic anhydride/acetic acid as a reusable solvent

Ran Xia¹, Chao Xia¹, Ying-Xing Yang¹, Li-Jie Liu¹, Lei-Shan Chen² and Li-Ping Sun²

Abstract
The highly practical acetylation of free nucleosides is achieved using acetic anhydride/acetic acid as a reusable solvent and acetylating regent. A series of nucleosides, including ribosyl, deoxyriboosyl, arabinosyl, acyclic and pyranosyl, and many clinical drugs were acetylated efficiently, even on large scale (200 g).

Keywords
acetylation, hydroxy group protection, late-stage modification, nucleosides, reusable solvent

Introduction
In recent years, modified nucleosides have become of significant interest due to their intriguing biological and pharmacological properties.¹,² In the context of the outbreak of COVID-19, nucleoside drugs have once again attracted the attention of scientists, and remdesivir,³ molnupiravir,⁴ azvudine⁵ have been used clinically for the treatment of COVID-19. The late-stage modification of nucleosides is an important method for obtaining a variety of nucleoside drugs.⁵ However, the direct modification of free nucleosides is often inhibited or there is a lack of selectivity due to the presence of hydroxy groups on the sugar rings.⁶,⁷ As a result, the development of methods for the protection of hydroxy groups has been the subject of intense research.⁸

In addition, the presence of an acetoxy group in pharmaceuticals often modulates their solubility and pharmacokinetics, while optimizing their pharmacological and...
physiochemical properties. The acetylation of hydroxy groups is often a powerful and low-cost approach, which is routinely carried out by reacting free nucleosides with acetic anhydride in the presence of a catalyst, for example, DMAP (4-dimethylaminopyridine), which is the most widely used catalyst for this purpose. In factories, the most common way to produce 1,2,3,5-tetraacetyl-β-D-ribose (1) is through the cleavage of glycosidic bonds of 2',3',5'-tri-O-acetylinosine (2a) catalyzed by boric acid. However, we found that the residual DMAP in 2a could inhibit the reaction or reduce the yield of compound 1 (Scheme 1). Therefore, more steps have to be taken to remove DMAP and to improve the purity of 2a.

This finding prompted us to improve the acetylation of the hydroxy groups of nucleosides, avoiding the existing deficiencies. Following on from our previous research on nucleosides, we herein report the practical acetylation of the hydroxy groups of nucleosides without an additional catalyst by using acetic anhydride/acetic acid as a reusable solvent.

We began our studies by using inosine (3a) as a model substrate to optimize the reaction conditions (Scheme 2, Table 1). Under the conditions of a previous report, we found that by using molecular sieves as the catalyst and acetic anhydride as the solvent, product 2a was obtained in 91% yield. Although molecular sieves were better than DMAP, pure acetic anhydride was used as the solvent, which increased the production cost. We speculated whether the acetic anhydride recovered during the work-up process, and containing some acetic acid generated during the reaction, could be reused. To our delight, the yield of 2a was 63% when recovered acetic anhydride (85% content) was used.

### Table 1. Optimization of the reaction conditions for the acetylation of inosine (3a).

| Entry | Temp (°C) | Ac2O content of Ac2O/AcOH | Time (h) | Yield (%) |
|-------|-----------|---------------------------|----------|-----------|
| 1     | 60        | 85%                       | 4        | 63        |
| 2     | 70        | 85%                       | 4        | 77        |
| 3     | 80        | 85%                       | 4        | 85        |
| 4     | 80        | 85%                       | 6        | 92        |
| 5     | 80        | 85%                       | 8        | 91        |
| 6b    | 80        | 71%                       | 6        | 91        |
| 7c    | 80        | 41%                       | 6        | 90        |
| 8d    | 80        | 29%                       | 6        | 80        |

*a Reaction conditions: 3a (1 mmol), solvent (2 mL); isolated yields are given.
*b The solvent was reused 2 times.
*c The solvent was reused 4 times.
*d The solvent was reused 5 times.
used as the solvent at 60 °C for 4 h (entry 1). The C-N glycosidic bond of 2a could be successfully cleaved by boric acid to produce 1. The yield of 2a increased when the reaction temperature was raised, and 80 °C proved to be the better choice (entries 2 and 3). Screening of the reaction time showed that 6 h was the best choice (entries 4 and 5). The yield of 2a was maintained, when using recovered acetic anhydride (following a second run); in this case, the content of acetic anhydride was 71% (entry 6). Further research showed that the solvent could be recycled at least 4 times, and that the content of acetic anhydride was 41% (entry 7). The yield of 2a was maintained, when using recovered acetic anhydride (following a second run); in this case, the content of acetic anhydride was 71% (entry 6). Further research showed that the solvent could be recycled at least 4 times, and that the content of acetic anhydride was 41% (entry 7). The yield of 2a was maintained, when using recovered acetic anhydride (following a second run); in this case, the content of acetic anhydride was 71% (entry 6). Further research showed that the solvent could be recycled at least 4 times, and that the content of acetic anhydride was 41% (entry 7).

Under the optimized reaction conditions and with 60% acetic anhydride/acetic acid as the solvent, a series of free nucleosides was subjected to the acetylation reaction under the optimized conditions (Scheme 3). To our delight, all the hydroxy groups were acetylated to afford satisfactory yields of products (2a-h, 86%–93%). This system could tolerate a variety of substituents on purine rings such as Cl, F, NH₂, and OH and the substitutes had little impact on the yields (2b vs 2d, 2e vs 2f). This catalytic system proved highly efficient for the ribosides 2a-d. It was worth mentioning that acetylation occurred exclusively on the hydroxy groups of the sugar, and the probable reason was that the amino groups were protonated under the reaction conditions.

Subsequently, a variety of nucleosides bearing other glycols or nucleobases (4a-4g) were then studied (Scheme 4). Acyclic, pyranosyl and pyrimidine nucleosides all reacted well to give the corresponding acetylated products 5a-g in good to high yields of 86%–94%. Famiclovir, an effective oral prodrug of the antiviral penciclovir for the treatment of herpes zoster was synthesized in 88% yield. Three pyrimidine nucleosides also reacted smoothly to afford products 5f-h in high yields. Xuriden (5h) is used to treat ultra-orphan indication hereditary whey aciduria. These results showed that this catalytic system tolerates various substituents.

Next, reactions on large scale demonstrated the robustness and preparative scale utility of this process.
2',3',5'-tri-O-Acetylinosine (2a) could be obtained successfully on 200 g scale in a comparable yield of 90%. More importantly, the product could be purified by recrystallization, thereby avoiding chromatography and/or a lengthy work-up process, which makes this route attractive for industrial applications.

**Conclusion**

In summary, we have developed a highly practical and improved process for the acetylation of the hydroxy groups of nucleosides. The acetic anhydride could be reused 4–5 times. A series of nucleosides including ribosyl, deoxyribosyl, arabinosyl, acyclic and pyranosyl worked well. What is more, the catalytic system tolerated different functional groups including F, Cl, and NH2. Importantly, the drugs nebuline, clofarabine, fludarabine, vidarabine, famciclovir and xuriden could be successfully acetylated, providing a simple method for the late-stage modification of these drugs. In addition, the reaction scale could be increased to 200 g, which is important for industrial applications. Further investigations on this reaction are underway in our laboratory.

**Experimental**

**General**

Melting points were recorded with a X-5 micro melting point apparatus and uncorrected (Beijing Taike Instruments Co., Ltd.). 1H and 13C NMR spectra were recorded on a Bruker AC 400 spectrometer (Bruker, Billerica, MA, USA) as CDCl3 or DMSO-d6 solutions. Chemical shifts are expressed in parts per million (δ) downfield from the internal standard tetramethylsilane. Multiplicities are reported as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), and dd (doublet of doublets), and coupling constants (J) are given in hertz (Hz). Mass spectra were obtained on a Waters Q-Tof MicroTM spectrometer (Waters Synapt).

**General procedure**

Free nucleoside (1 mmol) was added to 60% acetic anhydride-acetic acid (2 mL). The resulting mixture was heated to 80 °C and kept at this temperature for 4–8 h. Upon completion of the reaction, the solvent was removed under vacuum. The product was purified by column chromatography on silica gel (eluent: ethyl acetate/petroleum ether) to give the pure product.

**Large-scale procedure (200 g scale) for the synthesis of 2a.** Inosine (200 g, 746 mmol) was added to 60% acetic anhydride-acetic acid (800 mL). The resulting mixture was stirred and heated at 80 °C in an oil bath for 10 h. Upon completion of the reaction, the solvent was recovered under vacuum and the residual solvent was co-evaporated with ethanol (50 mL). The oily residue was recrystallized from H2O to give the pure product (90% yield).

\[
(2R,3R,4R,5R)-2-(Acetoxy methyl)-5-(6-oxo-1H-purin-9(6H)-yl)tetrahydrofuran-3,4-Diyl diacetate (2a): \text{ Yield:}
\]
(2R,3R,4R,5R)-2-(Acetoxymethyl)-5-(6-amino-9H-purin-9-yl)tetrahydrofuran-3,4-diy Diacetate (2b): Yield: 0.357 g (91%). White solid, m.p. 174–176 °C (Lit.18 174–175 °C). 

1H NMR (400 MHz, CDCl3): δ = 8.33 (s, 1H, 2-H); 7.95 (s, 1H, 8-H); 6.18 (d, J = 5.6 Hz, 1H, 1'-H); 6.14 (br s, 2H, NH2); 5.92 (t, J = 5.6 Hz, 1H, 2'-H); 5.66 (t, J = 4.8, 1H, 4'-H); 4.45–4.33 (m, 3H, 3' and 5'-H); 2.12 (s, 3H, OAc). 13C NMR (100 MHz, CDCl3): δ = 170.3 (OAc); 169.5 (OAc); 169.3 (OAc); 155.5 (6-C); 155.0 (4-C); 149.7 (2-C); 138.8 (8-C); 120.0 (5-C); 86.2 (1'-C); 80.2 (2'-C); 73.1 (4'-C); 70.6 (3'-C); 63.6 (5'-C); 21.0 (OAc). HRMS (ESI): m/z [M + H]+ calcd for C16H19N4O8: 394.1357; found: 394.1357.

(2R,3R,4R,5R)-2-(Acetoxymethyl)-5-(9H-purin-9-yl)tetrahydrofuran-3,4-diy Diacetate (2c): Yield: 0.351 g (93%). Colorless oil. 

1H NMR (400 MHz, CDCl3): δ = 9.17 (s, 1H, 6-H); 9.00 (s, 1H, 2-H); 8.27 (s, 1H, 8-H); 6.26 (d, J = 5.2 Hz, 1H, 1'-H); 5.97 (t, J = 5.2 Hz, 1H, 2'-H); 5.67 (t, J = 4.8, 1H, 4'-H); 4.46–4.35 (m, 3H, 3' and 5'-H); 2.15 (s, 3H, OAc), 2.11 (s, 3H, OAc); 2.07 (s, 3H, OAc). 13C NMR (100 MHz, CDCl3): δ = 170.3 (OAc); 169.5 (OAc); 169.3 (OAc); 155.5 (6-C); 155.0 (4-C); 149.7 (2-C); 138.8 (8-C); 120.0 (5-C); 86.2 (1'-C); 80.2 (2'-C); 73.1 (4'-C); 70.6 (3'-C); 63.6 (5'-C); 20.7 (OAc); 20.5 (OAc), 20.3 (OAc). HRMS (ESI): m/z [M + H]+ calcd for C14H18N3O6: 394.1357; found: 394.1357.
Yield: 0.253 g (94%). White solid, m.p. 96–98 °C. (Lit.24 96 °C).\(^1\)H NMR (400 MHz, CDCl\(_3\)): \(\delta = 8.77\) (s, 1H, 2-H), 8.29 (s, 1H, 8-H), 5.72 (s, 2H, -CH\(_2\)-), 4.18 (t, \(J = 4.8\) Hz, 2H, -CH\(_2\)-CH\(_2\)-), 3.77 (t, \(J = 4.8\) Hz, 2H, -CH\(_2\)-CH\(_2\)-), 2.01 (s, 3H, OAc). \(^{13}\)C NMR (100 MHz, DMSO-d\(_6\)): \(\delta = 181.2\) (OAc), 156.0 (6-C), 152.9 (4-C), 149.7 (2-C), 141.2 (8-C), 118.5 (5-C), 72.0 (-CH\(_2\)-CH\(_2\)-), 66.9 (-CH\(_2\)-CH\(_2\)-), 62.8 (-CH\(_2\)-CH\(_2\)-), 20.5 (OAc). HRMS (ESI): m/z [M + H]\(^+\) calcd for C\(_{10}\)H\(_{12}\)ClN\(_4\)O\(_3\): 271.0592; found: 271.0596.

2-(5Amino-9H-purin-9-yl)methoxy)ethyl Acetate (5B): Yield: 0.228 g (91%). White solid, m.p. 156–158 °C. (Lit.25 156–158 °C).\(^1\)H NMR (400 MHz, DMSO-d\(_6\)): \(\delta = 8.29\) (s, 1H, 2-H), 8.17 (s, 1H, 8-H), 7.33 (br s, 2H, NH\(_2\)), 5.56 (s, 2H, -CH\(_2\)-), 4.06 (t, \(J = 4.4\) Hz, 2H, -CH\(_2\)-CH\(_2\)-), 3.70 (t, \(J = 4.4\) Hz, 2H, -CH\(_2\)-CH\(_2\)-), 1.92 (s, 3H, OAc). \(^{13}\)C NMR (100 MHz, CDCl\(_3\)): \(\delta = 170.7\) (OAc), 159.6 (6-C), 152.7 (4-C), 150.8 (2-C), 143.6 (8-C), 132.7 (5-C), 72.6 (-CH\(_2\)-), 67.8 (-CH\(_2\)-CH\(_2\)-), 62.8 (-CH\(_2\)-CH\(_2\)-), 19.5 (OAc). HRMS (ESI): m/z [M + H]\(^+\) calcd for C\(_{10}\)H\(_{14}\)N\(_5\)O\(_4\): 252.1091; found: 252.1095.

2-(2-(6-Amino-9H-purin-9-yl)ethyl)propane-1,3-diy Diacetate (Famiciclovir) (5C): Yield: 0.282 g (88%). White solid, m.p. 100–102 °C. (Lit.26 100–102 °C).\(^1\)H NMR (400 MHz, DMSO-d\(_6\)): \(\delta = 8.57\) (s, 1H, 2-H), 8.05 (s, 1H, 8-H), 6.49 (br s, 2H, NH\(_2\)), 4.14 (t, \(J = 7.2\) Hz, 2H, 1'-H), 4.04 (d, \(J = 5.6\) Hz, 4H, 4' and 5'-H), 1.99 (s, 6H, OAc), 1.95–1.84 (m, 3H, 2' and 3'-H). \(^{13}\)C NMR (100 MHz, DMSO-d\(_6\)): \(\delta = 170.8\) (OAc), 160.9 (OAc), 153.4 (6-C), 149.5 (4-C), 143.1 (2-C), 127.4 (8-C), 63.9 (4' and 5'-C), 35.0 (1'-C), 28.3 (3'-C), 21.1 (OAc). HRMS (ESI): m/z [M + H]\(^+\) calcd for C\(_{16}\)H\(_{20}\)N\(_5\)O\(_4\): 322.1510; found: 322.1516.

Yield: 0.399 g (90%). White solid, m.p. 130–132 °C. (Lit.28 132 °C).\(^1\)H NMR (400 MHz, CDCl\(_3\)): \(\delta = 7.99\) (s, 1H, 2-H), 8.33 (s, 1H, 8-H), 6.81 (br s, 2H, NH), 5.97 (d, \(J = 9.6\) Hz, 1H, 1'-H), 5.68 (t, \(J = 9.6\) Hz, 1H, 3'-H), 5.48 (t, \(J = 9.6\) Hz, 1H, 2'-H), 5.31 (t, \(J = 9.6\) Hz, 1H, 4'-H), 4.33–4.28 (m, 1H, 6'-H), 4.18 (d, \(J = 12.4\) Hz, 1H, 5'-H), 4.07–4.04 (m, 1H, 6'-H), 2.09 (s, 3H, OAc), 2.08 (s, 3H, OAc), 2.05 (s, 3H, OAc), 1.78 (s, 3H, OAc). \(^{13}\)C NMR (100 MHz, CDCl\(_3\)): \(\delta = 170.4\) (OAc), 169.9 (OAc), 169.3 (OAc), 169.0 (OAc), 152.5 (6-C), 151.7 (4-C), 151.6 (2-C), 142.9 (8-C), 131.6 (5-C), 80.9 (1'-C), 75.3 (5'-C), 72.4 (4'-C), 70.2 (3'-C), 67.7 (2'-C), 61.5 (6'-C), 20.7 (OAc), 20.6 (OAc), 20.5 (OAc). HRMS (ESI): m/z [M + H]\(^+\) calcd for C\(_{16}\)H\(_{20}\)N\(_5\)O\(_4\): 466.1569; found: 466.1573.

Yield: 0.375 g (93%). Oil, (Lit. 30 93%). \(^1\)H NMR (400 MHz, CDCl\(_3\)): \(\delta = 9.91\) (s, 1H, NH), 7.88 (s, 1H, 6-H), 6.05 (d, \(J = 4.4\) Hz, 1H, 1'-H), 5.34 (t, \(J = 4.0\) Hz, 2H, 2' and 4'-H), 4.38–4.31 (m, 3H, 3' and 5'-H), 2.21–2.15 (m, 1H, 5'-H), 2.13 (s, 3H, OAc), 2.12 (s, 3H, OAc), 1.95 (s, 3H, 5-CH\(_3\)). \(^{13}\)C NMR (100 MHz, CDCl\(_3\)): \(\delta = 170.4\) (OAc), 169.8 (OAc), 163.0 (4-C), 149.4 (2-C), 136.0 (6-C), 121.0 (5-C), 91.8 (1'-C), 78.7 (4'-C), 70.0 (3'-C), 62.8 (5'-C), 35.3 (2'-C), 20.8 (OAc), 20.4 (OAc), 13.1 (5'-CH\(_3\)). HRMS (ESI): m/z [M + H]\(^+\) calcd for C\(_{16}\)H\(_{20}\)N\(_5\)O\(_4\): 327.1187; found: 327.1189.
Xuriden (5h): Yield: 0.333 g (90%). White solid, m.p. 127–128 °C. (Lit.11 127–130 °C). 1H NMR (400 MHz, CDCl3); δ = 9.68 (s, 1H, NH), 7.40 (d, J = 8.4 Hz, 1H, 5-H), 6.04 (d, J = 3.6 Hz, 1H, 1'-H), 5.79 (d, J = 7.2 Hz, 1H, 2'-H), 5.32 (s, 2H, 3' and 4'-H), 4.34 (s, 3H, 5'-H and 6-H), 2.13 (s, 3H, OAc), 2.12 (s, 3H, OAc), 2.09 (s, 3H, OAc). 13C NMR (100 MHz, CDCl3); δ = 170.2 (OAc), 170.1 (OAc), 169.7 (OAc), 162.8 (4-C), 150.3 (2-C), 139.3 (6-C), 103.4 (5-C), 87.5 (1'-C), 80.0 (4'-C), 72.7 (2'-C), 70.2 (3'-C), 63.2 (5'-C), 20.8 (OAc), 20.5 (OAc), 20.4 (OAc). HRMS (ESI): m/z [M + H]+ calcd for C15H19N2O9: 371.1085; found: 371.1089.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We thank the Key Scientific and Technical Projects of Henan Province (212102310333, 222102110403), the Training Plan for Young Key Teachers in Universities of Henan Province (20GGJS248), the Key Scientific and Technical Projects of Xinxiang City (GG2021016) and the Major Scientific and Technical Projects of Xinxiang City (21ZD008) for financial support.

ORCID iDs
Ran Xia https://orcid.org/0000-0001-8525-850X
Li-Ping Sun https://orcid.org/0000-0003-3009-4354

Supplemental material
Supplemental material for this article is available online.

References
1. Miller GJ. Science 2020; 369: 623.
2. Xu J, Chmela V, Green NJ, et al. Nature 2020; 582: 60–66.
3. Yu B and Chang J. Sig Transduct Target Ther 2020; 5: 236.
4. Halford B. C&EN 2021; 99: 44–45.
5. Sebastian D, Satishkumar S, Pradhan P, et al. J Org Chem 2022; 87: 18–39.
6. Bahrami Y and Franco C. Mar Drugs 2016; 14: 147.
7. Perkins JJ, Shurtleff VW, Johnson AM, et al. ACS Med Chem Lett 2021; 12: 662666.
8. Tranová L and Stýskala J. J Org Chem 2021; 86(19): 13265–13275.
9. Köckinger M, Hanselmann P, Roberge DM, et al. Green Chem 2021; 23: 2382–2390.
10. Hofle G, Steglich V and Vorbruggen H. Angew Chem Int Ed 1978; 17: 569.
11. Li YS, Zhang JJ, Mei LQ, et al. Org Prep Proced Int 2012; 44: 387–391.
12. Xia R, Xie MS, Niu HY, et al. Green Chem 2014; 16: 1077–1081.
13. Xia R, Niu HY, Qu GR, et al. Org Lett 2014; 16: 444–447.
14. Xia R, Sun LP, Yang XN, et al. J Chem Res 2015; 39: 335–367.
15. Xia R, Sun LP and Chen LS. J Org Chem 2016; 80: 13265–13275.
16. Mandolesi Sá M and Meier L. Synlett 2006; 2007.
17. Francon P, Janeba Z, Shibuya S, et al. J Org Chem 2002; 67: 6788–6796.
18. Čar Ž, Petrovič V and Tomič S. J Carbohydr Chem 2006; 25: 713–723.
19. Robins MJ and Uznański B. Can J Chem 1981; 59: 2601–2607.
20. Reist EJ, Calkins DF, Fisher LV, et al. J Org Chem 1968; 33: 1600–1603.
21. Xia R, Guo Z and Qin B. Chin J Org Chem 2014; 34: 1154–1160.
22. Bauman JG and Wirsching RC. WO 9412514 A1, 1994.
23. Lodyga-Chruscinska E, Oldziej S, Sochacka E, et al. J Inorg Biochem 2011; 105: 1212–1219.
24. Lanver A and Schmalz HG. Molecules 2005; 10: 508–515.
25. Robins MJ and Hatfield PW. Can J Chem 1982; 60: 547–553.
26. Kalayanov G, Jakša S, Scarica T, et al. Synthesis 2004; 12: 2026–2034.
27. Blackburn GM and Johnson AW. J Chem Soc 1960; 4347.
28. Barker GR and Foll GE. J Chem Soc 1957; 3794–3798.
29. Gilham PT and Khorana HG. J Am Chem Soc 1958; 80: 6212–6222.
30. Kumar V and Malhotra SV. Nucleic Nucleic Acids 2009; 28: 821–834.
31. El-Tayeb A, Qi A, Nicholas RA, et al. J Med Chem 2011; 54: 2878–2890.