

# Development of a Universal Nucleobase and Modified Nucleobases for Expanding the Genetic Code

This unit presents protocols for the synthesis and characterization of nucleosides with unnatural bases in order to develop bases for the expansion of the genetic alphabet or for nonselective pairing opposite natural bases.

The faithful pairing of nucleobases through complementary hydrogen-bond (H-bond) donors and acceptors forms the foundation of the genetic code. However, there is no reason to assume that the requirements for duplex stability and replication must limit the genetic alphabet to only two base pairs, or, for that matter, hydrogen-bonded base pairs. Expansion of this alphabet to contain a third base pair would allow for the encoding of additional information and would make possible a variety of in vitro experiments using nucleic acids with unnatural building blocks. Previous efforts to generate orthogonal base pairs have relied on H-bonding patterns that are not found with the canonical Watson-Crick pairs. However, in all cases, the unnatural bases were not kinetically orthogonal, and instead competitively paired with natural bases during polymerase-catalyzed DNA synthesis (Horlacher et al., 1995; Lutz et al., 1996, 1998a,b). Tautomeric isomerism, which would alter H-bond donor and acceptor patterns, likely contributes to this kinetic infidelity (Roberts et al., 1997a,b; Robinson et al., 1998; Beaussire and Pochet, 1999). An alternative strategy is centered around developing unnatural bases that form pairs based not on hydrogen bonds, but rather on interbase hydrophobic interactions. Such hydrophobic bases should not pair stably opposite natural bases due to the forced desolvation of the purines or pyrimidines. Additionally, the use of nucleobase analogs, which are not restricted to the shape of H-bonding topologies of natural bases, allows for the use of a wider range of analogs. This unit describes the design, synthesis, and characterization of unnatural base pairs involving 1- $\beta$ -D-2-deoxyriboseyl-*N*- and -*C*-nucleosides.

To be reasonable candidates for the modification of the genetic code, unnatural nucleosides must meet certain criteria. First, the unnatural bases must pair stably and selectively in duplex DNA. Second, the unnatural bases must be good substrates for DNA polymerases, being replicated with good efficiency and fidelity. Determination of these thermodynamic and kinetic parameters of the unnatural nucleosides is accomplished by incorporation into oligonucleotides and subsequent evaluation as described herein.

A general procedure for the synthesis of 1- $\beta$ -D-2-deoxyriboseyl-*N*-nucleosides containing pyrimidine-like unnatural hydrophobic bases using a modified Silyl-Hilbert-Johnson reaction (Hilbert and Johnson, 1930), namely Vorbrüggen glycosylation (Niedballa and Vorbrüggen, 1970), is first described (see Basic Protocol 1). This methodology is very useful for the synthesis of pyrimidine-like hydrophobic base pairs, because under the necessary reaction conditions the polar (but otherwise often rather insoluble) pyrimidine bases are readily converted by silylation into lipophilic silyl compounds, which are then soluble in organic solvents. Thus, the homogeneous coupling with sugar moieties is permitted.

Next, a general procedure is included for the synthesis of 1- $\beta$ -D-2-deoxyriboseyl-*N*-nucleosides containing purine-like bases using the metal salt procedure (see Basic Protocol 2), which entails the condensation of nucleobase sodium salts with a sugar halide (Kazimierczuk et al., 1984). The sodium salts of acidic heterocyclic systems such as imidazole, purine, triazole, and pyrazole are prepared in situ with NaH or analogous bases.

**BASIC  
PROTOCOL 1**

**Development of a  
Universal  
Nucleobase and  
Unnatural  
Nucleobases**

**1.5.2**

The synthesis of 3,5-dimethylphenyl-*C*-nucleoside by (1) condensation of in situ-generated Grignard reagent from 1-bromo-3,5-dimethylbenzene with 1- $\alpha$ -chloro-3,5-di-*O*-toluoyl-2-deoxyribofuranose and (2) methoxide-mediated deprotection of the bistoluoyl groups is also described (see Basic Protocol 3).

In addition, the unit includes a procedure for the synthesis of 1,4-dimethylnaphthalene-*C*-nucleoside (see Basic Protocol 4). The sugar precursor containing the aldehyde group is used; this is synthesized from 2-deoxyribose in seven steps (Eaton and Millican, 1988) and does not require the generation of any stereocenters beyond C1'. The condensation of the aryllithium reagent of 2-bromo-1,4-dimethylnaphthalene with the aldehyde, followed by in situ cyclization and deprotection of hydroxyl groups, affords the desired nucleoside.

An Alternate Protocol describes the synthesis of 3-methyl-2-naphthalene-*C*-nucleoside. A different sugar precursor containing an aldehyde group at C3' is generated from Felkin-Anh addition of an allylzinc nucleophile to isopropylidene-protected glyceraldehydes with high diastereoselectivity (Solomon and Hopkins, 1993). Mesylation, followed by treatment with excess trifluoroacetic acid, yields a diastereomeric mixture of separable *C*-nucleosides.

The purification of DNA containing these unnatural bases is then presented (see Basic Protocol 5). Purified DNA products are used to characterize the unnatural base pairs as candidate pairs for expanding the genetic code. Finally, the thermodynamic (see Basic Protocol 6) and kinetic characterization (see Basic Protocol 7) of the unnatural bases are described.

**NOTE:** All glassware used for reactions should be evacuated, flame-dried, and flushed with argon before use. All operations should be carried out in a well-ventilated fume hood. All reactions involving moisture-sensitive reagents should be performed under argon atmosphere (see Reagents and Solutions).

**NOTE:** It is recommended that all products in this unit be stored in a desiccator at  $\leq 4^{\circ}\text{C}$ .

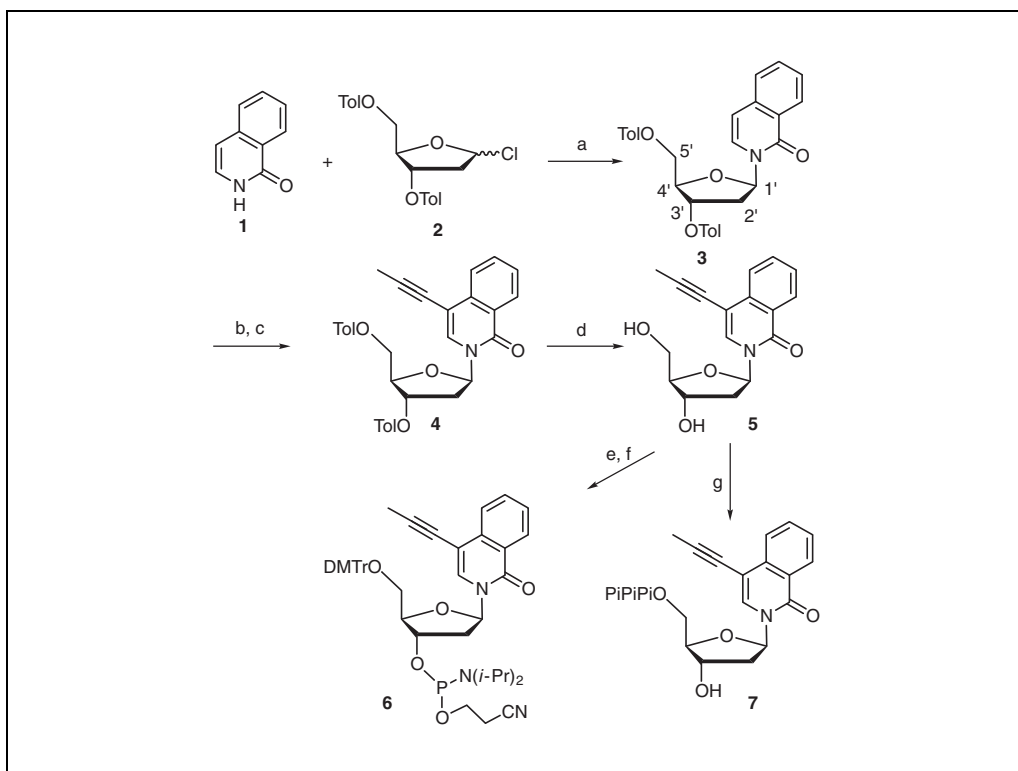
**GENERAL GUIDELINES FOR SYNTHESIS OF PYRIMIDINE-LIKE  
1- $\beta$ -D-2-DEOXYRIBOSYL-*N*-NUCLEOSIDES**

This protocol outlines a general procedure for the synthesis of unnatural hydrophobic pyrimidine-like *N*-nucleosides (Figs. 1.5.1 and 1.5.2; McMinn et al., 1999). Specific protocols are given for the synthesis of 7-propynylisocarbostyryl (PICS) triphosphate (**S.7**), as well as the phosphoramidite of PICS (**S.6**). Synthesis of the other unnatural nucleosides shown in Figure 1.5.2 can be accomplished using similar synthetic transformations, starting from the appropriate pyrimidine.

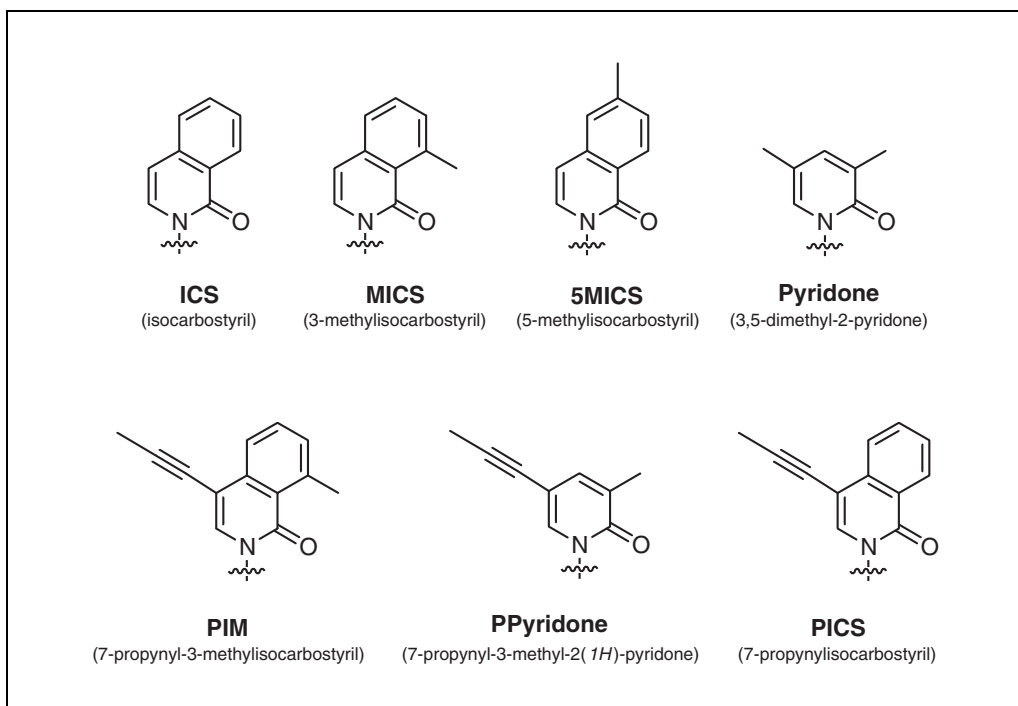
The introduction of propynyl groups, the deprotection of toluoyl groups, and the formation of the phosphoramidite and triphosphate described here are generally applied to all synthesis protocols in this unit.

**Materials**

Argon (see recipe)  
Isocarbostyryl (**S.1**; Aldrich) or pyrimidine of choice  
Acetonitrile, anhydrous  
*N,O*-Bis(trimethylsilyl)acetamide (Aldrich)  
Bis-toluoyl-protected chloroglycoside:  
1- $\alpha$ -chloro-3,5-di-*O*-toluoyl-2-deoxyribofuranose (**S.2**; Berry & Associates;  
Takeshita et al., 1987)



**Figure 1.5.1** General procedure for synthesis of unnatural hydrophobic pyrimidine-like *N*-nucleosides (see Basic Protocol 1). Reagents: **(A)** Bis-acetamide,  $\text{SnCl}_4$ ,  $\text{CH}_3\text{CN}$ ,  $0^\circ\text{C}$  to room temperature (steps 1 to 22); **(B)**  $\text{ICl}$ ,  $\text{CH}_2\text{Cl}_2$ , reflux to room temperature (steps 23 to 30); **(C)** propyne,  $(\text{Ph}_3\text{P})_2\text{PdCl}_2$ ,  $\text{CuI}$ , TEA,  $-78^\circ\text{C}$  to room temperature (steps 31 to 43); **(D)** 0.5 M sodium methoxide, methanol, room temperature (steps 44 to 49); **(E)** DMTrCl, pyridine, room temperature (steps 50 to 62); (f) 2-cyanoethyl-diisopropylchlorophosphoramidite, DIPEA,  $\text{CH}_2\text{Cl}_2$ ,  $0^\circ\text{C}$  (steps 63 to 71); (g) tetrabutylammonium pyrophosphate,  $\text{POCl}_3$ , tributylamine, Proton-Sponge, trimethyl phosphate,  $0^\circ\text{C}$  (steps 72 to 80).



**Figure 1.5.2** Unnatural hydrophobic *N*-nucleosides (see Basic Protocol 1).

$\text{SnCl}_4$ , anhydrous, freshly distilled in vacuo  
 Ethyl acetate  
 Hexanes  
 Ammonium molybdate solution (see recipe)  
 Saturated sodium bicarbonate ( $\text{NaHCO}_3$ ) solution  
 Saturated sodium chloride ( $\text{NaCl}$ ) solution  
 Sodium sulfate ( $\text{Na}_2\text{SO}_4$ ), anhydrous  
 Silica gel, 200 to 400 mesh, 60 Å  
 Dichloromethane ( $\text{CH}_2\text{Cl}_2$ ), freshly distilled from calcium hydride  
 Ethyl ether, anhydrous  
 Iodine monochloride ( $\text{ICl}$ ; Aldrich; packaged under nitrogen in Sure/Seal vials)  
 Saturated sodium thiosulfate ( $\text{Na}_2\text{S}_2\text{O}_3$ ) solution  
 Magnesium sulfate ( $\text{MgSO}_4$ ), anhydrous  
 Triethylamine (TEA), freshly distilled from calcium hydride  
 Dichlorobis(triphenylphosphine) palladium(II) [ $(\text{Ph}_3\text{P})_2\text{PdCl}_2$ ] (Aldrich)  
 Copper(I) iodide ( $\text{CuI}$ )  
 Dry ice/ethyl ether bath ( $-100^\circ\text{C}$ ) and dry ice/acetone bath  
 Propyne  
 Sodium methoxide (Aldrich)  
 Methanol, anhydrous, distilled from magnesium turnings and stored over 3A  
 molecular sieves  
 Ammonium chloride ( $\text{NH}_4\text{Cl}$ )  
 Pyridine, anhydrous, distilled and stored on NaOH protected from light  
 4,4'-Dimethoxytrityl chloride (DMTr-Cl; Aldrich)  
 Diisopropylethylamine (DIPEA), freshly distilled from calcium hydride  
 2-Cyanoethyl diisopropylchlorophosphoramidite (Aldrich)  
 Trimethyl phosphate  
 Proton-Sponge (Aldrich)  
 Phosphorous oxychloride ( $\text{POCl}_3$ ), freshly distilled  
 Tributylamine  
 Tetrabutylammonium pyrophosphate (TBAP), stored over Drierite  
 1 M triethylammonium bicarbonate (TEAB), pH 7  
 10-, 25-, and 100-mL two-neck round bottom flasks with 14/20 joints, oven-dried  
 Rubber septa  
 Vacuum system (oil pump) with manifold and cold trap  
 Silica gel 60 F<sub>254</sub> alumina-backed thin-layer chromatography (TLC) plates (Fisher)  
 Heat gun  
 UV light source  
 100-, 250-, and 500-mL separatory funnels  
 300-mL round bottom flask  
 Cotton  
 Glass funnel  
 Rotary evaporator (Büchi) equipped with a dry ice condenser and a vacuum system  
 Heavy-walled glass columns (1.5-cm i.d.; 10-, 15-, and 20-cm length) with glass  
 adapters attached to compressed air or nitrogen source (see flash  
 chromatography steps in APPENDIX 3E)  
 Sea sand  
 Reflux condenser with 14/20 joint  
 23- and 18-G needles  
 100-mL three-neck flask with 14/20 joints  
 Propyne inlet adapter (Aldrich, cat. no. Z41.577-4)  
 Teflon caps for 14/20 joints

Speedvac evaporator (Savant)  
High-performance liquid chromatography (HPLC) system  
Glass or disposable plastic syringes  
10-mL test tubes  
Oil bath with temperature controller

Additional reagents and equipment for thin-layer chromatography (TLC; *APPENDIX 3D*) and flash chromatography (*APPENDIX 3E*)

**NOTE:** Glass or disposable plastic syringes are generally used for addition of liquid reagents. The 10-mL test tubes are used for collection of eluting fractions in column chromatography. Oil baths with temperature controllers are used for heating reactions as necessary.

***Perform condensation of isocarbostyryl (S.1) with bis-toluoyl-protected chloroglycoside (S.2)***

1. Place a magnetic stir bar in an oven-dried 100-mL two-neck flask with two rubber septa, and place the flask on top of a magnetic stir plate.
2. Evacuate reaction flask on vacuum line, then flush with argon. Repeat this procedure three times and attach the flask to an argon line on the manifold.
3. Quickly remove the rubber septum on one of the side inlets, add 0.50 g (3.4 mmol) isocarbostyryl (**S.1**) to the flask under argon, then immediately reinsert the septum.
4. Transfer 10 mL acetonitrile to the flask under argon.
5. Add, in dropwise fashion, 0.85 mL (3.4 mmol) *N,O*-bis(trimethylsilyl)acetamide and stir vigorously for 40 min at room temperature.

*During this time the suspension is cleared.*

6. Add an additional 12 mL of acetonitrile to the clear reaction mixture, followed by 1.10 g (2.80 mmol) bis-toluoyl-protected chloroglycoside (**S.2**). Cool the reaction mixture in an ice-water bath.
7. Slowly add 0.18 mL (2 mmol)  $\text{SnCl}_4$  and continue to stir under argon at 0°C.

*Preparation of S.3 uses 1.0 eq S.1, 0.8 eq S.2, 1.0 eq N,O-bis(trimethylsilyl)acetamide, 0.6 eq  $\text{SnCl}_4$ , and 129.4 eq acetonitrile. Increasing the quantity of  $\text{SnCl}_4$  from catalytic to substoichiometric amounts may improve yield.*

8. Monitor the progress of the reaction by analytical TLC (*APPENDIX 3D*) as follows:
  - a. Occasionally withdraw a small sample (1 to 5  $\mu\text{L}$ ) using a capillary tube and spot on silica gel 60 F<sub>254</sub> alumina-backed plates.
  - b. Develop the plate (*APPENDIX 3D*) with 4:1 (v/v) hexanes/ethyl acetate.
  - c. Visualize by dipping plate into ammonium molybdate solution and heating with a heat gun. Examine with UV light source.

*For S.3,  $R_f = 0.4$  (see *APPENDIX 3D*).*

*To analyze the course of reactions by TLC, the anisaldehyde/sulfuric acid staining system (6 g of *p*-anisaldehyde, 50 mL of absolute ethanol and 2.5 mL of concentrated  $\text{H}_2\text{SO}_4$ ) is useful for nucleosides. For nucleosides containing a DMTr group, 3% TFA/ $\text{CH}_2\text{Cl}_2$  is also effective.*

*Ammonium molybdate is used to assay the presence of carbohydrate derivatives and is sufficient for visualization of S.3. Spots generally appear in various shades of purple and blue.*

### Work up and purify S.3

9. When TLC analysis indicates the reaction is complete, transfer the mixture to a 500-mL separatory funnel and add 250 mL ethyl acetate.
10. Extract twice, each time with 50 mL saturated  $\text{NaHCO}_3$  solution, then extract once with 100 mL saturated  $\text{NaCl}$  solution.
11. Insert cotton into a glass funnel and add a 2-cm-thick layer of anhydrous  $\text{Na}_2\text{SO}_4$  over the cotton. Filter the organic layer by gravity into a 300-mL round-bottom flask.
12. Remove solvents under reduced pressure with a rotatory evaporator equipped with a dry ice condenser.

*It is recommended that the bath temperature not exceed 45°C.*

13. Pack a  $1.5 \times 20$ -cm heavy-walled column with 60 g of 200 to 400 mesh silica gel in hexanes for flash column chromatography (APPENDIX 3E).

*Gas pressure (air or nitrogen) is applied via a glass adapter to the top of the column through the gas inlet to increase the flow rate to about 5 drops/sec.*

14. Dissolve the crude product in 1 mL of dichloromethane.
15. Transfer the solution using a 3-mL glass pipet onto the silica gel and let the solution sink into the column bed.
16. Rinse the flask with two 1-mL aliquots of dichloromethane, add each rinse to the surface of silica gel, and allow the rinse to enter the column.
17. Carefully add a 0.5-cm layer of sea sand to the top of the silica gel bed.

*This step is optional. Keeping the surface of the silica gel flat is important for good separation.*

18. Elute with a gradient of 8:1 to 4:1 (v/v) hexanes/ethyl acetate.

*Use a glass cylinder to freshly mix the solvents.*

19. Collect 10-mL fractions and analyze by TLC (see step 8).

*To achieve the best separation, TLC using 5:1 hexanes/ethyl acetate should give an  $R_f$  value of ~0.35 for the possible product.*

20. Transfer fractions containing pure product to a 300-mL round-bottom flask and evaporate to dryness with a rotary evaporator.
21. Add 3 mL of anhydrous ethyl ether to triturate the product. Remove the solvent and evacuate the product until the weight of the flask is not changed.
22. Confirm the desired product by  $^1\text{H}$  NMR, NOESY, and COSY analysis (see, e.g., Lambert et al., 1998).

*Flash chromatography gives a pair of diastereoisomers as the  $\alpha$  and  $\beta$  anomers. The desired product is the faster-migrating  $\beta$  anomer, which is a white foam. The assignment of  $\beta$ -stereochemistry at  $\text{C1}'$  for each nucleoside is based on NOESY data, in which  $\text{H1}'$  shows cross-peaks with both  $\alpha\text{-H4}'$  and  $\alpha\text{-H2}'$  (see Fig. 1.5.1).*

*The optimal yield of  $1'\text{-}\beta\text{-3',5'-O-toluoyl-2'-deoxyribosylisocarbostyryl}$  (S.3) is 48%.*

### Iodinate S.3 by Friedel-Crafts iodination

23. Assemble a 25-mL oven-dried two-neck flask with a 14/20 reflux condenser in the center neck, a rubber septum on the side neck, and a magnetic stirring bar. Maintain a smooth flow of argon using a 23-G needle inserted through another septum on top of the reflux condenser.

24. Mix 0.40 g (0.80 mmol) **S.3** (from step 21) with 5 mL freshly distilled dichloromethane ( $\text{CH}_2\text{Cl}_2$ ) and transfer solution to the flask under argon atmosphere.
25. Add, in a dropwise fashion, 0.96 mL of 1 M  $\text{ICl}$  in  $\text{CH}_2\text{Cl}_2$  (0.96 mmol  $\text{ICl}$ ).
26. Reflux the reaction mixture 1 min at  $45^\circ\text{C}$ , then cool to room temperature. Perform TLC analysis on the mixture as in step 8.

CAUTION: *Iodine monochloride is highly toxic and volatile and should be handled with care in a well-ventilated hood.*

#### **Work up the iodinated product**

27. When TLC analysis indicates that all of the starting material has disappeared during the cooling period, add 10 mL saturated  $\text{NaHCO}_3$  followed by dropwise addition of saturated  $\text{Na}_2\text{S}_2\text{O}_3$  until the reaction solution is clear.
28. Transfer the mixture to a 100-mL separatory funnel. Extract the aqueous layer three times, each time with 20 mL dichloromethane.
29. Combine the organic extracts and dry over anhydrous  $\text{MgSO}_4$ .
30. Repeat steps 11 and 12 and attach a vacuum line to dry the residue.

*The resulting crude iodinated product (not shown in the figures) is pure enough for use in the next step. It should be kept in darkness and used immediately.*

#### **Introduce propynyl group by Sonogashira coupling**

31. Equip a 100-mL, 14/20 three-neck flask with rubber septa on the two side necks, a 14/20 dry-ice condenser-trap in the center neck, and a magnetic stir bar. Place a gas inlet adapter on top of the condenser and attach it to an argon line on the manifold.
32. Evacuate the apparatus and then purge it with argon three times to remove air.
33. Transfer ~0.80 mmol of the iodinated compound and 15 mL triethylamine (TEA) to the flask and quickly add 0.01 g (0.02 mmol)  $(\text{Ph}_3\text{P})_2\text{PdCl}_2$  and 0.01 g (0.06 mmol)  $\text{CuI}$  under argon.

*$(\text{Ph}_3\text{P})_2\text{PdCl}_2$  and  $\text{CuI}$  are highly sensitive to oxygen and light, respectively. They should be handled very carefully and quickly. High quality of these reagents is essential to the success of Sonogashira coupling reactions.*

34. Cool the reaction vessel in dry ice/diethyl ether bath ( $-100^\circ\text{C}$ ).
35. Fill the condenser-trap with pulverized dry ice.
36. Quickly remove the rubber septum on one of the side inlets and insert a propyne inlet adapter.
37. Purge the flask with propyne for 10 min or until final volume is 50 mL.

*Preparation of **S.4** uses 1.0 eq **S.3**, 1.2 eq  $\text{ICl}$ , 116.2 eq  $\text{CH}_2\text{Cl}_2$ , 0.07 eq  $\text{CuI}$ , 0.02 eq  $(\text{Ph}_3\text{P})_2\text{PdCl}_2$ , 134.5 eq TEA, and >120 eq propyne.*

38. Turn off the propyne cylinder. Quickly remove the rubber septa and the propyne inlet, and seal inlets tightly with Teflon caps.
39. Remove dry-ice condenser and seal the inlet tightly with a Teflon cap. Continue to stir at room temperature for 5 hr.

### Work up and purify **S.4**

40. Cool the reaction mixture in dry ice/acetone bath for 5 min, then remove the cold bath.
41. Remove the caps carefully to vent the remaining propyne at ambient temperature. Evaporate the solvent with a rotary evaporatory.
42. Purify the crude product by flash chromatography (steps 13 to 19) using a 1.5 × 15-cm silica gel column packed with 30 g silica gel. Elute with a gradient of 5:1 to 3:1 (v/v) hexanes/ethyl acetate. Perform TLC on the eluate (see step 8) using 4:1 hexanes/ethyl acetate.

*For **S.4**,  $R_f = \sim 0.35$ .*

43. Combine appropriate fractions based on TLC analysis and evaporate to dryness.

*Flash column chromatography yields 0.11 g (64%) of **S.4** as a white foam.*

### Deprotect 3'- and 5'-OH groups

44. Equip a 10-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G needle threaded through the septum.
45. Add 0.10 g (0.20 mmol) **S.4** (from step 43) and 4 mL anhydrous methanol sequentially to the flask.
46. Add dropwise 1.2 mL of 0.5 M sodium methoxide in methanol (0.60 mmol sodium methoxide) to the reaction mixture and continue to stir under argon for 20 min.

*Preparation of **S.5** uses 1.0 eq **S.4**, 3 eq sodium methoxide, and 617 eq methanol.*

### Work up and purify **S.5**

47. Add ~30 mg  $\text{NH}_4\text{Cl}$  to quench the reaction and keep stirring for 10 min.
48. Evaporate solvents and purify the product by flash chromatography (steps 13 to 19) using a 1.5 × 10-cm silica gel column packed with 20 g silica gel. Elute with a gradient of 2:100 to 5:100 (v/v) methanol/ $\text{CH}_2\text{Cl}_2$ . Perform TLC on the eluate (see step 8) using 4:100 methanol/ $\text{CH}_2\text{Cl}_2$ .

*For **S.5**,  $R_f = \sim 0.35$ .*

49. Combine appropriate fractions based on TLC analysis and evaporate to dryness.

*The reaction is approximately quantitative and the product, 1- $\beta$ -2'-deoxyribosyl-7-propynylisocarbostyryl (**S.5**), is obtained as a white solid.*

### Tritylate 5'-OH by Williamson etherification

50. Equip an oven-dried two-neck 10-mL flask with a stir bar and two rubber septa. Keep the vessel under argon using an 18-G needle threaded through the septum.
51. Add 0.046 g (0.154 mmol) **S.5** (from step 49) and 4 mL anhydrous pyridine to the flask.
52. Using an 18-G needle, insert a vacuum line through one of the septa to remove most of the pyridine by vacuum evaporation.
53. Purge the reaction flask with argon and remove the vacuum line immediately.
54. Transfer 0.7 mL anhydrous pyridine to the flask to dissolve **S.5**.



55. Prepare a solution of 0.08 g (0.23 mmol) 4,4'-dimethoxytrityl chloride in 0.3 mL pyridine and add in a dropwise fashion to the reaction mixture over a period of 20 min.

*Tritylation uses 1.0 eq S.5, 1.5 eq DMTr-Cl, and 82.3 eq pyridine.*

56. Continue to stir at room temperature for an additional 20 min.

**Work up and purify tritylated nucleoside**

57. Combine the reaction mixture with 20 mL of ethyl acetate, 5 mL saturated NaHCO<sub>3</sub>, and 5 mL saturated NaCl to partition.
58. Transfer the mixture to a 250-mL separatory funnel. Rinse the reaction vessel twice, each time with 5 mL of ethyl acetate, and transfer to the funnel.
59. Separate the organic layer and extract the aqueous layer twice, each time with 10 mL of ethyl acetate.
60. Filter and dry the organic layer as in steps 11 and 12.
61. Purify the crude product by flash chromatography (steps 13 to 19) using a 1.5 × 10-cm silica gel column packed with 20 g silica gel. Elute with a gradient of 5:1 to 3:1 (v/v) hexanes/ethyl acetate. Perform TLC on the eluate (see step 8) using 3.5:1 hexanes/ethyl acetate.

*For the DMTr-protected product of the above reaction,  $R_f = -0.35$ .*

62. Combine appropriate fractions based on TLC analysis and evaporate to dryness.

*Chromatographic purification affords the DMTr-protected product (0.073 g, 79%) as a white foam.*

**Phosphitylate to give phosphoramidite S.6**

63. Equip a 10-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G needle threaded through the septum.
64. Prepare a solution of 0.073 g (0.121 mmol) tritylated product (from step 62) in 1.3 mL of freshly distilled dichloromethane and transfer to the flask.
65. Add 0.084 mL (0.48 mmol) DIPEA to the solution and cool in ice-water bath.
66. Add, in a dropwise fashion, 0.041 mL (0.16 mmol) 2-cyanoethyl diisopropylchlorophosphoramidite and allow the reaction mixture to reach ambient temperature over 15 min.

*Relative to the starting nucleoside (S.5 in step 51), preparation of S.6 uses 3.2 eq DIPEA, 1.1 eq 2-cyanoethyldiisopropylchlorophosphoramidite, and 135.2 eq dichloromethane.*

**Work up and purify S.6**

67. Dilute the reaction mixture with 50 mL dichloromethane. Transfer the reaction mixture to a 100-mL separatory funnel.
68. Wash with 5 mL saturated NaHCO<sub>3</sub> and 5 mL saturated NaCl.
69. Filter and dry as in steps 11 and 12.
70. Perform flash chromatography (steps 13 to 19). Elute with a gradient of 6:1 to 4:1 (v/v) hexanes/ethyl acetate. Perform TLC (see step 8) using 5:1 hexanes/ethyl acetate.

*For S.6,  $R_f = -0.35$ .*

71. Combine appropriate fractions based on TLC analysis and evaporate to dryness using a rotary evaporator.

*The phosphoramidite product (S.6) is a white foam (63 mg), prepared with 65% yield, and is stored in a desiccator at −20°C.*

*Phosphoramidite compounds are used for automated DNA synthesis; isolation and desalting of oligonucleotides is carried out by PAGE and electrophoretic dialysis (see Basic Protocol 5). Oligonucleotides are used for duplex oligonucleotide denaturation temperature measurements and kinetic studies of DNA extension (see Basic Protocols 6 and 7).*

#### **Prepare triphosphate S.7**

72. Equip a 10-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G needle threaded through the septum.
73. Quickly remove the septum on one of the side inlets, add 0.008 g (0.027 mmol) **S.5** (from step 49) under argon, and insert the rubber septum back to the flask.
74. Coevaporate moisture and methanol twice from 0.1 mL anhydrous pyridine.
75. Add 0.134 mL (0.23 mmol) trimethyl phosphate and 9 mg (0.04 mmol) Proton-Sponge, dissolve completely, and bring to 0°C.
76. Add 0.003 mL (0.032 mmol) POCl<sub>3</sub> and continue to stir at 0°C for 2 hr.
77. Add 0.042 mL (0.03 mmol) tributylamine and 0.023 g (0.040 mmol) TBAP and stir for 1 min.

#### **Work up and purify S.7**

78. Add 2.7 mL of 1 M TEAB, pH 7.
79. Stir the reaction mixture for 10 min and concentrate to dryness in a Speedvac evaporator.
80. Dissolve the crude product in 200 µL of 1:1 (v/v) DMSO/isopropanol and purify by HPLC (UNIT 10.5) using the following conditions:

Column: Rainin C18-Dynamax 60 Å column (7.8 × 300 mm)

Buffer A: 0.1 M TEAB, pH 7

Buffer B: Acetonitrile

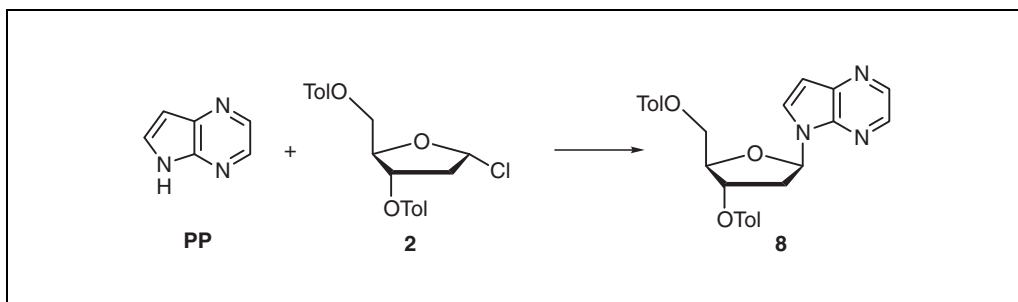
Gradient: 98% to 90% buffer A over 5 min, 90% to 75% buffer A over 25 min, 75% to 0% buffer A over 5 min, return to 98% A over 5 min.

Flow rate: 10 mL/min.

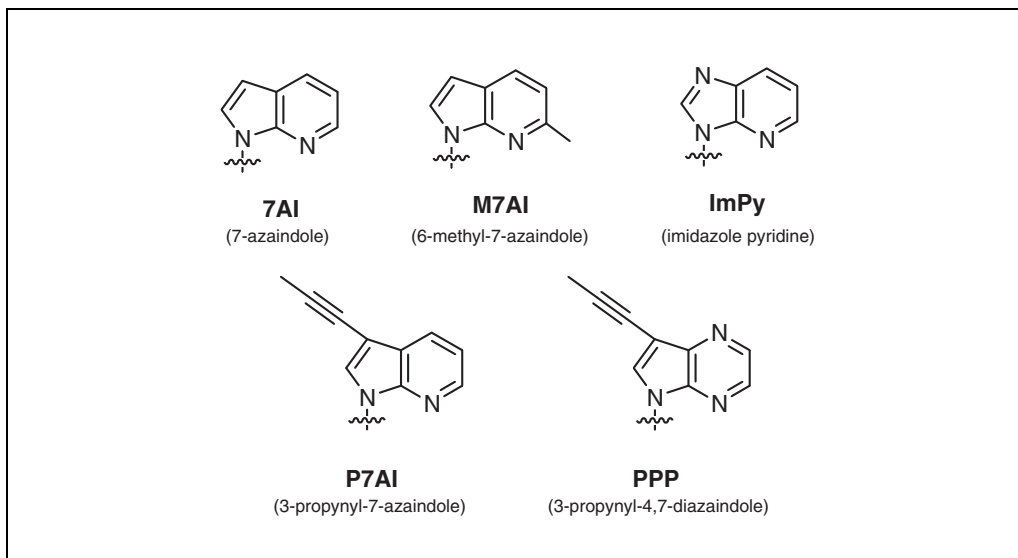
*Approximately 1 mg of the product triphosphate (S.7) is yielded with a retention time of 27.5 min, where the gradient is ~65% buffer A. The triphosphate compound could be used directly for kinetic studies of DNA incorporation.*

### **GENERAL GUIDELINES FOR SYNTHESIS OF PURINE-LIKE 1-β-D-2-DEOXYRIBOSYL-N-NUCLEOSIDES**

This protocol outlines a general procedure for the synthesis of unnatural hydrophobic purine-like *N*-nucleosides. The coupling strategy between 5*H*-pyrrolo[2,3-*b*]pyrazine (PP) and **S.2** (Fig. 1.5.1) to generate the nucleoside 1'-β-3',5'-*O*-toluoyl-2'-deoxyribosyl-*N*-5*H*-pyrrolo[2,3-*b*]pyrazine (**S.8**; Fig. 1.5.3) can be generally applied for the synthesis of other purine-like nucleosides such as those depicted in Figure 1.5.4 by starting with the appropriate purine analog. This protocol describes the synthesis of the 3',5'-toluoyl-protected nucleoside (**S.8**). Synthesis of the free nucleoside, phosphoramidite, and triphosphate are performed as in Basic Protocol 1.



**Figure 1.5.3** General procedure for synthesis of unnatural hydrophobic purine-like *N*-nucleosides



**Figure 1.5.4** Unnatural hydrophobic purine-like *N*-nucleosides (see Basic Protocol 2).

### Materials

Argon (see recipe)  
 5*H*-Pyrrolo[2,3-*b*]pyrazine or purine analog of choice  
 Acetonitrile, anhydrous  
 60% sodium hydride (NaH; see recipe)  
 Bis-toluoyl-protected chloroglycoside:  
   1- $\alpha$ -chloro-3,5-di-*O*-toluoyl-2-deoxyribofuranose (**S.2**; Berry & Associates;  
   Takeshita et al., 1987)  
 Ethyl acetate  
 Ethyl ether  
 Saturated sodium bicarbonate (NaHCO<sub>3</sub>)  
 Sodium sulfate (Na<sub>2</sub>SO<sub>4</sub>), anhydrous  
 Silica gel, 200-400 mesh, 60 Å  
 Hexanes  
 10-mL two-neck round-bottom flask with 14/20 joints, oven-dried  
 Rubber septa  
 18-G needle  
 Cotton  
 Glass funnel  
 1.5 × 30-cm heavy-walled glass column with glass-adapters attached to  
   compressed air or nitrogen source (see flash chromatography steps in *APPENDIX*  
   3*E*)

Sea sand  
 Silica gel 60 F<sub>254</sub> alumina-backed thin-layer chromatography (TLC) plates (Fischer)  
 Heat gun  
 UV light source  
 Rotary evaporator (Büchi) equipped with a dry ice condenser and a vacuum system  
 Additional reagents and equipment for purification and analysis (see Basic Protocol 1).

1. Equip a 10-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G syringe needle threaded through the septum.
2. Add 100 mg (0.84 mmol) 5*H*-pyrrolo[2,3-*b*]pyrazine and 4.5 mL anhydrous acetonitrile and cool the solution 5 min in an ice-water bath.
3. Add 36 mg of 60% NaH (0.9 mmol) in three aliquots to the reaction mixture. Remove the cold bath and stir at ambient temperature for 10 min.
4. Add 326.8 mg bis-toluoyl-protected chloroglycoside (**S.2**) in three aliquots to the flask and stir at ambient temperature for 10 min.

*Preparation of S.8 uses 1.0 eq 5*H*-pyrrolo[2,3-*b*]pyrazine, 1.1 eq NaH, 1.2 eq S.2, and 523 eq CH<sub>3</sub>CN.*

5. Add 25 mL of ethyl acetate and 20 mL of saturated aqueous NaHCO<sub>3</sub> to partition the reaction mixture.
6. Extract the aqueous layer twice, each time with 20 mL ethyl acetate.
7. Dry and filter the organic layer (see Basic Protocol 1, steps 11 and 12).
8. Perform flash chromatography (see Basic Protocol 1, steps 13 to 19) using a 1.5 × 10-cm column packed with 20 g silica gel. Elute with a gradient of 6:1 to 3:1 (v/v) hexanes/ethyl acetate. Analyze fractions by TLC (see Basic Protocol 1, step 8) using 4:1 hexanes/ethyl acetate.

*For S.8, R<sub>f</sub> = ~0.35.*

9. Combine the appropriate fractions and evaporate to dryness using a rotary evaporator.

*The bis-protected product (S.8) is prepared as a white foam (161 mg) at 41% yield.*

## BASIC PROTOCOL 3

### SYNTHESIS OF 3,5-DIMETHYLPHENYL-*C*-NUCLEOSIDE

This protocol describes the synthesis of the 3,5-dimethylphenyl-*C*-nucleoside, involving the condensation of the Grignard reagent derived from 1-bromo-3,5-dimethylbenzene with the bis-toluoyl-protected chloroglycoside (**S.2**; Fig. 1.5.1; see Fig. 1.5.5 for the reaction). Deprotection yields the free nucleoside (**S.11**), followed by generation of the phosphoramidite (**S.12**) and triphosphate (**S.13**) derivatives. Two other *C*-nucleosides are prepared using a similar strategy (TM and 2Nap in Fig. 1.5.6A). Deprotection and formation of the phosphoramidite and triphosphate all follow the general procedures as described in Basic Protocol 1.

#### Materials

Argon (see recipe)  
 Magnesium metal turnings  
 Tetrahydrofuran (THF), anhydrous  
 1-Bromo-3,5-dimethylbenzene (**S.9**; Aldrich)

Bis-toluoyl-protected chloroglycoside:

1- $\alpha$ -chloro-3,5-di-*O*-toluoyl-2-deoxyribofuranose (**S.2**; Takeshita et al., 1987)

Ethyl acetate

Saturated ammonium chloride (NH<sub>4</sub>Cl) solution

Sodium sulfate, anhydrous

Silica gel, 200-400 mesh, 60 Å

Hexanes

Ethyl ether, anhydrous

Methanol, anhydrous

Sodium methoxide (Aldrich or synthesized from sodium metal and methanol)

Dichloromethane (CH<sub>2</sub>Cl<sub>2</sub>), freshly distilled from calcium hydride

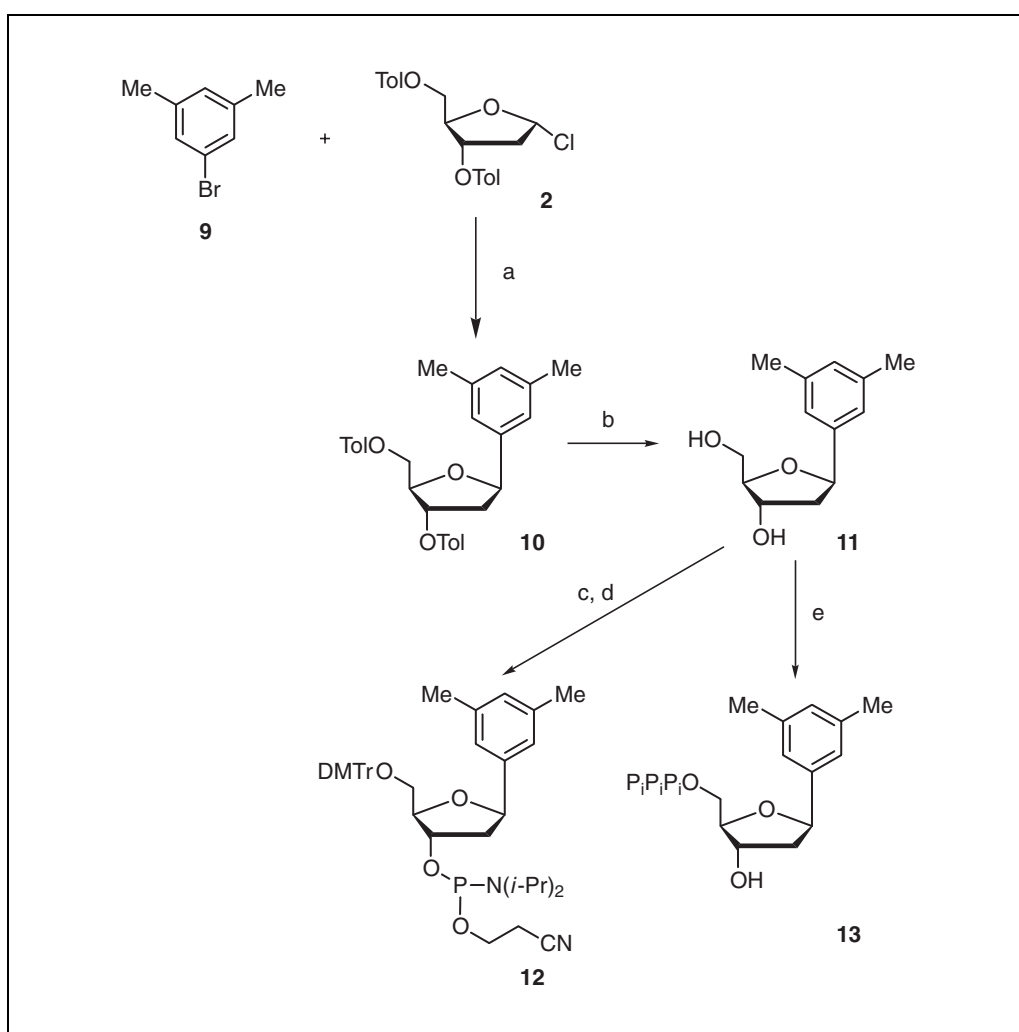
Pyridine, anhydrous

Triethylamine (TEA), freshly distilled from calcium hydride

4,4'-Dimethoxytrityl chloride (DMTr-Cl; Aldrich)

Saturated sodium bicarbonate (NaHCO<sub>3</sub>) solution

4-Dimethylaminopyridine (DMAP; Aldrich)



**Figure 1.5.5** Synthesis of 3,5-dimethylphenyl-*C*-nucleoside and its phosphoramidite and triphosphate derivatives (see Basic Protocol 3). Reagents: (A) Mg<sup>0</sup>, then **S.9**, THF (steps 1 to 9); (B) sodium methoxide, methanol (steps 10 to 15); (C) DMTrCl, TEA, pyridine, CH<sub>2</sub>Cl<sub>2</sub> (steps 16 to 24); (D) 2-cyanoethyl-diisopropylchlorophosphoramidite, triethylamine, CH<sub>2</sub>Cl<sub>2</sub> (steps 25 to 31); (E) Proton-Sponge, trimethyl phosphate, POCl<sub>3</sub>, tributylamine, tributylammonium pyrophosphate, DMF (steps 32 to 38).

2-Cyanoethyl diisopropylchlorophosphoramidite (Aldrich)  
 Trimethyl phosphate (Aldrich)  
 Proton-Sponge (Aldrich)  
 Phosphorus oxychloride (POCl<sub>3</sub>)  
 Tributylamine  
*n*-Tetrabutylammonium pyrophosphate (TBAP; Sigma)  
 Dimethylformamide (DMF)  
 Triethylammonium bicarbonate (TEAB; Fluka)  
 Dimethylsulfoxide (DMSO)  
 Isopropanol  
 25-mL and 100-mL two-neck round-bottom flasks with 14/20 joints, oven-dried  
 Reflux condenser with 14/20 joint  
 Rubber septa  
 18- and 23-G needles  
 125-mL separatory funnel  
 Cotton  
 Glass funnel  
 Rotary evaporator (Büchi) equipped with a dry ice condenser and a vacuum system  
 Heavy-walled glass columns (1.5-cm i.d.; 10- and 15-cm length) with glass  
 adapters attached to compressed air or nitrogen source (see flash  
 chromatography steps in *APPENDIX 3E*)  
 Sea sand  
 300-mL round-bottom flask  
 Filter paper  
 Lyophilizer (e.g., Labconco freeze-dry system)  
 Additional reagents and equipment for purification and analysis (see Basic  
 Protocol 1)

***Perform condensation of 1-bromo-3,5-dimethylbenzene (S.9) with bis-toluoyl-protected chloroglycoside (S.2)***

1. Assemble a 25-mL oven-dried two-neck round-bottom flask with a 14/20 reflux condenser in the center neck, a rubber septum on the side neck, and a magnetic stirring bar. Maintain a smooth flow of argon using a 23-G needle inserted through another septum on top of the reflux condenser.
2. Quickly add 52 mg (2.139 mmol) magnesium metal and 2 mL THF to the flask under argon atmosphere.
3. Add, in a dropwise fashion, 293  $\mu$ L (2.156 mmol) 1-bromo-3,5-dimethylbenzene (**S.9**) and heat the resulting suspension to 50°C for 1 hr using a temperature-controlled oil bath.
4. Prepare a solution of 126 mg (0.324 mmol) bis-toluoyl-protected chloroglycoside (**S.2**) in 1 mL THF in another round-bottom flask under argon atmosphere. Add all of the prepared Grignard reagent (step 3) to this solution and stir 14 hr at room temperature.
5. During the 14-hr incubation, prepare another batch of Grignard reagent (repeat steps 1 to 3). At the end of the 14-hr incubation, add another 100  $\mu$ L Grignard reagent to the reaction mixture (step 4) and continue to stir 1 hr at room temperature.

*Synthesis of S.10 uses 1.0 eq S.9, 0.992 eq magnesium metal, 0.150 eq S.2, and 17.16 eq THF.*

***Work up and purify S.10***

6. Add 10 mL of ethyl acetate and 10 mL of saturated aqueous NH<sub>4</sub>Cl and transfer to a 125-mL separatory funnel. Rinse the reaction vessel twice, each time with 5 mL ethyl acetate, and add to the funnel.

7. Extract separate, and retain the organic layer. Extract the aqueous layer twice, each time with 15 mL of ethyl acetate, and combine these organic layers with the original organic layer.
8. Filter and dry the organic layer (see Basic Protocol, steps 11 to 12). Perform flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with a gradient of 95:5 to 85:15 (v/v) hexanes/ethyl acetate. Perform TLC (see Basic Protocol 1, step 8) using 9:1 hexanes/ethyl acetate.

*For S.10,  $R_f = \sim 0.30$ .*

9. Evaporate eluate and analyze product by NMR (see Basic Protocol 1, steps 20 to 22).

*Approximately 13 mg (9% yield) of purified  $\beta$ -product (S.10) and 48 mg (32%) of  $\alpha$ -product should be obtained.*

#### **Deprotect 3'- and 5'-OH groups**

10. Equip a 100-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G syringe needle threaded through a septum.
11. Add 264 mg (0.576 mmol) S.10 (from step 9) and 10 mL anhydrous methanol to the flask.
12. Add, in a dropwise fashion, 2 mL of 1 M sodium methoxide (2.0 mmol) in methanol to the reaction mixture and keep stirring for 45 min.

*Synthesis of S.11 uses 1.0 eq S.10, 3.5 eq sodium methoxide, and 514 eq methanol.*

#### **Work up and purify S.11**

13. Add 100 mg  $\text{NH}_4\text{Cl}$  to quench the reaction and continue to stir for 10 min.
14. Evaporate solvents and purify the crude product by flash chromatography (see Basic Protocol 1, steps 13 to 19) using a  $1.5 \times 10$ -cm silica gel column packed with 20 g silica gel. Elute with a gradient of 1:99 to 5:95 (v/v) methanol/ $\text{CH}_2\text{Cl}_2$ . Perform TLC (see Basic Protocol, step 8) using 5:95 methanol/ $\text{CH}_2\text{Cl}_2$ .

*For S.11,  $R_f = \sim 0.30$ .*

15. Combine appropriate fractions based on TLC and evaporate to dryness in a rotary evaporator.

*Flash chromatography gives approximately 105 mg (82% yield) of purified product (S.11) as a white solid.*

#### **Tritylate 5'-OH**

16. Equip a 100-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G syringe needle threaded through one of the septa.
17. Add 90 mg (0.405 mmol) of S.11 (from step 15), 2 mL anhydrous pyridine, and 2 mL  $\text{CH}_2\text{Cl}_2$  to the flask.
18. Add, in a dropwise fashion, 0.30 mL (2.15 mmol) TEA, followed by 175 mg (0.516 mmol) 4,4'-DMTr-Cl in two portions over a 30-min period, and continue to stir for 2 hr at room temperature.

*Tritylation uses 1.0 eq S.11, 1.27 eq DMTr-Cl, 5.31 eq triethylamine, 61 eq pyridine, and 77 eq dichloromethane.*

#### **Work up and purify tritylated product**

19. Add 20 mL ethyl acetate and 10 mL saturated  $\text{NaHCO}_3$  and transfer the mixture to a 125-mL separatory funnel. Rinse the reaction vessel twice, each time with 5 mL ethyl acetate, and add to the funnel.

20. Extract separate, and retain the organic layer. Extract the aqueous layer twice, each time with 20 mL ethyl acetate.
21. Combine the organic layers and dry over anhydrous  $\text{Na}_2\text{SO}_4$ .
22. Filter through filter paper into a 300-mL round-bottom flask and evaporate under reduced pressure in a rotary evaporator.
23. Purify the crude product by flash chromatography (see Basic Protocol 1, steps 13 to 19) using a  $1.5 \times 10$ -cm silica gel column packed with 20 g silica gel. Elute with a gradient of 1:1 to 4:1 (v/v) hexanes/ethyl acetate. Perform TLC (see Basic Protocol, step 8) using 1:1 hexanes/ethyl acetate.

*For the tritylated S.11,  $R_f = \sim 0.3$ .*

24. Combine the appropriate fractions based on TLC analysis and evaporate to dryness in a rotary evaporator.

*Flash chromatography affords 185 mg (87% yield) of purified DMTr-protected product as a white foam.*

#### **Prepare phosphoramidite S.12 by phosphorylation**

25. Equip a 100-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G syringe needle threaded through one of the septa.
26. Add 185 mg (0.353 mmol) tritylated product (from step 24) and 3.5 mL anhydrous  $\text{CH}_2\text{Cl}_2$  to the flask.
27. Add a catalytic amount (2 mg, 0.016 mmol) of DMAP, followed by 0.350 mL (2.512 mmol) triethylamine and 0.160 mL (0.718 mmol) 2-cyanoethyl diisopropylamino-chlorophosphoramidite.
28. Continue to stir for 30 min at room temperature.

*Synthesis of S.12 uses 1.0 eq tritylated nucleoside, 0.045 eq DMAP, 7.1 eq TEA, 2.0 eq 2-cyanoethyl diisopropylaminochlorophosphoramidite, and 155 eq dichloromethane.*

#### **Work up and purify S.12**

29. Add 20 mL ethyl acetate and 20 mL saturated  $\text{NaHCO}_3$  and repeat steps 19 to 22 of this protocol.
30. Purify the crude product by flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with a gradient of 1:9 to 3:7 (v/v) ethyl acetate/hexane containing 5% triethylamine. Perform TLC using 3:7 ethyl acetate/hexane.

*For S.12,  $R_f = \sim 0.83$ .*

31. Combine appropriate fractions based on TLC and evaporate to dryness.

*Flash chromatography affords 238 mg (93% yield) of phosphoramidite product (S.12) as a white foam.*

#### **Prepare triphosphate S.13 by phosphorylation of S.11**

32. Equip a 25-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G needle threaded through the septum.
33. Add 16 mg (0.072 mmol) free nucleoside (S.11, from step 15) to the flask.
34. Add 0.36 mL (3.08 mmol) trimethyl phosphate and 23 mg (0.107 mmol) Proton-Sponge. Cool to  $0^\circ\text{C}$ .
35. Add 8  $\mu\text{L}$  (0.090 mmol)  $\text{POCl}_3$  in a dropwise fashion and continue to stir for 2 hr at  $0^\circ\text{C}$ .



36. Add 105  $\mu\text{L}$  (0.441 mmol) tributylamine, followed by a solution of 62 mg (0.171 mmol) TBAP in 0.8 mL DMF, and stir for 1 min.

*Synthesis of S.13 uses 1.0 eq S.11, 43 eq trimethyl phosphate, 1.5 eq Proton-Sponge, 1.3 eq  $\text{POCl}_3$ , 6.1 eq tributylamine, 2.4 eq TBAP, and 143 eq DMF.*

#### **Work up and purify S.13**

37. Add 7 mL of 1 M TEAB. Dilute the resulting crude solution ~10 fold with water and lyophilize.
38. Dissolve in 200 mL of 1:1 (v/v) DMSO/isopropanol and purify by HPLC (UNIT 10.5) using the following conditions:

Column: Rainin C18-Dynamax 60 Å column  
Buffer A: 0.1 M TEAB, pH 7.5  
Buffer B: Acetonitrile  
Gradient: 96% to 70% buffer A from 0 to 30 min  
Flow rate: 10 mL/min.

*Approximately 1 mg of triphosphate (S.13) is obtained as a white solid after lyophilization.*

### **SYNTHESIS OF 1,4-DIMETHYLNAPHTHALENE-C-NUCLEOSIDE**

This protocol describes the synthesis of 1,4-dimethylnaphthalene-C-nucleoside (DMN, see Fig. 1.5.6B), as derived from the condensation of the aryllithium species onto an aldehyde followed by a ring-closing reaction, shown in Figure 1.5.7 (Ogawa et al., 2000a). The steps below detail the synthesis of the free nucleoside (S.18). The preparation of the phosphoramidite and triphosphate are performed as described in Basic Protocol 3. A similar strategy may be used to synthesize the 1-methyl-3-naphthalene nucleoside (3MN; Fig. 1.5.6B), starting from 3-bromo-1-methylnaphthalene (Ogawa, et al., 2000b).

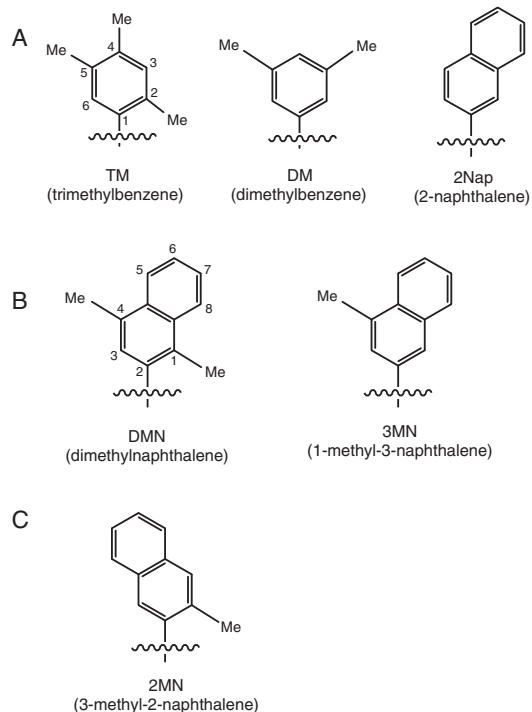
#### **Materials**

Argon (see recipe)  
2-Bromo-1,4-dimethylnaphthalene (S.14; Aldrich; Sharma, 1993)  
Tetrahydrofuran (THF), anhydrous  
*n*-Butyllithium (Aldrich)  
Cyclohexane  
Protected aldehyde of sugar precursor (S.15; Eaton and Millican, 1988)  
Ethyl acetate  
Saturated sodium bicarbonate ( $\text{NaHCO}_3$ ) solution  
Sodium sulfate ( $\text{Na}_2\text{SO}_4$ ), anhydrous  
Silica gel, 200-400 mesh, 60 Å  
Hexane  
Ethyl acetate  
Pyridine, anhydrous  
Triethylamine (TEA), freshly distilled from calcium  
Methanesulfonyl chloride  
Acetic acid, glacial  
Methanol, anhydrous  
Trifluoroacetic acid  
Dichloromethane ( $\text{CH}_2\text{Cl}_2$ ), freshly distilled from calcium hydride  
2,3-Dichloro-5,6-dicyano-1,4-benzoquinone (DDQ; Aldrich)  
25-mL, 50-mL, and 100-mL two-neck round bottom flasks with 14/20 joints, oven-dried  
Rubber septa  
18-G needles

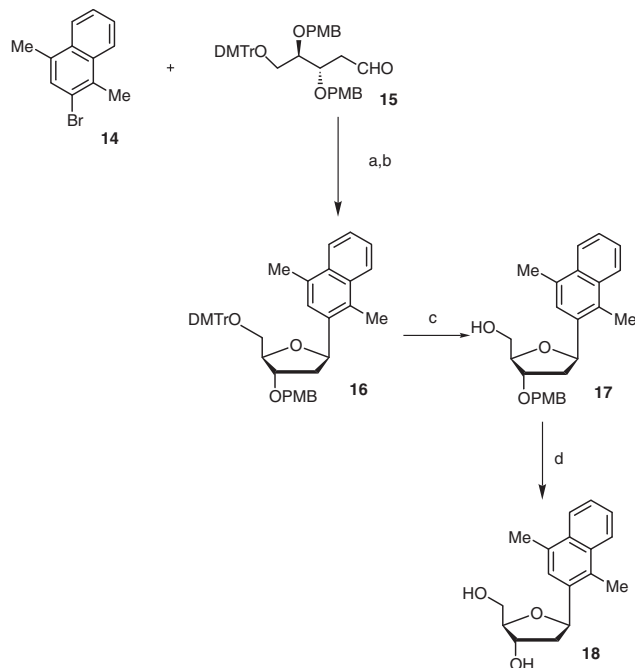
#### **BASIC PROTOCOL 4**

#### **Synthesis of Modified Nucleosides**

### **1.5.17**



**Figure 1.5.6** Unnatural hydrophobic C-nucleosides. Synthetic strategies are presented in **(A)** Basic Protocol 3, **(B)** Basic Protocol 4, and **(C)** Alternate Protocol.



**Figure 1.5.7** Synthesis of 1,4-dimethylnaphthalene-C-nucleoside (see Basic Protocol 4). Reagents: **(A)** *n*-butyllithium,  $-78^{\circ}\text{C}$ , then **S.15**, THF (steps 1 to 8); **(B)** methanesulfonyl chloride, triethylamine, pyridine,  $0^{\circ}\text{C}$  (steps 9 to 14); **(C)** acetic acid (steps 15 to 20); **(D)** DDQ (steps 21 to 26).

Dry ice/acetone bath ( $-78^{\circ}\text{C}$ )

Cotton

Glass funnel

Rotary evaporator (Büchi) equipped with a dry ice condenser and a vacuum system

Heavy-walled glass columns (1.5-cm i.d.; 10-, 15-, and 20-cm length) with glass adapters attached to compressed air or nitrogen source (see flash chromatography steps in *APPENDIX 3E*)

Sea sand

Additional reagents and equipment for purification and analysis (see Basic Protocol 1)

**NOTE:** *n*-Butyllithium is flammable and sensitive to moisture. It should be handled carefully and quickly.

***Perform condensation of 2-bromo-1,4-dimethylnaphthalene (S.14) with protected aldehyde of sugar precursor (S.15)***

1. Equip a 100-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G needle threaded through one of the septa.
2. Add 640 mg (2.72 mmol) 2-bromo-1,4-dimethylnaphthalene (**S.14**) and 15 mL THF to the flask under argon atmosphere and cool to  $-78^{\circ}\text{C}$  in a dry ice/acetone bath.
3. Slowly add dropwise 2.0 mL of 2 M *n*-butyllithium (4.0 mmol) in cyclohexane and continue to stir for 15 min at  $-78^{\circ}\text{C}$ .
4. Add, in a dropwise fashion, a solution of 1.214 g (1.801 mmol) of the protected aldehyde of the sugar precursor (**S.15**) in 5 mL THF and continue to stir for 1 hr at  $-78^{\circ}\text{C}$ .
5. Remove the flask from the bath and warm to room temperature over 1.5 hr.

*Synthesis of the 1,4-dimethylnaphthalene derivative uses 1.0 eq S.14, 1.47 eq n-butyllithium, 0.66 eq S.15, and 90 eq THF.*

***Work up and purify product***

6. Add 10 mL ethyl acetate and 10 mL saturated  $\text{NaHCO}_3$  to partition. Filter and dry the organic layer (see Basic Protocol 1, steps 11 and 12).
7. Perform flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with a gradient of 9:1 to 8:2 (v/v) hexane/ethyl acetate. Perform TLC (see Basic Protocol 1, step 8) using 7:3 hexane/ethyl acetate.

*For this product,  $R_f \sim 0.23$ .*

8. Combine the appropriate fractions based on TLC and evaporate to dryness with a rotary evaporator.

*Flash chromatography gives the C1 S-isomer (712 mg, 48% yield) and the C1 R-isomer (625 mg, 42% yield). S and R isomers were assigned based on the  $\alpha$  or  $\beta$  anomer that was formed after cyclization.*

***Perform in situ cyclization of mesylate***

9. Equip a 25-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G syringe needle threaded through one of the septa.
10. Add 712 mg (0.857 mmol) of the *S*-diastereomer (step 8) and 9 mL anhydrous pyridine to the flask and cool to  $0^{\circ}\text{C}$ .

11. Add 0.57 mL (4.09 mmol) triethylamine followed by 90  $\mu$ L (1.17 mmol) methanesulfonyl chloride, and warm slowly to room temperature over 1.5 hr with stirring.

*Synthesis of S.16 uses 1.0 eq S-diastereomer, 4.8 eq triethylamine, 1.37 eq methanesulfonyl chloride, and 130 eq pyridine.*

#### **Work up and purify S.16**

12. Add 2 mL saturated  $\text{NaHCO}_3$  to quench the reaction.
13. Perform flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with a gradient of 95:5 to 85:15 (v/v) hexane/ethyl acetate. Perform TLC using 7:3 hexane/ethyl acetate.

*For S.16,  $R_f = \sim 0.68$ .*

14. Collect appropriate fractions based on TLC and evaporate to dryness with a rotary evaporator.

*Flash chromatography yields 176 mg (30% yield) of purified product (S.16).*

#### **Detritylate 5'-OH group**

15. Equip a 25-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G syringe needle threaded through one of the septa.
16. Add 190 mg (0.274 mmol) S.16, 6 mL acetic acid, and 1 mL methanol to the flask.
17. Add 1 mL trifluoroacetic acid in a dropwise fashion ( $\sim 10$  drops) and continue stirring for 20 min at room temperature.

*Detritylation uses 1.0 eq S.16, 383 eq acetic acid, 90 eq methanol, and 47.4 eq trifluoroacetic acid.*

#### **Work up and purify S.17**

18. Concentrate the reaction solution with a rotary evaporator.
19. Remove solvents (see Basic Protocol, step 12) and perform flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with a gradient of 4:1 to 1:1 (v/v) hexane/ethyl acetate. Perform TLC using ethyl acetate.

*For S.17,  $R_f = \sim 0.75$ .*

20. Combine appropriate fractions based on TLC and evaporate to dryness with a rotary evaporator.

*Flash chromatography affords 93 mg (86% yield) of purified product (S.17).*

#### **Remove 3'-PMB group**

21. Equip a 50-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G syringe needle threaded through one of the septa.
22. Add 8 mg, 0.020 mmol S.17, 1 mL  $\text{CH}_2\text{Cl}_2$ , and 1 drop  $\text{H}_2\text{O}$  to the flask.
23. Add 7 mg (0.031 mmol) DDQ and continue to stir for 1 hr at room temperature.

*Deprotection uses 1.0 eq S.17, 780 eq  $\text{CH}_2\text{Cl}_2$ , 277.5 eq  $\text{H}_2\text{O}$ , and 1.55 eq DDQ.*

#### **Work up and purify S.18**

24. Add 15 mL ethyl acetate and 15 mL saturated  $\text{NaHCO}_3$  to partition. Filter and dry the organic layer (see Basic Protocol 1, steps 11 and 12).

25. Perform flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with a gradient of 1:99 to 5:95 (v/v) methanol/CH<sub>2</sub>Cl<sub>2</sub>. Perform TLC using 3:97 methanol/CH<sub>2</sub>Cl<sub>2</sub>.

*For S.18,  $R_f = \sim 0.3$ .*

26. Combine appropriate fractions based on TLC and evaporate to dryness using rotary evaporator.

*Flash chromatography affords 3 mg (55% yield) of pure product (S.18).*

### SYNTHESIS OF 3-METHYL-2-NAPHTHALENE-C-NUCLEOSIDE

This protocol describes the synthesis of the 3-methyl-2-naphthalene-C-nucleoside (Fig. 1.5.6C), which was derived from the condensation of the aryllithium species onto another aldehyde followed by in situ cyclization as shown in Figure 1.5.8 (Ogawa et al., 2000a). The preparation of phosphoramidite and triphosphate follows the general procedures described in Basic Protocol 3.

**Additional Materials** (also see Basic Protocols 3 and 4)

2-Bromo-3-methylnaphthalene (**S.19**; prepared according to Lambert et al., 1979)

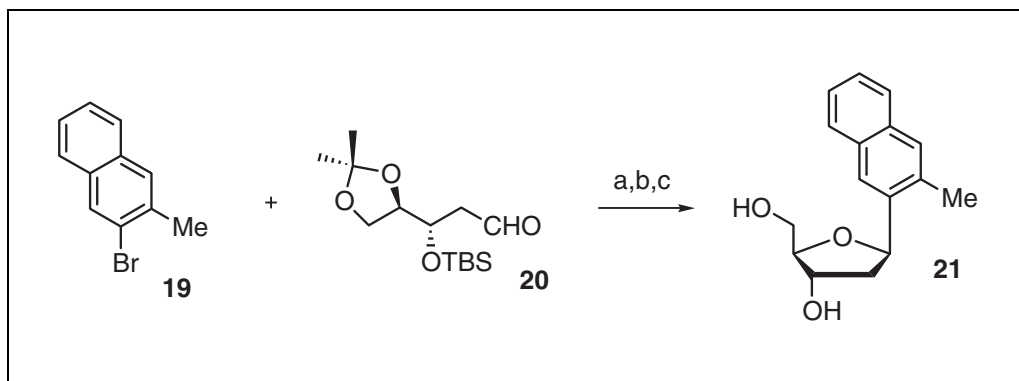
Protected aldehyde of sugar precursor (**S.20**, prepared according to Solomon and Hopkins, 1993)

**Perform condensation of 2-bromo-3-methylnaphthalene (S.19) with the protected aldehyde of the sugar precursor (S.20)**

1. Place a magnetic stir bar in a dry 100-mL two-neck flask with two rubber septa, and place the flask on top of a magnetic stir plate.
2. Evacuate the reaction flask on vacuum line and then flush it with argon. Repeat this procedure three times and attach the flask to an argon line on the manifold.
3. Add 595 mg (2.70 mmol) 2-bromo-3-methylnaphthalene (**S.19**) and 8.9 mL THF to the flask under argon atmosphere and cool to  $-78^{\circ}\text{C}$  in a dry ice/acetone bath.
4. Add dropwise 1.50 mL of 2 M *n*-butyllithium (3.0 mmol) in cyclohexane over 10 min and continue to stir 15 min at  $-78^{\circ}\text{C}$ .
5. Add dropwise a solution of 502 mg (1.74 mmol) aldehyde (**S.20**) in 4.4 mL THF over 5 min and continue to stir 30 min at  $-78^{\circ}\text{C}$ .
6. Remove the flask from the bath and warm to room temperature.

*During this time the reaction color changes from deep green to brown.*

*Condensation uses 1.0 eq S.19, 1.1 eq *n*-butyllithium, 0.64 eq S.20, and 61 eq THF.*



**Figure 1.5.8** Synthesis of 3-methyl-2-naphthalene-C-nucleoside (see Alternate Protocol). Reagents: (A) *n*-butyllithium,  $-78^{\circ}\text{C}$ , then **S.2**, THF (steps 1 to 9); (B) methanesulfonyl chloride, triethylamine, pyridine,  $0^{\circ}\text{C}$  (steps 10 to 16); (C) TFA (steps 17 to 22).

**ALTERNATE  
PROTOCOL**

**Synthesis of  
Modified  
Nucleosides**

**1.5.21**

**Work up and purify product**

7. Add 10 mL ethyl acetate and 10 mL saturated  $\text{NaHCO}_3$  to partition. Filter and dry the organic layer (see Basic Protocol 1, steps 11 and 12).
8. Perform flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with a gradient of 9:1 to 5:1 (v/v) hexane/ethyl acetate. Perform TLC (see Basic Protocol 1, step 8) using 9:1 hexane/ethyl acetate.

*For this product,  $R_f = \sim 0.30$ .*

9. Combine appropriate fractions based on TLC and evaporate to dryness with a rotary evaporator.

*Flash chromatography gives a crude mixture of both diastereomers (532 mg, 1.24 mmol).*

**Mesylate hydroxy group**

10. Equip a 100-mL two-neck flask with two rubber septa and a stir bar. Keep the vessel under argon using an 18-G syringe needle threaded through one of the septa.
11. Add 510 mg (1.19 mmol) product (step 9, both diastereomers) and 32 mL  $\text{CH}_2\text{Cl}_2$  to the flask and cool to  $0^\circ\text{C}$ .
12. Add 0.332 mL (2.38 mmol) triethylamine, followed by 0.120 mL (1.55 mmol) methanesulfonyl chloride, and warm slowly to room temperature over 30 min with stirring.

*Relative to S.19 (step 3), mesylation uses 0.88 eq triethylamine, 0.57 eq methanesulfonyl chloride, and 185 eq  $\text{CH}_2\text{Cl}_2$ .*

**Work up and purify product**

13. Add 30 mL saturated aqueous  $\text{NaHCO}_3$  and transfer to a 125-mL separatory funnel. Rinse the reaction vessel twice, each time with 5 mL ethyl acetate and add to the funnel.
14. Separate the organic layer and extract the aqueous layer twice, each time with 15 mL ethyl acetate.
15. Filter and dry the organic layer (see Basic Protocol 1, steps 11 and 12), and perform flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with a gradient of 95:5 to 85:15 (v/v) hexanes/ethyl acetate. Perform TLC using 9:1 hexanes/ethyl acetate.

*For this product,  $R_f = \sim 0.30$ .*

16. Evaporate eluate and analyze product by NMR (see Basic Protocol 1, steps 20 to 22).

*The resulting crude product is pure enough for use in the next step.*

**Perform in situ cyclization of mesylate**

17. Set up the reaction flask as in steps 1 and 2 above.
18. Add 605 mg (1.19 mmol) crude product (step 16) and 56 mL of 4:1 (v/v) TFA/methanol and stir 20 min at room temperature.

*Relative to S.19 (step 3), cyclization uses 216 eq trifluoroacetic acid and 103 eq methanol.*

**Work up and purify S.21**

19. Concentrate the solution with a rotary evaporator and then neutralize the residual TFA with a minimum volume of saturated  $\text{NaHCO}_3$ .
20. Transfer to a separatory funnel and extract four times, each time with 20 mL  $\text{CH}_2\text{Cl}_2$ .

21. Filter and dry the organic layer (see Basic Protocol 1, steps 11 and 12), and perform flash chromatography (see Basic Protocol 1, steps 13 to 19). Elute with 3:97 (v/v) methanol/CH<sub>2</sub>Cl<sub>2</sub>. Perform TLC using the same solvent

For **S.21**,  $R_f = \sim 0.2$ .

22. Collect the appropriate fractions and evaporate to dryness with a rotary evaporator.

Flash chromatography affords the desired  $\beta$ -anomer (**S.21**) at 108 mg (0.418 mmol, 24% yield over 3 steps).

## PURIFICATION OF DNA OLIGONUCLEOTIDES

Oligonucleotides containing an unnatural base are synthesized using an ABI 392 DNA/RNA synthesizer with automatic removal of the 5'-trityl protecting group (so-called trityl-off procedure; *APPENDIX 3C*). The oligonucleotides are then deprotected and cleaved from the CPG support according to the instrument instruction manual. Figure 1.5.9 shows the sequences of four oligonucleotides that are used in the thermodynamic and kinetic analyses described in Basic Protocols 6 and 7. dC-CPG (500 Å, 1.0  $\mu$ mol) is used for **S.22**, dG-CPG (500 Å, 1.0  $\mu$ mol) is used for **S.23**, dA-CPG (500 Å, 1.0  $\mu$ mol) is used for **S.24**, the universal support (500 Å, 1.0  $\mu$ mol) is used for **S.25** and dA-CPG (500 Å, 1.0  $\mu$ mol) is used for **S.26**. To be suitable for use in Basic Protocols 6 and 7, the oligonucleotides are purified by preparative denaturing PAGE (*UNIT 10.4* and *APPENDIX 3B*), excised from the gel, electroeluted using a Schleicher & Schuell Elutrap electrophoresis chamber, and ethanol precipitated to concentrate and desalt the sample. The oligonucleotides are stored at  $-20^\circ\text{C}$ .

### Materials

CPG-bound DNA oligonucleotide (*APPENDIX 3C*)  
Concentrated ammonium hydroxide  
95% (v/v) formamide in 10 mM Tris-Cl, pH 8.5 (*APPENDIX 2A*), with and without 0.05% (w/v) xylene cyanol and bromphenol blue  
TBE buffer (see *APPENDIX 2A*)  
Absolute ethanol, prechilled to  $-20^\circ\text{C}$   
5 M NaCl, ice cold  
80% ethanol, ice cold  
1.5-mL screw-cap vial  
60° and 80°C Dri-baths (Thermolyne)  
Speed-Vac evaporator (Savant)  
50-mL polypropylene centrifuge tubes  
UV/Vis spectrophotometer  
Electroelution unit (Schleicher and Schuell)  
Silica gel-coated preparative TLC plate  
UV lamp (hand held)  
Refrigerated centrifuge and rotor appropriate for sample size  
Additional reagents and equipment for denaturing PAGE (*UNIT 10.4* and *APPENDIX 3B*)

### Deprotect oligonucleotide and cleave from support

1. Transfer the CPG-bound DNA oligonucleotide to a 1.5-mL screw-cap vial.
2. Add 1.5 mL concentrated ammonium hydroxide, tighten the cap, and incubate 12 hr at 80°C (for 24-mer) or 60°C (for 13-mer or 45-mer).
3. Cool to 0°C and evaporate to dryness in a Speed-Vac evaporator.

## BASIC PROTOCOL 5

### Synthesis of Modified Nucleosides

#### 1.5.23

```

22  5'-dCGCATGXGTACGC
23  5'-dGCGTACXCATGCG
24  5'-dTAAACGACTCACTATAGGGAGA
25  5'-dTAAACGACTCACTATAGGGAGAX
26  5'-dCGCTAGGACGGCATTGGATCGXTCTCCCTATAGTGAGTCGTATTA

```

**Figure 1.5.9** Oligonucleotides used for characterization of unnatural nucleobases.

***Purify crude oligonucleotide***

4. Add 300  $\mu$ L water and 300  $\mu$ L of 95% formamide/10 mM Tris·Cl, pH 8.5 (without xylene cyanol and bromphenol blue) and mix thoroughly.
5. Prepare a 12-well denaturing PAGE gel (*UNIT 10.4* and *APPENDIX 3B*) as follows, according to the length of the oligonucleotide:
  - 13-mer: 15% acrylamide; 42 cm long  $\times$  33 cm wide  $\times$  1.5 mm thick
  - 24-mer: 20% acrylamide; 42 cm long  $\times$  33 cm wide  $\times$  1.5 mm thick
  - 45-mer: 10% acrylamide; 42 cm long  $\times$  33 cm wide  $\times$  1.5 mm thick.
6. Load 60  $\mu$ L of 95% formamide/10 mM Tris·Cl, pH 8.5, with 0.05% (w/v) xylene cyanol and bromphenol blue in each of the two outermost wells; load 60  $\mu$ L of the sample in each of the ten innermost wells.
7. Purify the DNA to single-nucleotide resolution by denaturing PAGE (*UNIT 10.4* and *APPENDIX 3B*) at 65 W until the full-length oligonucleotide has migrated two-thirds of the way down the gel, as indicated by the dye markers.
8. Excise the full-length DNA product by UV shadowing (*UNIT 10.4*).

***Isolate, desalt, and concentrate pure oligonucleotide***

9. Place the gel slice in a 50-mL polypropylene centrifuge tube and crush thoroughly with a glass stirring rod.
10. Add 1 $\times$  TBE electrophoresis buffer mix, and pour the gel slurry into the sample chamber of a properly assembled and oriented electroelution apparatus, according to the instrument instruction manual.
11. Electroelute the sample and monitor its progress by withdrawing small samples and measuring absorbance at 260 nm ( $A_{260}$ ).
12. Combine the most concentrated fractions.
13. Add 3 vol absolute ethanol (prechilled to  $-20^{\circ}\text{C}$ ) and  $\frac{1}{20}$  vol 5 M NaCl and freeze at  $-20^{\circ}\text{C}$  for at least 30 min.
14. Centrifuge for at least 20 min at  $>15,000 \times g$ ,  $4^{\circ}\text{C}$ .
15. Carefully decant the supernatant, wash the DNA pellet twice with ice cold 80% ethanol, and allow it to air dry.
16. Dissolve DNA in an appropriate volume of sterile deionized water for determination of concentration and downstream applications.



## DETERMINATION OF THERMODYNAMIC STABILITY OF UNNATURAL BASE PAIRS

*Thermal stability* or *thermodynamic stability* refers to the melting temperature ( $T_m$ ) of a DNA duplex that contains the unnatural base pair. *Thermodynamic selectivity* refers to the difference in  $T_m$  between the desired unnatural base pair and the most stable mispair of an unnatural base with a natural base. Oligonucleotides **S.22** and **S.23** (Fig. 1.5.9) provide the sequence context for determination of the stability and selectivity of the unnatural nucleosides. Duplex oligonucleotide samples are prepared with the unnatural nucleoside in position X of, for example, **S.22** and paired against **S.23** containing each natural base as well as the same unnatural base (self-pair) and/or any other unnatural base to be tested. The complete set of pairs is also prepared with the unnatural nucleoside in **S.23** to determine if there are sequence context effects (e.g., CXT versus TXC). The samples are heated and evaluated for temperature-dependent changes in absorption using a Cary 300 Bio UV-Vis spectrophotometer. Duplex melting curves are collected in triplicate, and melting temperatures are determined by the derivative method contained in the Cary WinUV-Vis software. For a more detailed description of  $T_m$  experiments, see *UNIT 7.3*.

### Materials

Complementary 13-mer oligonucleotides (e.g., **S.22** and **S.23** in water; see Basic Protocol 5)

2×  $T_m$  buffer (see recipe)

Variable-temperature double-beam spectrophotometer (Cary 300 BIO UV-Vis spectrophotometer) with optically matched cuvettes

1. Prepare sample cells with 3  $\mu$ M of each complementary oligonucleotide in 1×  $T_m$  buffer. Prepare reference cells with 1×  $T_m$  buffer alone.
2. Arrange cells in the spectrophotometer and monitor the sample absorbance at 260 nm while heating at a rate of 0.5°C/min from 20° to 80°C.
3. Collect three curves for each sample and analyze data by the first-derivative method provided with the instrument software.

## KINETIC ANALYSIS OF UNNATURAL BASE PAIR INCORPORATION, SELECTIVITY AND REPLICABILITY

A rapid gel kinetic assay is used for the determination of steady-state parameters for the incorporation and extension of the unnatural base pairs, as well as to evaluate the selectivity, or fidelity of these steps. Initial velocities are determined by radiolabeled primer extension reactions at various concentrations of 5'-nucleoside triphosphate (natural or unnatural). To ensure *single completed hit* conditions, a final ratio of 30:1 primer-template to enzyme is used, and less than 20% of the primer is extended. An example of a primer-template set for determining incorporation kinetics is shown in Figure 1.5.9 with **S.24** as the primer and **S.26** as the template. The reactions can be rapidly initiated and quenched by hand and analyzed by denaturing PAGE; a PhosphorImager (Molecular Dynamics) is used to quantify gel band intensities corresponding to the extended primer. Velocities are plotted against triphosphate concentration and fit to a rectangular hyperbola. Preliminary determination of  $K_M$  is followed by kinetic analysis using triphosphate concentrations evenly distributed about the  $K_M$ . Unnatural nucleosides are characterized both in the template strand and as the incoming triphosphate. Kinetic data are collected in triplicate. In general, these kinetic determinations follow the experimental protocols developed by Goodman and co-workers (Creighton et al., 1995).

## BASIC PROTOCOL 6

## BASIC PROTOCOL 7

### Synthesis of Modified Nucleosides

#### 1.5.25

## Materials

4  $\mu\text{M}$  DNA oligonucleotide primer (e.g., **S.24**; see Basic Protocol 5)  
 10 U/ $\mu\text{L}$  T4 polynucleotide kinase and 10 $\times$  buffer (New England Biolabs)  
 10 mCi/mL [ $\gamma$ - $^{33}\text{P}$ ]ATP (2500 Ci/mmol)  
 QIAquick Nucleotide Removal Kit (Qiagen)  
 1  $\mu\text{M}$  DNA oligonucleotide template (e.g., **S.26**; see Basic Protocol 5)  
 10 U/ $\mu\text{L}$  exonuclease-free Klenow fragment (Amersham Pharmacia Biotech) and  
 10 $\times$  random-prime buffer (see recipe)  
 Unnatural triphosphates (see Basic Protocols 1 to 4 and Alternate Protocol)  
 100 mM stock of each natural dNTP (Amersham Pharmacia Biotech; sequencing  
 grade)  
 Enzyme dilution buffer (see recipe)  
 Quench solution (see recipe)  
 1 $\times$  TBE buffer (*APPENDIX 2A*)  
 37°, 25°, and 100°C Dri-Baths (Thermolyne) with heating blocks that  
 accommodate 1.5- or 0.5-mL tubes  
 Thin putty knife  
 Geiger counter  
 Chromatography paper (Whatman 3MM CHR, 35  $\times$  45 cm)  
 Gel dryer (Bio-Rad Model 583)  
 Storage Phosphor Screen (Kodak or Molecular Dynamics, 35  $\times$  43 cm)  
 PhosphorImager (e.g., STORM Model 860; Molecular Dynamics)  
 ImageQuant and compatible spreadsheet (MS Excel) and graphing (Kaleidagraph)  
 software  
 Additional reagents and equipment for denaturing PAGE (*UNIT 10.4* and *APPENDIX 3B*)

**NOTE:** Use autoclaved deionized water for all reagents.

### Radiolabel primer

1. Combine 25  $\mu\text{L}$  of 4  $\mu\text{M}$  oligonucleotide primer with 3.1  $\mu\text{L}$  of 10 $\times$  T4 polynucleotide kinase buffer, 2  $\mu\text{L}$  of [ $\gamma$ - $^{33}\text{P}$ ]ATP stock, and 1  $\mu\text{L}$  of T4 polynucleotide kinase. Mix and incubate at 37°C for at least 30 min.
2. Remove excess nucleotide with the QIAquick Nucleotide Removal Kit, according to the manual, and observing proper radioactive waste disposal.
3. Dilute buffer EB (from QIAquick kit) five-fold and use 100  $\mu\text{L}$  to elute the DNA. Store at 4°C.

*Allow the wetted column to sit for a few minutes before the final spin. Assuming 80% recovery of the labeled primer, this protocol yields 100  $\mu\text{L}$  of 800 nM 5'-radiolabeled primer.*

### Prepare analytical gel

5. Pour a 15% denaturing polyacrylamide gel (*UNIT 10.4* and *APPENDIX 3B*) that is 42 cm long, 33 cm wide, and 0.5 mm thick, and has 34 wells. Make sure there are no trapped air bubbles. Allow the gel at least 1 hr to fully polymerize before use.

### Anneal primer and template

6. Combine 25  $\mu\text{L}$  of water, 10  $\mu\text{L}$  of 10 $\times$  random-prime buffer, 8  $\mu\text{L}$  of 1  $\mu\text{M}$  template oligonucleotide, and 5  $\mu\text{L}$  of 5'-radiolabeled 800 nM primer oligonucleotide (from step 1) in a 1.5-mL microcentrifuge tube, mix well, then microcentrifuge briefly at maximum speed to bring all the solution to the bottom of the tube.

*This gives a two-fold excess of template over primer, necessary to ensure that greater than 95% of the primer is annealed to the template. The use of different primer or template sequences may require that this ratio be optimized.*

7. Place the solution in a Dri-bath preheated to 100°C and immediately turn off the heating element. Allow the primer-template to anneal by slow cooling to room temperature in the Dri-bath (~2 hr).

*This yields a final volume of 48  $\mu\text{L}$ , which, when combined with 2  $\mu\text{L}$  of enzyme, will yield sufficient volume for nine reactions of different triphosphate concentrations. Although it is best to use freshly annealed primer-template, annealed primer-templates can be stored at 4°C overnight. Warm to room temperature before use.*

### **Perform reactions**

8. Prepare 5- $\mu\text{L}$  aliquots of 2 $\times$  concentrations of the desired triphosphate range by combining the appropriate volume of 1  $\mu\text{M}$ , 10  $\mu\text{M}$ , 100  $\mu\text{M}$ , 1 mM, or 10 mM of the previously diluted dNTP and water to a final volume of 5  $\mu\text{L}$ .

*For the incorporation experiments, the unnatural triphosphate is used (in conjunction with the primer-template complex S.24-S.26) to determine the rate of unnatural base pair synthesis. To measure incorporation selectivity, each of the four natural triphosphates is used in conjunction with this primer-template. For the extension experiments, including both extension efficiency as well as extension fidelity, each of the four natural triphosphates is used in conjunction with primer-template S.25-S.26.*

*It is a good idea to make a reasonably large number of small-volume aliquots of the various concentrations of dNTP to reduce the number of freeze-thaw cycles to which the dNTP is subjected. It is generally desirable to run a control for each experiment with only 5  $\mu\text{L}$  of water and no triphosphate.*

9. Add 2  $\mu\text{L}$  of the appropriate concentration of Klenow fragment freshly diluted in enzyme dilution buffer to cooled annealed primer-template. Mix thoroughly, but do not vortex.

*This yields 50  $\mu\text{L}$  of a 2 $\times$  reaction solution that is 80 nM in primer-template, 2 $\times$  in all buffer components, and 0.04 $\times$  of the concentration of diluted enzyme used. This 2 $\times$  reaction solution will be diluted to 1 $\times$  when combined with the triphosphate aliquots to initiate the reaction. To ensure single completed hit conditions, make certain that the final enzyme concentration is no more than  $\frac{1}{30}$  that of the final primer-template concentration, final concentration referring to that achieved in step 10.*

10. Add 5  $\mu\text{L}$  enzyme-primer-template solution (from step 9) to each 5- $\mu\text{L}$  aliquot of 2 $\times$  triphosphate (from step 8), mix by pipetting up and down, close the tube, and place in a 25°C Dri-bath. Initiate all reactions in turn as quickly as possible ( $\leq 10$  sec between starts).
11. Quench the reactions in turn with 20  $\mu\text{L}$  quench solution precisely when the timer reaches the desired length for the experiment.

### **Perform electrophoresis**

12. Warm gel (prepared in step 5) by running it at 65 W in 1 $\times$  TBE buffer until it reaches a temperature of at least 37°C. Wash the wells by pipetting 1 $\times$  TBE into each a couple of times.
13. Load 8  $\mu\text{L}$  quench solution into the three outermost lanes as well as any unused lanes. Load 8  $\mu\text{L}$  of each quenched reaction into up to 28 of the innermost lanes. Run at 65 W.

*Up to 28 reactions can be loaded in one row of the gel. Four rows (112 reactions) can be run in one gel by allowing the first row to run at 65W for 30 min, stopping the gel and loading another row. The second row is run for 25 min, the third row for 20 min, and the fourth row until the dark blue dye band of the second row exits the gel. The wells are washed prior to loading each of the four rows.*

14. Remove the gel from the apparatus and cool it under running water. Separate the plates by using a thin putty knife, taking care to notice which plate the gel will remain stuck to and allowing it to do so.
15. Find the region of the gel containing the extended primers by surveying with a Geiger counter.
16. Place a large piece of plastic wrap over the gel, smooth it out, and cut the relevant area.
17. Carefully transfer the gel (covered on one side with plastic wrap) to a 35 × 45-cm sheet of chromatography paper, gel-side down, and smooth it out. Dry the gel completely (time may vary with apparatus) and allow it to cool nearly to room temperature.
18. Cut the chromatography paper down to size, if necessary, and transfer the gel to a storage phosphor screen, plastic-wrap side facing the screen. Allow image to develop overnight.

#### **Analyze data**

19. Remove the gel from the storage screen and read the plate with a PhosphorImager.
20. Using ImageQuant software, draw rectangles around primer and extended primer and obtain a volume report of the quantitative data from the gel quantitative data. Use the control reaction (no triphosphate) as the background to be subtracted in each series of triphosphate concentrations.
21. Calculate the fraction of primer extended at each triphosphate concentration as:  

$$v = I_p / (I_0 + I_p)$$
 where  $I_0$  is the primer band count and  $I_p$  is the product band count.
22. Subtract the value obtained for the control reaction from each reaction.  
*This will force the data set to the origin. Make sure that <20% of the primer is extended at the highest triphosphate concentration to ensure single completed hit conditions.*
23. Multiply each corrected value by the final concentration of primer template (40 nM) and divide by the final concentration of enzyme and the reaction time. This gives  $k_{\text{cat}}$  values in  $\text{min}^{-1}$ .
24. Plot  $k_{\text{cat}}$  against triphosphate concentration and fit the curve to the nonlinear form of the Michaelis-Menten equation to obtain  $k_{\text{cat}}$  and  $K_M$ .
25. After a preliminary determination of  $K_M$ , adjust the experimental conditions to obtain data points in the Michaelis plot that evenly flank the experimentally determined  $K_M$  to triphosphate concentrations ~5-fold under and over this value.
26. Collect data in triplicate and average the values for  $k_{\text{cat}}$  and  $K_M$ .

## REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2A*; for suppliers, see *SUPPLIERS APPENDIX*.

### ***Ammonium molybdate solution***

2.0 g ammonium molybdate  
7.0 g ceric sulfate  
250 mL 10% (v/v) H<sub>2</sub>SO<sub>4</sub>  
Store indefinitely at 4°C

### ***Argon***

Argon should be purged through two Drierite desiccators and then introduced into the reaction flask via Tygon tubing installed on a manifold. To ensure a slightly positive pressure of argon in the reaction flask, flow rates should be adjusted by regulators on the argon tank so that a slow stream of argon bubbles (one bubble per sec) is obtained in the bubbler connected to the manifold.

### ***Enzyme dilution buffer***

50 mM potassium phosphate, pH 7.0 (*APPENDIX 2A*)  
1 mM DTT  
0.1 mM EDTA  
50% (v/v) glycerol  
Store up to 6 months at –20°C

### ***Quench solution***

95% formamide  
5 mM Tris·Cl, pH 8 (*APPENDIX 2A*)  
20 mM EDTA  
0.05% (w/v) xylene cyanol  
0.05% (w/v) bromphenol blue

### ***Random-prime buffer, 10×***

500 mM Tris·Cl, pH 7.5 (*APPENDIX 2A*)  
100 mM MgCl<sub>2</sub>  
10 mM DTT  
500 µg/mL acetylated BSA  
Store indefinitely at –20°C

### ***Sodium hydride, 60%***

Purchase sodium hydride (NaH) as a 60% dispersion in mineral oil (Aldrich). Add to a flask and then add hexanes with a pipet. Mix to rinse and remove hexanes with a pipet. Repeat several times until the mineral oil is completely removed, then remove residual hexanes thoroughly using a high-vacuum oil pump.

### ***T<sub>m</sub> buffer, 2×***

20 mM PIPES, pH 7  
20 mM MgCl<sub>2</sub>  
200 mM NaCl  
Store indefinitely at room temperature

## COMMENTARY

### Background Information

#### *Synthesis of unnatural nucleosides*

The first synthesis of natural nucleosides was published in the late 1940s. Since then, nucleoside synthesis has enjoyed a long and rich history. To date, there are five typical methods for the synthesis of *N*-nucleosides (Vorbrüggen and Ruh-Pohlenz, 2001): (1) coupling of heterocyclic salts with halo sugars; (2) Silyl-Hilbert-Johnson reaction followed by Vorbrüggen procedure; (3) addition of heterocycle to glycals; (4) derivation of glycosylamine derivatives; and (5) enzymatic transglycosylation. The first two methods have proven to be the most useful for nucleosidic bond formation. The use of sodium salts (Kazimierczuk et al., 1984) gives considerably higher yield of nucleoside product and is a significant improvement upon the Fischer-Helferich protocol (Fischer and Helferich, 1914), which uses toxic mercuric reagents.

The reaction of peracylated sugars with per-silylated heterocyclic nucleobases in the presence of a Lewis acids such as  $\text{SnCl}_4$  (Niedballa and Vorbrüggen, 1974) or trimethylsilyl triflate (TMSOTf; Vorbrüggen et al., 1981) has become a routine method for the synthesis of pyrimidine and purine nucleosides. TMSOTf is often the catalyst of choice, as it gives higher yields of  $\beta$  anomer and avoids the emulsion problems associated with tin residue during workup.

Natural and modified *C*-nucleosides, in addition to being candidates for unnatural bases with interesting thermal or replication properties, have attracted great interest due to their anticancer and antiviral activities (Postema, 1992; Chaudhuri et al., 1997). There are many methods for synthesizing aromatic *N*- and *C*-nucleosides. For the purpose of preparing *C*-nucleoside-containing oligonucleotides, two main approaches have been used: (1) nucleophilic displacement of a halogen group from C1 of the deoxyribose by aryl Grignard reagents (Chaudhuri et al., 1997; Ogawa et al., 2000a) and (2) *C*-nucleophilic addition to the aldehyde of the deoxyribose precursor (Solomon and Hopkins, 1993; Eaton and Millican, 1988). The merit of the former method is its simplicity and the fact that the desired free nucleoside is directly obtained after deprotection of the hydroxy group. The latter method generally affords the desired product in higher yield than the former. As an alternative method for synthesis of *C*-nucleoside, ribonolactone

has been used with aryl reagents (Hildbrand et al., 1997). The step of nucleophilic addition in this method results in the intermediate formation of a corresponding hemiacetal, which selectively yields the  $\beta$  anomer upon reduction. Removing the 2'-hydroxy group, however, requires additional steps.

#### *Universal bases and unnatural base pairs*

Unnatural nucleosides that possess hydrophobic base analogs have been shown to form stable pairs in duplex DNA, and have also been demonstrated to be efficiently recognized by DNA polymerases, both when present in a template and as a nucleoside triphosphate. In fact, a wide variety of such unnatural bases, with no H bonding or shape complementarity, are surprisingly able to replace a natural base in duplex DNA or during enzymatic replication. These bases are not constrained to resemble the natural purines or pyrimidines, and a much wider variety of design strategies may be pursued. A variety of predominantly hydrophobic nucleobases have been developed with properties such as stable but nonselective pairing in duplex DNA (universal thermodynamic base), efficient but nonselective polymerase recognition (universal kinetic base), or orthogonal stability and replication (third base pair for expanding the genetic code).

This unit describes the synthesis of both *N*- and *C*-nucleosides. The *N*-nucleosides retain the natural glycosidic linkage and maintain the minor groove carbonyl group hydrogen bond acceptor. The minor groove hydrogen bond acceptor may be important for stability in duplex DNA by binding water molecules, and also for polymerase recognition. The *C*-glycosides are aromatic, in order to retain planarity, and, to date, those examined (i.e., TM, DM, 2Nap, DMN, 3MN, and 2MN) do not possess minor groove hydrogen bond acceptors.

The current set of bases has resulted from the derivatization of several scaffolds that were originally found to be successful: the isocarbostyryl framework (Fig. 1.5.2), the 7-azaindole framework (Fig. 1.5.4), and the aromatic ring framework (Fig. 1.5.6). The thermodynamic and kinetic characterization of the modified scaffold structures was used to direct their further derivatization in a largely empirical fashion. Although this process has been largely empirical, a few general trends are beginning to emerge (McMinn et al., 1999; Berger

et al., 2000, 2002; Ogawa et al., 2000a,b; Wu et al., 2000; Tae et al., 2001).

The development of universal bases that pair nondiscriminantly opposite any natural base would have useful applications for the design of hybridization probes where sequence ambiguities exist. There has been some success in the design of nucleobase analogs that can hybridize nonselectively to each of the native bases (reviewed in Loakes, 2001), but analogs that can universally hybridize without significant duplex destabilization are rare. Hydrophobic base analogs without H-bonding groups that pack efficiently in duplex DNA show little selectivity in pairing with native bases and are promising universal bases (Berger et al., 2000). For example, **MICS** and **5MICS** (Fig. 1.5.2) form stable base pairs with each of the natural bases, when incorporated into the central position of either oligodeoxyribonucleotide strand of a 13-mer duplex (**S.22** and **S.23**, shown in Fig. 1.5.9). When incorporated into **S.22** at position X (Figure 1.5.9) and paired opposite each of the four natural bases in **S.23**, the average  $T_m$  for **MICS** and **5MICS** is  $54.6 \pm 0.7^\circ\text{C}$  and  $55.3 \pm 0.6^\circ\text{C}$ , respectively. This compares favorably with the stability of the same duplex with a native dT:dA pair ( $T_m = 58.7^\circ\text{C}$ ). The spread in measured  $T_m$  values for both **MICS** and **5MICS** opposite each natural base is the same,  $1.4^\circ\text{C}$ . A single **MICS** or **5MICS** paired opposite any natural base is significantly more stable than a mispair in the same sequence context ( $\Delta T_m = 3.8^\circ$  to  $17^\circ\text{C}$  for mispairs among the native bases in the same sequence context). When **MICS** and **5MICS** are examined in the opposite strand (position X in **S.23**), paired opposite natural bases, the measured  $\Delta T_m$ 's are slightly larger ( $\Delta T_m = 1.9^\circ$  and  $2.2^\circ\text{C}$  for **MICS** and **5MICS**, respectively).

Less effort has been directed toward the identification of universal bases for DNA replication (Berger et al., 2000). No base analog is known that can act as a universal template for incorporation of each natural nucleoside triphosphate. Such bases would have important applications in the generation of oligonucleotide libraries. A variety of predominantly hydrophobic nucleobases that efficiently direct Klenow fragment to incorporate native triphosphates have been identified. Despite a compromise in absolute rate, when present in template, **MICS** directs the insertion of each native triphosphate with approximately equal efficiency (Berger et al., 2000). Each natural triphosphate is inserted by Klenow fragment opposite **MICS** with an efficiency of  $5.8 \times 10^3$

to  $2.4 \times 10^4 \text{ M}^{-1} \text{ min}^{-1}$ , with dATP and dTTP inserted slightly faster than dCTP and dGTP. Klenow fragment inserts each native triphosphate opposite **MICS** with only a four-fold variation in efficiency, and can synthesize full-length DNA with high concentrations of dNTPs.

The unnatural nucleoside triphosphate, **dPIMTP**, behaves as a universal chain terminator. The triphosphate is inserted opposite any natural base in the template with remarkable efficiency (only ~40-fold reduced relative to insertion of a correct triphosphate) and with only 5-fold discrimination. After incorporation of **dPIMTP**, Klenow fragment is unable to continue DNA synthesis, even under forcing conditions (0.5-hr incubation time and 1 mM dNTPs). Therefore, **PIM** inserts efficiently but randomly into the growing oligonucleotide strand and then terminates synthesis. The average length of synthesized oligonucleotide may be tuned by choice of chain terminator concentration. The efficient insertion of **dPIMTP** opposite any native template results in the efficient generation of random-length oligonucleotides, with low concentrations of a single chain terminator. This contrasts with the conventional dideoxy method of chain termination, in which each dideoxynucleotide must be used in great excess to compete with more efficient incorporation of natural nucleoside triphosphates. For example, Klenow fragment prefers the natural dNTP substrates by several thousand-fold relative to the chain-terminator dideoxy analogs. In this regard, **dPIMTP** may have practical applications, for example, in the generation of random-length oligonucleotides for sequencing by mass spectrometry.

In contrast to nucleobases designed to be universal bases, nucleobases designed as third-base-pair candidates should have orthogonal properties relative to the natural bases. They must form stable pairs in duplex DNA, as well as during enzymatic replication, only with their designed base pair partners. Much attention has been focused on unnatural nucleobases with phenyl, naphthyl, isocarboxystyryl, or 7-azaindole scaffolding. Modifications may be made and then evaluated in terms of base-pair stability ( $T_m$ ) and replicability (steady-state kinetics). For example, **d2MNTP** is efficiently inserted by Klenow fragment opposite **2MN** (**S.21**; Fig. 1.5.8) in the template (self-pair formation) with an efficiency of  $4.4 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$ , but is also efficiently inserted opposite dA ( $1.0 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$ ). It was reasoned that this might result from favorable interactions in the developing

minor groove between the **2MN** methyl group and the methine of dA. Therefore, the nucleobase was modified, resulting in **3MN**. The mispairing between **3MN** and dA was selectively reduced ( $7.6 \times 10^3 \text{ M}^{-1} \text{ min}^{-1}$ ), resulting in a **3MN:3MN** unnatural base pair that is synthesized only 10-fold more slowly than natural base pairs ( $2.4 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$ ) and with at least two orders of magnitude selectivity against all possible mispairs.

Based on thermodynamic and kinetic characterization, a variety of self-pairs and heteropairs have been identified that are attractive candidates as third base pairs. These include the **7AI:7AI**, **PICS:PICS**, and **3MN:3MN** self-pairs, as well as the **7AI:ICS** and **PP:MICS** heteropairs (Figure 1.5.10). It should be noted that the use of a self-pair in no way limits the effort to expand the genetic code. In fact, addition of a self-pair to the genetic alphabet would result in 64 new codons. Each of these unnatural base pairs are stable and replicable with reasonable efficiency and selectivity.

The rate-limiting step for the synthesis of DNA with each of the unnatural base pairs is extension, i.e., the continued primer extension after incorporation of an unnatural base. This step may be optimized by design and direct evaluation of steady-state extension rates, or by examining different polymerases. Progress along both lines has been published. For example, it was found that Klenow fragment and polymerase  $\beta$  have complementary recognition of the **7AI:7AI** self-pair. Although Klenow fragment efficiently inserts d**7AITP** opposite **7AI** in the template, it is unable to continue primer extension. Polymerase  $\beta$ , while unable to synthesize the base pair, extends it with a

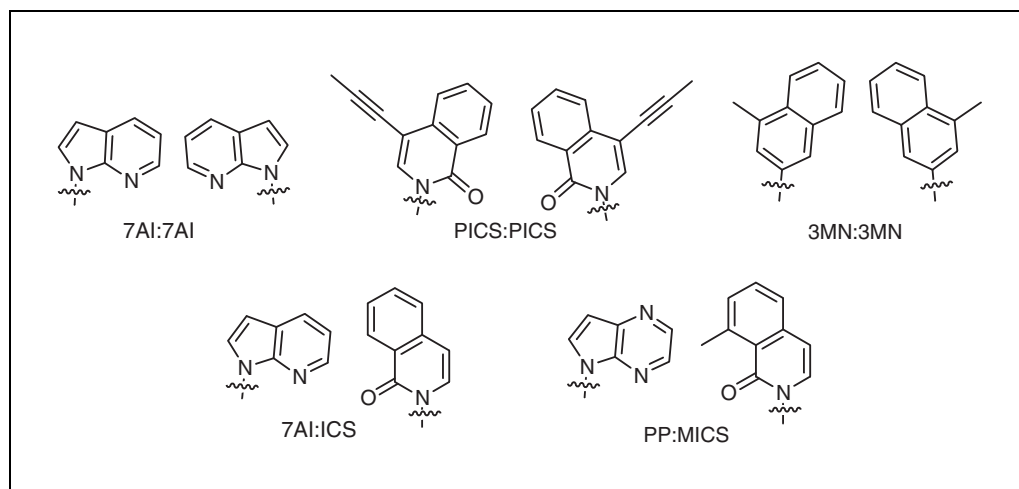
wild-type-like efficiency. The action of both polymerases allows for the synthesis of DNA containing the **7AI:7AI** self-pair in addition to dG:dC and dA:dT (Lee et al., 2001). Further derivativization of the unnatural bases is also being evaluated with respect to extension. In particular, the effects of aromatic surface area and a minor groove hydrogen-bond acceptor are being evaluated.

It is apparent that hydrophobicity is a suitable force for controlling internucleobase interactions in duplex DNA as well as during enzymatic replication. The approach is expected to be more versatile than one based on purine or pyrimidine analogs because a wider range of nucleobase analogs may be examined. It is perhaps surprising that such a wide range of nucleobase structures are tolerated in the duplex and polymerase environments. Nonetheless, this wide tolerance of structures may be taken advantage of to design nucleobase analogs with desired thermal and kinetic properties for use as a universal base or to expand the genetic code.

### Compound Characterization

**S.4:**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.39 (1H, dd,  $J = 8.1, 1.1 \text{ Hz}$ ), 7.94–7.98 (4H, m), 7.91 (1H, d,  $J = 7.8 \text{ Hz}$ ), 7.70 (2H, m), 7.51 (1H, m), 7.20–7.27 (4H, m), 6.86 (1H, dd,  $J = 8.6, 5.4 \text{ Hz}$ ), 5.62 (1H, m), 4.79 (1H, dd,  $J = 12.2, 3.5 \text{ Hz}$ ), 4.70 (1H, dd,  $J = 12.1, 3.2 \text{ Hz}$ ), 4.60 (1H, m), 2.85 (1H, ddd,  $J = 14.3, 5.5, 1.6 \text{ Hz}$ ), 2.42 (3H, s), 2.38 (3H, s), 2.32 (1H, ddd,  $J = 14.6, 8.4, 6.8 \text{ Hz}$ ), 1.94 (3H, s).

**S.5:**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  8.28 (1H, m), 7.97 (1H, s), 7.94 (1H, dd,  $J = 8.1, 0.6 \text{ Hz}$ ), 7.76 (1H, m), 7.54 (1H, m), 6.63 (1H, dd,



**Figure 1.5.10** Unnatural base pairs for expansion of the genetic alphabet.



$J = 7.3, 6.2$  Hz), 4.4 (1H, m), 3.98 (1H, m), 3.83 (1H, dd,  $J = 11.9, 3.5$  Hz), 3.76 (1H, dd,  $J = 12.0, 3.8$  Hz), 2.41 (1H, ddd,  $J = 13.7, 6.2, 3.5$  Hz), 2.17 (1H, ddd,  $J = 13.6, 7.4, 6.5$  Hz), 2.10 (3H, s).  $^{13}\text{C}$  NMR (125 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  160.0, 135.2, 131.7, 129.0, 126.0, 125.9, 123.5, 101.5, 87.6, 86.5, 84.1, 71.9, 69.8, 60.3, 39.7. HRMS (MALDI) anal. calcd. for  $\text{C}_{17}\text{H}_{17}\text{NO}_4\text{Na}$  ( $\text{MNa}^+$ ) 322.1055; found 322.1045.

**S.6:**  $^1\text{H}$  NMR (400 M Hz,  $\text{CDCl}_3$ )  $\delta$  8.40 (1H, d,  $J = 8.1$  Hz), 7.93 (2H, m), 7.70 (1H, m), 7.50 (3H, m), 7.37–7.41 (4H, m), 7.18–7.30 (4H, m), 6.82 (4H, m), 6.71 (1H, ddd,  $J = 12.8, 7.6, 5.7$  Hz), 4.56 (1H, m), 4.24 (1H, m), 3.74–3.83 (8H, m), 3.52–3.65 (4H, m), 3.45 (1H, m), 3.33 (1H, m), 2.67 (1H, m), 2.61 (2H, t,  $J = 6.5$  Hz), 2.42 (2H, t,  $J = 6.5$  Hz), 2.26 (1H, m), 1.82 (3H, d,  $J = 11.9$  Hz), 1.17 (9H, m), 1.06 (3H, d,  $J = 6.8$  Hz).  $^{31}\text{P}$  NMR (160 MHz,  $\text{CDCl}_3$ )  $\delta$  51.4, 50.7.

**S.7:**  $^{31}\text{P}$  NMR (160 M Hz, 50 mM Tris (pH 7.5), 2mM EDTA in  $\text{D}_2\text{O}$ )  $\delta$  -6.11 (d,  $J = 19.2$  Hz), -10.38 (d,  $J = 19.9$  Hz), 21.55 (dd,  $J = 19.9, 19.2$  Hz).

**S.8:**  $^1\text{H}$  NMR (400 M Hz,  $\text{CDCl}_3$ )  $\delta$  8.38 (1H, m), 8.08 (2H, d,  $J = 6.4$  Hz), 8.02 (2H, d,  $J = 6.5$  Hz), 7.78 (1H, d,  $J = 2.1$  Hz), 7.36 (2H, d,  $J = 6.4$  Hz), 7.34 (1H, s), 7.33 (2H, d,  $J = 6.7$  Hz), 6.91 (1H, dd,  $J = 6.9, 4.4$  Hz), 6.83 (1H, m), 5.87 (1H, m), 4.80 (1H, dd,  $J = 9.7, 3.2$  Hz), 4.73 (1H, dd,  $J = 9.4, 3.2$  Hz), 4.69 (1H, m), 3.08 (1H, ddd,  $J = 11.6, 6.6, 5.0$  Hz), 2.83 (1H, ddd,  $J = 11.4, 4.5, 1.8$  Hz), 2.52 (3H, s), 2.50 (3H, s).

*Deprotected S.8:*  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  8.37 (1H, d,  $J = 2.7$  Hz), 8.24 (1H, d,  $J = 2.7$  Hz), 8.05 (1H, d,  $J = 3.8$  Hz), 6.68 (1H, dd,  $J = 8.1, 6.0$  Hz), 6.67 (1H, d,  $J = 3.8$  Hz), 4.53 (1H, m), 3.99 (1H, dd,  $J = 6.6, 4.1$  Hz), 3.77 (1H, dd,  $J = 12.0, 3.5$  Hz), 3.70 (1H, dd,  $J = 12.0, 4.1$  Hz), 2.72 (1H, ddd,  $J = 13.8, 7.7, 6.0$  Hz), 2.35 (1H, ddd,  $J = 13.5, 6.2, 3.0$  Hz); HMRS anal. calcd. for  $\text{C}_{11}\text{H}_{14}\text{N}_3\text{O}_3$  ( $\text{MH}^+$ ) 236.1035; found 236.1034.

**S.11:**  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.99 (4H, m), 7.28 (2H, d,  $J = 8.1$  Hz), 7.24 (2H, d,  $J = 8.1$  Hz), 7.02 (2H, s), 6.92 (1H, s), 5.62 (1H, m), 5.20 (1H, dd,  $J = 10.9, 5.0$  Hz), 4.71 (1H, dd,  $J = 11.8, 4.0$  Hz), 4.66 (1H, dd,  $J = 11.8, 3.6$  Hz), 4.53 (1H, m), 2.50 (1H, m), 2.44 (3H, s), 2.41 (3H, s), 2.66 (7H, m). HRMS anal. calcd. for  $\text{C}_{29}\text{H}_{31}\text{O}_5$  ( $\text{MH}^+$ ) 459.2171; found 459.2179.

**S.12:**  $^{31}\text{P}$  NMR (140 MHz,  $\text{CDCl}_3$ )  $\delta$  148.5, 148.3.

**S.18:**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.07 (1H, m), 7.98 (1H, m), 7.50 (2H, m), 7.45 (1H,

s), 5.24 (1H, dd,  $J = 11.5, 1.7$  Hz), 4.23 (1H, m), 3.96 (2H, m), 3.81 (1H, m), 2.67 (3H, s), 2.62 (3H, s), 2.10 (1H, m), 1.88 (1H, ddd,  $J = 14.2, 11.6, 2.3$  Hz).  $^{13}\text{C}$  NMR (150 MHz,  $\text{CDCl}_3$ )  $\delta$  135.5, 132.8, 132.7, 132.1, 127.6, 125.7, 125.3, 124.7, 124.5, 123.8, 76.0, 70.9, 69.2, 68.0, 36.5, 19.5, 13.8. HRMS anal. calcd. for  $\text{C}_{13}\text{H}_{24}\text{NO}_3$  ( $\text{MNH}_4^+$ ) 290.1756; found 290.1764.

**S.21:**  $^1\text{H}$  NMR (600 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  8.03 (1H, s), 7.78 (1H, d,  $J = 7.1$  Hz), 7.71 (1H, d,  $J = 7.1$  Hz), 7.59 (1H, s), 7.37 (2H, m), 5.44 (1H, dd,  $J = 10.1, 5.3$  Hz), 4.35 (1H, m), 3.98 (1H, m), 3.76 (2H, m), 2.47 (3H, s), 2.36 (1H, ddd,  $J = 13.2, 5.7, 1.8$  Hz), 1.90 (1H, m).  $^{13}\text{C}$  NMR (150 MHz,  $\text{CDCl}_3$ )  $\delta$  139.8, 133.8, 133.6, 133.1, 128.6, 128.2, 127.3, 126.1, 125.6, 124.1, 88.2, 78.2, 73.8, 63.5, 42.8, 19.1. HRMS anal. calcd. for  $\text{C}_{16}\text{H}_{18}\text{O}_3$  ( $\text{MNa}^+$ ) 281.1154; found 281.1158.

## Critical Parameters and Troubleshooting

Generally, any reaction that involves moisture-sensitive reagents should be handled with care. The reaction equipment, including flasks, rubber septa, adapters, gas or liquid condensers, and syringes, must be predried in an oven or under an infrared lamp and flushed with dry argon before use. Anhydrous solvents are critical to the success of reactions. It is best to freshly distill solvents and use moisture-sensitive reagents in situ.

**Glycosylation.** The Lewis acids used in Silyl-Hilbert-Johnson reactions such as Tin(IV) tetrachloride are generally moisture-sensitive. The quality of these reagents dramatically affects the yield of product.

**Iodination.** The reaction time is very critical. Excess time may cause significant side reactions.

**Propynylation.** The reaction system should be absolutely air-free, since copper(I) iodide and the palladium catalyst used are generally air- and/or light-sensitive. The pressure tube should be sealed very well, otherwise propyne collected at low temperature will evaporate readily when the reaction is stirred at ambient temperature.

**Deprotection.** There are no particular difficulties with this step.

**Tritylation.** The substrate should be thoroughly dried by coevaporation with anhydrous pyridine prior to use. Slow addition of the solution of DMTr-Cl in pyridine or batchwise addition of DMTr-Cl solid is necessary; other-

wise, the yield of product will be greatly reduced.

**Phosphitylation.** 2-Cyanoethyl diisopropylchlorophosphoramidite should be strictly stored at  $-20^{\circ}\text{C}$  and warmed to ambient temperature before use. Quickly transferring this reagent via a short needle to the reaction mixture produces optimal results.

**Phosphorylation.** The moisture contained in the substrate should be removed by azeotropic evaporation from a small amount of anhydrous pyridine before the other reactants are added.

**$S_N2$  displacement reactions.** Sodium hydride should be dried and used in situ and quickly added into the reaction flask in several batches, because dry sodium hydride decomposes very readily during storage.

**Evaporation of solvents.** A rotatory evaporator is routinely used to remove solvents. The efficiency of evaporation is largely dependent on the vacuum (normally between 10 to 20 mm Hg) and the dry-ice condenser. It is better to maintain the bath temperature below  $45^{\circ}\text{C}$  to prevent any chemical change of the crude product. To avoid loss of compound caused by a sudden burst of solvents into the splash-guard adapter, the evaporation flask should be cooled by slow rotary evaporation under low vacuum before it is immersed into the bath.

**Grignard reaction.** This reaction should be carried out with exclusion of moisture. All apparatus must be thoroughly dried in a hot ( $>120^{\circ}\text{C}$ ) oven before use. Use anhydrous solvents.

**Deprotection of toluoyl group.** There are no particular difficulties with this reaction; it will work effectively.

**Nucleophilic addition by aryllithium reagents.** The reaction is moisture-sensitive. Anhydrous solvent is important for the reaction. Do not use *n*-butyllithium reagent with a milky precipitate, because such a reagent is decomposed in the presence of water.

**Removal of protecting group of hydroxy group.** Difficulties are seldom encountered in this step.

### Anticipated Results

An overall good yield of the free nucleoside (48% to 68%) of unnatural hydrophobic nucleobases can be achieved based on the strategies described in this unit. The yields of the nucleosidic glycosylation are generally good enough for the preparation of phosphoramidites and triphosphates, although stereoselectivity of the desired  $\beta$ -anomer versus  $\alpha$ -anomer may be mediocre in some cases.

The routine deprotection, phosphitylation, and phosphorylation are generally high-yielding.

Although the Grignard addition described in Basic Protocol 3 generally favors the  $\alpha$ -anomer, acid treatment of the corresponding benzylic ethers in refluxing xylene will result in epimerization at C1' to yield the  $\beta$ -anomer (Ren et al., 1996). Conversely, the aryllithium addition described in Basic Protocol 4 and the Alternate Protocol generally affords the desired product in a good yield. The product is a mixture of diastereomers, but is separable by column chromatography.

### Time Considerations

The estimated time to complete Basic Protocol 1 in its entirety is 3 weeks. The estimated time required to complete Basic Protocol 2 is 2 days. Basic Protocol 3 should take ~6 days to complete. For Basic Protocol 4 and Alternate Protocol, allow ~1 day for the Grignard reaction and several hours for all of the other reactions.

For Basic Protocol 5, 2 to 3 days should be sufficient time to prepare the oligonucleotide. On day 1, the oligonucleotide is synthesized and the deprotection should be done overnight. Several hours should be allowed for gel pouring and polymerization; this step may also be done on day 1 or in the morning of day 2. The gel must be kept hydrated, regardless, by wrapping with plastic wrap. The electrophoresis step can take several hours. After separation, the excised gel slice may be stored at  $-20^{\circ}\text{C}$  for further purification at a later date. To perform the electroelution, allow  $<2$  hr. Ethanol precipitation can then be done overnight at  $-20^{\circ}\text{C}$  and the final steps done the following morning.

For Basic Protocol 7, about 1 hr should be allowed for radiolabeling the primer. Allow 2 hr for annealing the primer-template duplex, during which time the gel should be poured and allowed to polymerize. Annealed primer-templates can be stored at  $4^{\circ}\text{C}$  overnight, and the gel can also be stored overnight at room temperature (it must be protected from dehydration by first washing the wells with buffer, then covering with paper towels and wrapping tightly in plastic wrap). Allow the annealed primer-templates to return to room temperature before use. The time it takes to aliquot the triphosphates and run the reactions depends on the number and nature of the reactions. After final loading of the gel, it can be run at 90 W for about 1.5 hr. Excising the portion of the gel containing the radioactive bands and subsequent drying onto filter paper should take no

more than 1 hr. Finally exposure of the phosphor storage screen is routinely done overnight.

## Literature Cited

- Beaussire, J.J. and Pochet, S. 1999. Recognition of 2'-deoxyisoinosine triphosphate by the Klenow fragment of DNA polymerase I. *Nucleosides Nucleotides* 18:403-410.
- Chaudhuri, N.C., Ren, R.X.F., and Kool, E.T. 1997. C-Nucleoside derived from simple aromatic hydrocarbons. *Synlett*. 341-347.
- Creighton, S., Bloom, L.B., and Goodman, M.F. 1995. Gel fidelity assay measuring nucleotide misinsertion, exonucleolytic proofreading, and lesion bypass efficiencies. *Methods Enzymol.* 262:232-256.
- Eaton, M.A.W. and Millican, T.A. 1988. New methodology for C-nucleoside synthesis: Preparation of 1,2-dideoxy-1-(3-pyridyl)-D-ribofuranose. *J. Chem. Soc. Perkin Trans. I.* 545-547.
- Fischer, E. and Helferich, B. 1914. Synthetische Glucoside der Purine. *Chem. Ber.* 47:210.
- Hilbert, G.E. and Johnson, T.B. 1930. Researches on pyrimidines. CXV. Alkylation on nitrogen of the pyrimidine cycle by application of a new technique involving molecular rearrangements. *J. Am. Chem. Soc.* 52:2001.
- Hildbrand, S., Blaser, A., Parel, S.P., and Leumann, C.J. 1997. 5-Substituted 2-aminopyridine C-nucleosides as protonated cytidine equivalents: Increasing efficiency and selectivity in DNA triple-helix formation. *J. Am. Chem. Soc.* 119:5499-5511.
- Horlacher, J., Hottiger, M., Podust, V.N., Hübscher, U., and Bennier, S.A. 1995. Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns. *Proc. Natl. Acad. Sci. U.S.A.* 92:6329-6333.
- Kazimierczuk, Z., Cottam, H.B., Revankar, G.R., and Robins, R.K. 1984. Synthesis of 2'-deoxytubercidin, 2'-deoxyadenosine, and related 2'-deoxynucleosides via a novel direct stereospecific sodium salt glycosylation procedure. *J. Am. Chem. Soc.* 106:6379-6382.
- Kornberg, A. and Baker, T.A. 1992. DNA Replication, 2nd ed. W.H. Freeman and Co., New York.
- Lambert, J.B., Fabricus, D.M., and Hoard, J.A. 1979. Bond localization approach to the carbon analog of the Claisen rearrangement. Thermolysis of 4-aryl-1-butenes. *J. Org. Chem.* 44:1480-1485.
- Lambert, J.B., Shurvell, H.F., Lightner, D.A., and Cooks, R.G. 1998. Organic Structural Spectroscopy. Prentice Hall, Englewood Cliffs, N.J.
- Lee, E.L., Wu, Y., Xia, G., Schultz, P.G., and Romesberg, F.E. 2001. Efforts toward expansion of the genetic alphabet: Replication of DNA with three base pairs. *J. Am. Chem. Soc.* 123:7439-7440.
- Lutz, M.J., Held, H.A., Hottiger, M., Hübscher, U., and Benner, S.A. 1996. Differential discrimination of DNA polymerases for variants of the non-standard nucleobase pair between xanthosine and 2,4-diaminopyrimidine, two components of an expanded genetic alphabet. *Nucl. Acids Res.* 24:1308-1313.
- Lutz, M.J., Horlacher, J., and Benner, S.A. 1998a. Recognition of 2'-deoxyisoguanosine triphosphate by HIV-reverse transcriptase and mammalian cellular DNA polymerases. *Bioorg. Med. Chem. Letts.* 8:499-504.
- Lutz, M.J., Horlacher, J., and Benner, S.A. 1998b. Recognition of a non-standard base pair by thermostable DNA polymerases. *Bioorg. Med. Chem. Letts.* 8:1149-1152.
- McMinn, D.L., Ogawa, A.K., Wu, Y., Liu, J., Schultz, P.G., and Romesberg, F.E. 1999. Efforts toward expansion of the genetic alphabet: DNA polymerase recognition of a highly stable, self-pairing hydrophobic base. *J. Am. Chem. Soc.* 121:11585-11586.
- Niedbala, U. and Vorbrüggen, H. 1970. A general synthesis of pyrimidine nucleosides. *Angew. Chem. Int. Ed. Engl.* 9:461.
- Niedbala, U. and Vorbrüggen, H. 1974. A general synthesis of N-glycosides. I. Synthesis of pyrimidine nucleosides. *J. Org. Chem.* 39:3654-3660.
- Ogawa, A.K., Wu, Y., McMinn, D.L., Liu, J., Schultz, P.G., and Romesberg, F.E. 2000a. Efforts toward the expansion of the genetic alphabet: Information storage and replication with unnatural hydrophobic base pairs. *J. Am. Chem. Soc.* 122:3274-3287.
- Ogawa, A.K., Wu, Y., Berger, M., Schultz, P.G., and Romesberg, F.E. 2000b. Rational design of an unnatural base pair with increased kinetic selectivity. *J. Am. Chem. Soc.* 122:8803-8804.
- Postema, M.H.D. 1992. Recent developments in the synthesis of C-glycosides. *Tetrahedron* 48:8545-8599.
- Ren, R.X.F., Chaudhuri, N.C., Paris, P.L., Rumney S., and Kool, E.T. 1996. Naphthalene, phenanthrene, and pyrene as DNA base analogues: Synthesis, structure, and fluorescence in DNA. *J. Am. Chem. Soc.* 118:7671-7678.
- Roberts, C., Chaput, J.C., and Switzer, C. 1997a. Beyond guanine quartets: Cation-induced formation of homogenous and chimeric DNA tetraplexes incorporating iso-guanine and guanine. *Chem. & Biol.* 4:899-908.
- Roberts, C., Bandaru, R., and Switzer, C. 1997b. Theoretical and experimental study of isoguanine and isocytosine: Base pairing in an expanded genetic system. *J. Am. Chem. Soc.* 119:4640-4649.
- Robinson, H., Gao, Y.-G., Bauer, C., Roberts, C., Switzer, C., and Wang, A.H. 1998. 2'-Deoxyisoguanosine adopts more than one tautomer to form base pairs with thymidine observed by high-resolution crystal structure analysis. *Biochemistry* 37:10897-10905.

- Sharma, P.K. 1993. A new synthesis of trans-3,4-dihydroxy-anti-1,2-epoxy-1,2,3,4-tetrahydro-7,12-dimethylbenz[a]anthracene. *Synth. Commun.* 23:389-394.
- Solomon, M.S. and Hopkins, P.B. 1993. Chemical synthesis and characterization of duplex DNA containing a new base pair: A nondisruptive, benzofused pyrimidine analog. *J. Org. Chem.* 58: 2232-2243.
- Takeshita, M., Chang, C.N., Johnson, F., Will, S., and Grollman, A.P. 1987. Oligodeoxynucleotides containing synthetic abasic sites. *J. Biol. Chem.* 262:10171-10179.
- Vorbrüggen, H. and Ruh-Pohlenz, C. 2001. Handbook of Nucleoside Synthesis. John Wiley & Sons, New York.
- Vorbrüggen H., Kroliekiewicz, K., and Bennua, B. 1981. Nucleoside synthesis with trimethylsilyl triflate and perchlorate as catalysts. *Chem. Ber.* 114:1234-1235.
- Wu, Y., Ogawa, A.K., Berger, M., McMinn, D.L., Schultz, P.G., and Romesberg, F.E. 2000. Efforts toward expansion of the genetic alphabet: Optimization of interbase hydrophobic interactions. *J. Am. Chem. Soc.* 122:7621-7632.

### Key References

- Berger, M., Wu, Y., Ogawa, A.K., McMinn, D.L., Schultz, P.G., and Romesberg, F.E. 2000. Universal bases for hybridization, replication and chain termination. *Nucl. Acids Res.* 28:2911-2914.

*This paper describes the evaluation of a variety of unnatural bases as universal bases for hybridization and replication.*

- Berger, M., Luzzi, S.D., Henry, A.A., and Romesberg, F.E. 2002. Stability and selectivity of unnatural DNA with five-membered-ring nucleobase analogues. *J. Am. Chem. Soc.* 124:1222-1226.

*This paper describes the evaluation of 5-membered ring nucleobase analogs.*

- Lee et al., 2001. See above.

*This paper describes the replication of DNA containing the 7AI self-pair with the Klewnow fragment/pol β binary polymerase system.*

- Ogawa et al., 2000a. See above.

*This paper describes the synthesis of the C-nucleosides reported herein.*

- Ogawa et al., 2000b. See above.

*This paper describes the rational design of the 3MN:3MN self-pair.*

---

Contributed by Floyd E. Romesberg,  
Chengzhi Yu, Shigeo Matsuda, and  
Allison A. Henry  
The Scripps Research Institute  
La Jolla, California