Synthesis of Regioselectively Protected Forms of Cytidine Based on Enzyme-catalyzed Deacetylation as the Key Step

Atsuhito Kuboki,* Takashi Ishihara,* Eiko Kobayashi,* Hiromichi Ohta,* Takeshi Ishii,** Ayumu Inoue,** Satoshi Mitsuda,** Tatsuo Miyazaki,** Yasuhiro Kajihara,*** and Takeshi Sugai*†

*Department of Chemistry, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan
**Biotechnology Laboratory, Sumitomo Chemical Co. Ltd., 4-2-1 Takatsukasa, Takarazuka 665-0051, Japan
***Department of System Function, Yokohama City University, 22-2 Seto, Kanazawa-ku, Yokohama 236-0027, Japan

Received September 2, 1999; Accepted October 7, 1999

\[ N^4\text{-Acetylcytidine (77\%)} \] and \[ 2',3'-\text{O,N'-triacetylcytidine (95\%)} \] were obtained from the hydrolysis of a common precursor, the peracetylated form of cytidine with Aspergillus niger lipase (Amano A) and Burkholderia cepacia esterase (SC esterase S), respectively, under very mild conditions. The experimental procedure for the conversion of triacetylcytidine to a corresponding phosphoramidite (82\%), an intermediate for sugar nucleotide synthesis, is also elaborated.

Key words: cytidine; esterase; lipase; selective deacetylation; phosphorylation

The importance of the partially protected form of cytidine (1a) has been addressed in synthetic nucleoside and nucleotide chemistry. For example, the \( N^4\)-acylated form (1b) is the starting material for ddC, an antiviral drug.\(^1\)\(^-\)\(^4\) Another representative is the \( 2',3'-\text{O,N'-triacetylated form (1e)} \), whose liberated hydroxy group at the 5'-position is phosphorlated toward the synthesis of nucleotides.\(^5\)\(^-\)\(^7\) Extensive studies have so far been devoted to attempts at the regioselective introduction of protective groups.\(^8\)\(^-\)\(^16\) The protection/deprotection of aminopyrimidine nucleosides requires: 1) sufficiently mild conditions to allow the alkaline-labile \( N\)-acyl group (\( N^4\) position) to remain intact; and 2) as high selectivity as possible between the primary/secondary hydroxyl groups for preparation of the partially protected forms. Enzyme-catalyzed hydrolysis\(^20\)\(^-\)\(^21\) and acetylation\(^22\)\(^-\)\(^25\) seem attractive to meet these requirements. Indeed, the protease-catalyzed regioselective hydrolysis of acetylated forms of uridine and some related pyrimidine nucleosides has been reported.\(^21\) This report prompted us to investigate the enzyme-catalyzed hydrolysis of the peracetylated form (1d) of cytidine.

\[^1\] To whom correspondence should be addressed. Fax: +81-45-563-5967; E-mail: sugai@chem.keio.ac.jp

Materials and Methods

General procedures. Melting point (mp) data are uncorrected. Optical rotation values were recorded with a Jasco DIP 360 polarimeter. IR spectra were measured as films for oils and as KBr discs for solids with a Jasco FT/IR-410 instrument. \(^1\)H-NMR spectra were recorded at 270 MHz with a JEOL JNM-EX 270 or at 400 MHz with a JEOL JNM-EX 400 spec-
HPLC analysis of the cytidine derivatives (1a-f). The retention times of 1a-f by an HPLC analysis of all of the enzyme-catalyzed hydrolysis reactions were each first assigned as follows: column, Wakosil-II 5C18 (6.0 mmφ x 250 mm); eluent, acetonitrile-water (1:1) at 0.5 ml/min; detection, 254 nm; R, (min) 8.6 (1a); 8.9 (1b); 9.5 (1e and/or 1f); 10.7 (1c); 12.7 (1d).

The quantitative estimation of 1c and 1b was calculated by comparing their peak areas with that of 1a (254 nm), because of the low availability of any authentic samples other than 1a for constructing a calibration line prior to pursuing the reaction process. A standard solution of 1a was prepared by dissolving cytidine (1a, 10.5 mg, 0.043 mmol) in water and adjusting the total volume to 25 ml. An aliquot of a solution containing 1c was mixed with the standard solution of 1a (300 µl, containing 5.18 x 10^-4 mmol) and analyzed by HPLC. The amount of 1c (R, 10.7 min) was estimated by comparing the peak area with that of 1a (R, 8.6 min), estimation of the substrate for enzyme-catalyzed hydrolysis (1d) was performed in a similar manner.

The difference in R, between 1a and 1b was too small to estimate 1b, so another standard, benzamide (R, 14.0 min), was utilized as a liaison to overcome this problem. A solution was prepared by dissolving benzamide (5.2 mg) in water, and the total volume was adjusted to 25 ml. A 300-µl volume of this benzamide solution was mixed with the standard solution of 1a (300 µl, containing 5.18 x 10^-4 mmol, as already described) and analyzed by HPLC. The amount of benzamide (R, 14.0 min) was estimated by comparing the peak area with that of 1a (R, 8.6 min), and this benzamide solution was then used to estimate the amount of 1b.

2',3',5'-O,N'-Tetraacetylcytidine (1d). This was prepared according to the reported procedure with a slight modification. Sodium acetate (3.70 g, 44.8 mmol) was mixed with acetic anhydride (48.0 ml, 51.0 mmol), and the mixture was stirred at 120°C until the sodium acetate had been completely dissolved. To this mixture was added cytidine (1a, 15.0 g, 61.8 mmol) in two portions, keeping the reaction temperature at 120°C, and the mixture was stirred for 1 hr at that temperature. Ice-cooled water (150 ml) was then added, and the mixture was concentrated in vacuo. The residue was dissolved in water, and the solution was desalted. The resulting solution was revealed to contain 1d (24.9 g, 60.6 mmol), which was estimated as already described, the yield of 1d being 98%. This solution was employed for the next reaction without further purification. [α]D +49.4° (c 0.64, H2O; the concentration of 1d in the solution was estimated as described); 1H-NMR (400 MHz, D2O) δ: 2.16 (3H × 3, s), 2.25 (3H, s), 4.43 (1H, dd, J = 3.9, 12.5 Hz), 4.47 (1H, dd, J = 2.7, 12.5 Hz), 4.61 (1H, ddd, J = 2.7, 3.9, 5.8 Hz), 5.44 (1H, dd, J = 5.8, 5.8 Hz), 5.59 (1H, dd, J = 3.9, 5.8 Hz), 6.06 (1H, d, J = 3.9 Hz), 7.38 (1H, d, J = 7.6 Hz), 8.15 (1H, d, J = 7.6 Hz).

Screening of the enzymes. The candidates were Pseudomonas cepacia lipase (Amano PS), Candida rugosa lipase (Sigma L-1754), Bacillus licheniformis protease (subtilisin, Sigma P5380), Burkholderia cepacia lipase (Sumitomo SC lipase A),<sup>27</sup> B. cepacia lipase esterase (Sumitomo SC esterase S),<sup>28</sup><sup>29</sup> Candida antarctica lipase (Chirazyme L-2, Boehringer), and Aspergillus niger lipase (Amano A). A small portion of each enzyme (10-12 mg) was added to a solution of 1d (3% w/v, 1 ml) in a phosphate buffer (pH 7.2, 0.1 M), and the progress of the reaction was monitored by silica gel TLC. Values of the starting material and products were as follows: 1d, 0.80; 1c, 0.69; 1e and/or 1f, 0.57; 1b, 0.48, developed by ethyl acetate:2-propanol:water = 36:8:1.

Enzyme preparation from commercial Aspergillus niger lipase (Amano A). All of the procedures were carried out at 4°C, and the phosphate buffer (0.1 M, pH 7.2, 50 ml) was also pre-cooled at 4°C. Commercial Aspergillus niger lipase (Amano A, 1000 mg) was suspended in the buffer and filtered through a Nalgene disposable tissue culture filter unit (50 mmφ, 20 µm pore size). The resulting filtrate was passed through an Amicon YM10 ultrafiltration membrane (62 mmφ). When the volume of the solution inside the cell had reached 2 ml, the protein solution was diluted with the buffer solution (20 ml). This ultrafiltration procedure was conducted five times to remove the low-molecular-weight impurities from the enzyme protein. Finally, the protein solution was diluted with the buffer solution, and the volume was adjusted to 10 ml. The protein concentration was estimated by the CBB method, being calibrated with γ-globulin (Sigma G7516) as the standard. The enzyme solution thus obtained (from 100 mg of commercial lipase) contained 5.4 mg of protein, and was employed in the subsequent experiments.
Enzymatic Synthesis of Cytidine Derivatives

365

N'‐Acetylcytidine (1b). A solution of 1d (900 mg, 2.18 mmol, 21.6 ml total volume) was mixed with a phosphate buffer solution (0.0125 M, pH 7.2, 27 ml). To this mixture was added the enzyme preparation from Amano A (protein weight of 29.1 mg), and the mixture was stirred at 37°C, while keeping its pH at 7.0 by adding a 1 M NaOH aqueous solution. After 15.5 hr, the hydrolysis of almost all of the O-acetyl groups was confirmed by an HPLC analysis. The mixture was filtered through a Nalgene disposable tissue culture filter unit, and then through on Amicon YM10 membrane. The resulting filtrate was desalted and lyophilized to give a crude product (678 mg) which was recrystallized from hot water to give 1b (481 mg, 1.68 mmol, 77%) as colorless needles, mp 207–208°C [lit.10 mp 208–209°C]. The yield of 1b was estimated by an HPLC analysis. [α]D +23.7° (c 1.00, H2O); IR νmax cm⁻¹: 3473, 3263, 2927, 2844, 2360, 2339, 1717, 1643, 1582, 1491, 1432, 1399, 1307, 1229, 1194, 1132, 1073, 1049, 970, 938, 816, 788, 685, 653, 592; 1H-NMR (400 MHz, D₂O) δ: 2.21 (3H, s), 3.82 (1H, dd, J = 2.3, 12.0 Hz), 3.97 (1H, d, J = 12.0 Hz), 4.16 (2H, dd, J = 2.3, 4.2 Hz), 4.29 (1H, dd, J = 3.4, 4.2 Hz), 4.30 (1H, dd, J = 2.7, 3.4 Hz), 5.87 (1H, d, J = 2.7 Hz), 7.32 (1H, d, J = 7.6 Hz), 8.32 (1H, d, J = 7.6 Hz). The NMR spectrum was identical with that reported previously.13

Overexpression and purification of the esterase. Burkholderia cepacia esterase (Sumitomo SC esterase S) was produced by recombinant E. coli JM109/pAL108 which had been constructed by the standard gene cloning method. Plasmid pAL108 contained the esterase gene15 that had been isolated from the Burkholderia cepacia SC-20 strain (Sumitomo Chemical Co.) and was inserted in the expression vector of pUC119. E. coli JM109/pAL108 was grown on an LB medium [Tryptone (1%), yeast extract (0.5%), NaCl (1%), 3000 ml] supplemented with glycerol (30 g), ampicillin (150 mg), and an IPTG solution (1 mM, 150 μl) at 37°C for 15 hr in a jar-fermenter. The bacterial cells were harvested from the culture broth by centrifugation and re-suspended in a phosphate buffer solution (0.1 M, pH 7.0, 100 ml), before the cells were disrupted with a sonicator. Removal of the debris of disrupted cells from the suspension by centrifugation gave a solution of the cell-free extract. Finely powdered ammonium sulfate was added to the solution to 80% saturation. The resulting precipitate obtained by centrifugation was re-suspended in the phosphate buffer solution (0.1 M, pH 7.0, 50 ml), and dialyzed three times in 2000 ml each of the same buffer solution. The crude enzyme solution after dialysis was lyophilized to give a powdered preparation of the crude esterase, this being named SC esterase S (4.9 g). The protein content of this preparation was 77.4%.

2',3',5'-O,N'-Triacetylcytidine (1c). A solution of 1d (1000 mg, 2.43 mmol, 24 ml total volume) was diluted with water, and the total volume was adjusted to 40 ml. To this was added the enzyme preparation of SC esterase S as just described (protein weight, 30 mg), and the mixture was stirred at 47°C, while keeping its pH at 7.0 by adding a 1 M NaOH aqueous solution. After 1 hr, the progress of the hydrolysis was confirmed to be 97% by an HPLC analysis. The mixture was filtered through a Nalgene disposable tissue culture filter unit, and the resulting filtrate was desalted and lyophilized to give a crude product. This crude product was dissolved in pyridine, and the resulting solution was adsorbed to silica gel (1.5 g). This was evacuated by a vacuum pump for 2 hr. The powdery adsorbent was charged into the top of a column of silica gel (7.5 g, dispersed with ethyl acetate:methanol:water = 15:1:1), and eluted with a mixture of ethyl acetate:methanol = 15:1. Desired product 1c (850 mg, 2.31 mmol, 95%) was obtained as an amorphous solid. The yield of 1c was estimated by HPLC as already described. Mp 148–149°C (colorless fine needles, recrystallized from acetic acid–ether); [α]D +62.0° (c 0.70, H2O); IR νmax cm⁻¹: 3474, 3293, 3138, 2935, 2453, 2362, 1666, 1561, 1499, 1435, 1376, 1241, 1077, 912, 813, 788, 669, 597; 1H-NMR (400 MHz, D₂O) δ: 2.12 (3H × 2, s), 2.21 (3H, s), 3.82 (1H, dd, J = 4.2, 5.7 Hz), 3.93 (1H, dd, J = 2.8, 12.9 Hz), 4.40 (1H, ddd, J = 2.8, 4.2, 5.7 Hz), 5.36 (1H, dd, J = 5.6, 5.7 Hz), 5.52 (1H, dd, J = 4.2, 5.6 Hz), 6.09 (1H, d, J = 4.2 Hz), 7.35 (1H, d, J = 7.6 Hz), 8.32 (1H, d, J = 7.6 Hz). Anal. Found: C, 48.41; H, 5.42; N, 11.11%. Calcd. for C₁₅H₁₉O₅N₃: C, 48.78; H, 5.19; N, 11.38%.

2-Cyanoethyl 2',3',5'-O,N'-Triacetylcytidine-5'-yl N',N'-Diisopropyl phosphoramidite (1g). The experiment was carried out according to the reported procedure5 with a slight modification. DMF, acetonitrile, ethyl acetate had been dried over 4A molecular sieves prior to use in the subsequent experiment and workup procedure. A 30-ml two-necked reaction flask was charged with powdered 4A molecular sieves (200 mg) and flame-dried under vacuum. To this was added 1c (100 mg, 0.271 mmol) and then DMF (0.24 ml). After the substrate had been completely dissolved in DMF, acetonitrile (0.8 ml) was added, and the resulting mixture was stirred overnight under Ar. In addition, a flame-dried 30-ml conical flask was charged with acetonitrile (1.5 ml), disopropylamine (76 μl, 2 eq.), 1H-tetrazole (38 mg, 2 eq.), and 4A molecular sieves (granules, 200 mg) in this order, and the mixture was stirred overnight under Ar.

To the solution of 1c, the foregoing mixture of disopropylamine and 1H-tetrazole was transferred via a cannula, and the resulting mixture was stirred
for 10 min at room temperature. The mixture was then cooled to 0°C, and 2-cyanoethyl N,N,N',N' -tetraisopropylphosphoramidite (Sigma C8539, 181 µl, 95% purity, d 0.949, 2 eq.) was added while stirring. The reaction temperature was raised to room temperature, and stirring was continued for 2 hr. The disappearance of the starting material was confirmed by a thin-layer chromatographic analysis (silica gel; ethyl acetate:methanol = 7:1), and the reaction was quenched by adding a sodium hydrogen carbonate solution.

The mixture was extracted several times with distilled ethyl acetate. The combined extract was successively washed with a sodium hydrogen carbonate solution and brine, and then dried over anhydrous magnesium sulfate:potassium carbonate = 4:1. The mixture was filtered, and the filtrate was concentrated in vacuo to give 1g. This material was employed for the subsequent coupling reaction with carbohydrates. The content of 1g at this stage was estimated by diluting the residue with distilled and predried ethyl acetate, the total volume being adjusted to 25 ml. To a 10-ml portion of this solution was added benzhydryl (10 mg), and the mixture was stirred until it became a homogeneous solution. Benzhydryl worked as an inert internal standard for measuring the 1H-NMR spectrum, and no reaction between 1g and benzhydryl was apparent by this estimation procedure. A small portion of the resulting solution was concentrated in vacuo, and the 1H-NMR spectrum of the residue was measured in CDCl3. The content of desired product 1g was estimated by comparing the signals of δ 6.26 (0.5H, d, J = 4.9 Hz, assigned to the cytidine H-1' proton of the diastereomeric mixture of 1g) and δ 6.33 (0.5H, d, J = 5.9 Hz, H-1') with that of benzhydryl, δ 5.80 (1H, s). In the conversion of 1c to 1g, the yield of 1g was estimated to be 82%.

1H-NMR (400 MHz, CDCl3) δ: 2.02 and 2.03 (each 1.5H, s), 2.06 and 2.08 (each 1.5H, s), 2.22 (3H, s), 2.62 (1H, m, -OCH2CH2CN of one diastereomer), 2.68 (1H, d, J = 2.0 Hz, 5.6, 5.8 Hz, -OCH2CH2CN of the other diastereomer), 3.40-3.52 (1H, m, -OCH2CH2CN of one diastereomer), 3.52-3.65 (1H, m, -OCH2CH2CN of the other diastereomer), 3.76-4.03 (2H, H-5'), 4.17 (0.5H, dd, J = 3.4, 6.1 Hz, H-4'), 4.31 (0.5H, dd, J = 3.4, 5.4 Hz, H-4'), 5.27-5.34 (1H, H-2', H-3'), 5.35 (0.5H, dd, J = 3.9, 5.9 Hz, H-2'), 5.44 (0.5H, dd, J = 3.4, 4.9 Hz, H-3'), 6.26 (0.5H, d, J = 4.9 Hz, H-1'), 6.33 (0.5H, d, J = 5.9 Hz, H-1') 7.41 and 7.43 (each 0.5H, s), 10.10 (1H, br. s).

If starting material 1c had not been completely consumed, desired product 1g could be separated from unreacted starting material 1c by chromatographic separation of the crude reaction mixture. Rf values were 1g, 0.30; 1c, 0.15. For example, the crude mixture obtained on the scale just described was chromatographed with silica gel (50 g, suspended in a pre-mixed solvent of ethyl acetate:triethylamine = 100:1). The elution of 1g was achieved by the same solvent system.

Results and Discussion

The hydrolytic action of various enzymes were screened on the substrate (1d),20 the candidates being listed in the materials and methods section. Amano A was selected from them due to its highest activity toward all of the primary and secondary acetates in 1d.

The time-dependent profile for hydrolysis is shown in Fig. 2. It is clear that the hydrolysis of the primary acetate (5' position) was faster than that of secondary acetates (2' and 3' positions). Even after a prolonged incubation, no hydrolysis of the acetamide (N1 position) was apparent, although this C-N bond is concomitantly cleaved under conventional basic conditions as mild as a diluted aqueous ammonia solution for the removal of O-protecting acyl groups.

To facilitate product isolation, the removal of low-molecular-weight additives as stabilizers for the commercial preparation of the lipase was important prior to use. With this partially purified lipase (30 mg of protein), the hydrolysis of 1d (900 mg) was completed within 16 hr at 30°C under a controlled condition of pH 7.0. After removing the enzyme protein and subsequent desalting of the crude product, simple crystallization from hot water provided pure 1b (77%).

To prepare 1c, an esterase, whose gene had been

![Fig. 2. Time-Course Characteristics of the Hydrolysis of 1d with Aspergillus lipase (Amano A). ○ 1b; △ 1c; ■ 1d; □ 1e and/or 1f.](image-url)
cloned from *B. cepacia* and overexpressed in *E. coli* was the best choice as the selective enzyme working on the primary acetate (5' position). For example, by treating 1d (500 mg) with esterase (5 mg) at 37°C for 1.5 hr, the hydrolysis of the primary acetate was accomplished, and desired 1c (91%) was obtained, together with overhydrolyzed 1e and 1f (total 9%). The resulting mixture was salted out, and the desired product (1e) was precipitated. Although the purity of 1e was high, the recovery was as low as 37%. Our work at this stage was focused upon the elaboration of more regioselective reaction conditions and an improved purification procedure.

As has been suggested in the some reports, the reaction was attempted at a low temperature (4°C); however, the selectivity between the primary and secondary acetates was apparently reduced. This was probably due to the hydrolysis of primary acetate being too slow, compared with that at the other undesired positions. This observation was based on Sakai’s quantitative estimation of the change of selectivity which depends on the reaction temperature. In contrast, the enzyme worked very efficiently immediately after starting the incubation with 1d at a temperature as high as 47°C. At this temperature, when almost all of 1d had been hydrolyzed to triacetate 1c, the activity reached nearly zero, due to denaturation of the enzyme. The HPLC analysis showed no overhydrolyzed products such as 1e and/or 1f in the reaction mixture. Subsequent removal of the debris of the denatured enzyme and desalting provided the desired product in a crude state in a 97% yield. This material could be purified by a silica gel column chromatography. A dry charge of the material at the top of silica gel column was essential for high recovery (98%) and high purity of the desired product. It has been suggested that even a small amount of such contaminants as water and inorganic material would have a deleterious effect on the phosphorylation of 1c by the phosphoramidite method. Our purified 1c worked very well as the precursor of 1g. Starting material 1c was carefully dehydrated in the reaction flask with 4A molecular sieves prior to the phosphorylation reaction, and the yield of 1g was as high as 82%.

In conclusion, N4-acetylcytidine (1b, 77%) and 2',3'-O,N4-triacetylcytidine (1c, 95%) were obtained from the hydrolysis of a common precursor (1d) with *Aspergillus niger* lipase (Amano A) and *Burkholderia cepacia* esterase (SC esterase S), respectively, under very mild conditions.

**Acknowledgments**

The authors thank Amano Pharmaceutical Ltd. and Roche Diagnostics for generously presenting the enzymes at the stage of screening. This work was part of the “Science and Technology Program on Molecules, Supra-Molecules and Supra-Structured Materials” of the Academic Frontier Promotional Project by Japan’s Ministry of Education (Monbusho) and is acknowledged with thanks. This work was also supported by the “Research for the Future” Program (JSPS-RFTF 97100302) from Japan Society for the Promotion of Science and Kato Bioscience Foundation, which is acknowledged with thanks. The authors thank Mr. Atsushi Otake of this laboratory for his literature survey.

**References**


11) Ogilvie, K. K., Shifman, A. L., and Penny, C. L., The Synthesis of Oligoribonucleotides. III. The Use of Silyl Protecting Groups in Nucleoside and


