Part I. Tandem Aldol-Allylation Reactions Promoted by Strained Silacycles

Part II. Design and Synthesis of Modified Fluorescent Nucleotides for DNA Sequencing by Synthesis

Qinglin Meng

Submitted in partial fulfillment of

the requirements for the degree of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2006

Illumina Ex. 1116 Illumina v. Columbia IPR2018-00291, -00318, -00322, -00385, and -00797

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.



UMI Microform 3237290 Copyright 2007 by ProQuest Information and Learning Company. All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

> ProQuest Information and Learning Company 300 North Zeeb Road P.O. Box 1346 Ann Arbor, MI 48106-1346

© 2006

,

Qinglin Meng All Rights Reserved

ABSTRACT

- Part I. Tandem Aldol-Allylation Reactions Promoted by Strained Silacycles
- Part II. Design and Synthesis of Modified Fluorescent Nucleotides for DNA Sequencing by Synthesis

Qinglin Meng

The first part of this thesis presents the tandem aldol-allylation reactions promoted by strained 1,3-dioxa-2-silacyclopentane rings. Our studies have proven that both the aldehyde- and ketone-derived allylenolsilanes could react with aldehydes smoothly in the absence of catalysts to afford homoallylic 1,3-diols with high diastereoselectivity. Pinacol was employed to constitute the strained 1,3-dioxa-2-silacyclopentane ring as an effective platform to activate the reactivity of silane reagents due to the effect of *strain-release Lewis acidity*. The tandem aldol-allylation reactions proceed via a cyclic transition state with a six-membered chair-like conformation (*type I* allylation). Homoallylic 1,3-diols with up to four stereocenters can be generated using tandem aldol-allylation reactions in one step conveniently and rapidly, which has become an important method to construct 1,3-diol fragments for the total synthesis of natural products.

The second part of this thesis describes the design and synthesis of both photocleavable

and chemically cleavable 3' modified fluorescent nucleotides. Four 3' allyl modified nucleoside triphosphates (3'-O-allyl-dNTP-NH₂) were synthesized from readily available chemicals in reasonable yields after which modified fluorescent nucleotides were prepared via coupling reactions of these nucleoside triphosphates with corresponding fluorescent dyes. These fluorescent nucleotides have been applied for DNA sequencing by synthesis both in solution phase and on a glass chip and have proven to be good reversible terminators for DNA SBS. They could be accurately incorporated into a growing DNA strand in the DNA polymerase reaction and the 3' capping group could be cleaved conveniently in aqueous palladium (II)-phosphine solution after fluorophore detection. This valuable approach has offered an alternative DNA analysis method to overcome the disadvantages and limitations of conventional electrophoresis-based sequencing methods.

Table of Contents

Part I. Tandem Aldol-Allylation Reactions Promoted by Strained Silacycles

Chapter 1.	Introduction	2
References		11

Chapter 2. Tandem Aldol-Allylation Reactions Promoted by a Strained 1	,3,2-
Dioxasilolane Cycle	13
2.1 Tandem aldol-allylation reactions using aldehyde-derived allylenolsilanes	14
2.2 Tandem aldol-allylation reactions using ketone-derived allylenolsilanes	21
2.3 Discussions on tandem aldol-allylation reactions of allylenolsilanes	32
2.4 Summary and Outlook	38
2.5 Experimental section	39
References	63

Part II. Design and Synthesis of Modified Fluorescent Nucleotides for DNA Sequencing by Synthesis

Chapter 1.	Introduction to DNA Sequencing by Synthesis	66
1.1 Structure	s and properties of DNA, nucleotides and nucleosides	67

1.2 Sanger dideoxy DNA sequencing	.71
1.3 DNA sequencing by synthesis	.73
1.4 DNA pyrosequencing	.76
1.5 DNA sequencing using mass spectrometry	80
1.6 DNA sequencing by nanopore	.82
References	.85

Chapter 2. Design and Synthesis of Modified Fluorescent Nucleotides for DNA
Sequencing by Synthesis
2.1 Background
2.2 Design of modified fluorescent nucleotides for DNA SBS96
2.3 Synthesis of 3' allyl modified nucleoside triphosphates (1, 2, 3 and 4)102
2.3.1 Synthesis of 3'-O-allyl-7-deaza-7-(3-aminoprop-1-ynyl)-2'-deoxyadenosine-5'-
triphosphate 1102
2.3.2 Synthesis of 3'-O-allyl-7-deaza-7-(3-aminoprop-1-ynyl)-2'-deoxyguanosine-5'-
triphosphate 2108
2.3.3 Synthesis of 3'-O-allyl-5-(3-aminoprop-1-ynyl)-2'-deoxycytidine-5'-triphosphate
3118
2.3.4 Synthesis of 3'-O-allyl-5-(3-aminoprop-1-ynyl)-2'-deoxyuridine-5'-triphosphate
4120
2.4 Synthesis of cleavable linkers

2.4.1 Synthesis of chemically cleavable allyl linker 5	
2.4.2 Synthesis of photocleavable linker 65124	
2.5 Synthesis of fluorescent nucleotides 73 , 77 , 81 , 85 , 86 , 87 , 88 , 89 125	
2.6 Confirmation of allyl position in the modified nucleoside molecules: O-allyl or	
<i>N</i> -allyl129	
2.7 Experimental section	
References159	

Chapter 3. Application of Modified Fluorescent Nucleotides as Reversible
Terminators for DNA Sequencing by Synthesis162
3.1 Survey of methods to cleave 3' allyl group and allyl linker moiety162
3.2 Application of photocleavable modified fluorescent nucleotides as reversible
terminators for DNA sequencing by synthesis167
3.3 Application of photocleavable modified fluorescent nucleotides as reversible
terminators for DNA sequencing on a glass chip172
3.4 Application of chemically cleavable modified fluorescent nucleotides as reversible
terminators for DNA sequencing by synthesis175
3.5 Proposed photocleavage and deallylation mechanisms of modified fluorescent
nucleotides in DNA sequencing by synthesis183
3.6 Experimental section

References	
------------	--

Chapter 4.	Summary and	Outlook	.192
------------	-------------	---------	------

Appendix:	¹ H and ¹³ C NMR Spectra	195
Part I		196
Part II		247

List of Schemes, Figures and Tables

Part I.

Chapter 1. Introduction

Chapter 2. Tandem Aldol-Allylation Reactions Promoted by a Strained 1,3,2-Dioxasilolane Cycle

Scheme 2-2. Noncatalyzed tandem aldol-allylation reactions of allylenolsilane 4 containing a strained 1,3-dioxa-2-silacyclopentane ring with aldehydes by Xiaolun

Wang15

Scheme 2-3. Synthesis of allylenolsilanes 8 and 9 containing a strained 1,3-dioxa-2silacyclopentane cycle, their noncatalyzed tandem aldol-allylation reactions with aldehyde, and stereochemical confirmation of the major tandem products 11 and 12.....17 Scheme 2-4. Xiaolun Wang and Andrew J. Nation's work on noncatalyzed tandem aldol-allylation reaction of substituted allylenolsilanes 15-19 with cyclohexanecarboxaldehyde......21 Scheme 2-5. Synthesis of acetone-derived allylenolsilane 25, its noncatalyzed tandem aldol-allylation reaction with aldehyde, and stereochemical confirmation of tandem Scheme 2-6. Synthesis of acetone-derived allylenolsilanes 33-36 with methyl-substituted Scheme 2-7. Noncatalyzed tandem aldol-allylation reactions of allylenolsilanes 33-36 Scheme 2-9. Synthesis of allylenolsilanes 54 and 55, their noncatalyzed tandem aldol-allylation reactions with aldehyde, and stereochemical confirmation of tandem Scheme 2-10. Reaction pathways of 55 with aldehyde to form 59a, 59b and 60......31 Scheme 2-11. Attempted synthesis of 63-68 and noncatalyzed tandem aldol-allylation

Part II.

Chapter 1. Introduction to DNA Sequencing by Synthesis

Figure 1-1. Names and structures of the four 2'-deoxyribonucleoside 5'-triphosphates
(dNTP's)67
Figure 1-2. Watson-Crick base pairing of double-stranded DNA68
Figure 1-3. Possible labeling sites of fluorescent dyes on the nucleoside triphosphate
molecule76
Figure 1-4. A nucleic acid stretch passes through a nanopore by the applied electric field
(left) and electronic signatures are generated as a result (right)

Chapter 2. Design and Synthesis of Modified Fluorescent Nucleotides for DNA Sequencing by Synthesis

SBS	101
Scheme 2-7. Milton's synthetic strategy of 1	102
Scheme 2-8. Synthetic routes of 6 from 13	103
Scheme 2-9. Our synthesis of 1 from 13	104
Scheme 2-10. Milton's synthetic strategy of 2	108
Scheme 2-11. Synthetic route of 21 from 25	109
Scheme 2-12. Alternative synthetic route of 21 from 29	
Scheme 2-13. Our first attempted synthesis of 37 from 25	112
Scheme 2-14. Attempted synthetic route toward 2 from 36	114
Scheme 2-15. Another attempted synthetic route toward 2 from 36	115
Scheme 2-16. Our final synthesis of 24 from 43	116
Scheme 2-17. Complete synthetic route of 2 from 25	117
Scheme 2-18. Milton's synthetic strategy of 3	118
Scheme 2-19. Milton's synthetic strategy of 4	120
Scheme 2-20. Synthetic strategy of 4 in our previous work	121
Scheme 2-21. Our modified synthetic route of 4 from 52	122
Scheme 2-22. Synthesis of allyl linker 5	124
Scheme 2-23. Synthesis of photocleavable linker 65	
Scheme 2-24. Synthesis of 3'-O-allyl-dNTP-allyl-Dyes 73, 77, 81 and 85	126-127
Scheme 2-25. Synthesis of 3'-O-allyl-dNTP-PC-Dyes 86-89	
Figure 2-1. A transition state complex of DNA template, primer, DNA pol	ymerase and a

2',3'-dideoxycytidine triphosphate (ddCTP) in the DNA polymerase reaction established by H. Pelletier et al. The left figure illustrates the incorporation process of ddCTP onto the 3' end of the primer catalyzed by Mg^{2+} and polymerase active residues. The right figure shows the steric interaction of polymerase active residues with ddCTP in the Figure 2-2. Structures of the four photocleavable fluorescent nucleotides and photocleavage of the 2-nitrobenzyl linker by near-UV irradiation......94 Figure 2-3. Coupling constants of selected β -anomers of 2'-deoxy-D-ribofuranosyl-Figure 2-4. Coupling constants of selected α -anomers of 2'-deoxy-D-ribofuranosyl-Figure 2-5. Substituent effect on the iodination site of 7-deazaguanine......110

Chapter 3. Application of Modified Fluorescent Nucleotides as Reversible Terminators for DNA Sequencing by Synthesis

Scheme 3-2. Genêt's method to cleave an allyl group from allyl esters, allyl carbonates or

allyl carbamates using an aqueous palladium-phosphine system164
Scheme 3-3. Thayumanavan's strategy to cleave an allyl group from allyl esters and allyl
carbamates using a palladium (0) catalyst165
Scheme 3-4. Smith's cleavage of an allyl group by palladium (II) chloride in acetic acid
and sodium acetate
Scheme 3-5. Extension with 3'-O-allyl dNTP-PC-Dye on a DNA template170-171
Scheme 3-6. Illustrative description of immobilizing a DNA template on a glass chip via
a copper (I)-catalyzed 1,3-dipolar cycloaddition reaction173
Scheme 3-7. Extension with 3'-O-allyl dGTP-allyl-Bodipy-FL-510 90 on a repeat C
template178
Scheme 3-8. Extension with 3'-O-allyl dNTP-PC-Dye on a DNA template181-182
Scheme 3-9. Photolysis mechanism of 2-nitrobenzyl linker moiety
Scheme 3-10. Proposed mechanism of cleaving allyl groups from allyl ethers and allyl
carbamates in aqueous palladium-phosphine solution184
Figure 3-1. Structures of fluorescent 3'-O-allyl-dNTP-PC-Dye nucleotides 86-89168
Figure 3-2. Application of photocleavable modified fluorescent 3'-O-allyl-dNTP-PC-
Dye nucleotides 86~89 for DNA sequencing by synthesis on a glass chip175
Figure 3-3. Structures of fluorescent 3'-O-allyl-dNTP-allyl-Dye nucleotides 73, 77, 81
and 85 179
Table 3-1. Deallylation of 3'-O-allylthymidine by Na ₂ PdCl ₄ and TPPTS in water166

Acknowledgements

The first part of this thesis was finished under the guidance of Professor James L. Leighton. Here I would like to thank him for his great support and instruction to me as well as his knowledge and professionalism during the years I was working in his group. It was my pleasure to work together with Xiaolun Wang, Andrew Nation and Nicholas Perl in this tandem aldol-allylation project. I appreciate their intelligence and motivation so much. I also want to express my gratitude to Dr. Katsumi Kubota, Christopher Hamblett and other members in the Leighton group, for all their help.

The second part of this thesis was finished under the guidance of Professor Jingyue Ju and Professor Nicholas J. Turro. I would like to express my most sincere gratitude and appreciation for their guidance and support. Their enthusiasm and professionalism in science will inspire me all my life. Dr. James J. Russo gave me very helpful suggestions on my thesis. I owe special thanks to Dr. Xiaopeng Bai, who gave me so much help in every aspect since I joined this group. We worked together on the triphosphate preparation and he also worked on the HPLC purification of fluorescent nucleotides. Xiaopeng and Shenglong Zhang prepared the fluorescent 3'-O-allyl-dNTP-PC-Dye nucleotides. Dye Hyun Kim, Dr. Hameer Ruparel, Dr. Lanrong Bi, Dr. Xiaoxu Li, Dr. Zengmin Li, Dr. Mong Marma and Shundi Shi worked on the biological evaluation of the fluorescent nucleotides in Chapter 3 of Part II. I also appreciate the great help and suggestions from Dr. Zengmin Li, Dr. Lanrong Bi, Dr. Zhiqiang Liu and other people with whom I have worked.

New England Biolabs is acknowledged for generously providing the 9° N_m DNA polymerase.

Finally, I wish to thank my parents and family for their unselfish support over so many years. I could not have gone so far without their encouragement.

I owe special gratitude to Yingmei, my fiancée. It is her love that makes me happy and helps me overcome the difficulties in my life. She has been and will always be my greatest fortune throughout my life.

for my parents and Yingmei

Part I. Tandem Aldol-Allylation Reactions Promoted by Strained Silacycles

Chapter 1. Introduction

Polyketide units frequently appear in many natural products with stereocenters and constitute very important components of diverse natural products, such as rifamycin B, oleandomycin and rapamycin (Figure 1-1).¹ Many of them have shown very important pharmacological activities. Development of convenient and efficient methods to synthesize stereo-controlled 1,3-diol fragments has attracted the enthusiasm of many chemists, especially in the pharmacology industry. Many methods for preparing 1,3-diol fragments with stereoselectivity have been investigated to date, including alkylation, aldol reaction, reduction or the use of homoallylic alcohols as starting material.² Traditional strategies utilize stepwise transformations to construct such non-adjacent stereocenters, in which the two stereocenters are typically generated stereoselectively in separate steps. A more efficient but challenging approach for this purpose is to employ a tandem reaction strategy so that the two stereocenters can be created stereoselectively in



Figure 1-1. Examples of some polyketides containing 1,3-diol fragment.¹

a single step.³ By this concise method the target products can be produced conveniently and rapidly with higher efficacy and less effort.

Since Mukaiyama reported the aldol addition reactions of silvl enol ethers with aldehydes and ketones promoted by Lewis acids over three decades ago (now named the Mukaiyama reaction),⁴ silvl enol ethers have proven to be powerful reagents for aldol reactions to form new carbon-carbon bonds in organic chemistry. In general Lewis acids are necessary to activate the carbonyl group of aldehydes and ketones. Since then, the scope of this kind of reaction has been extended to asymmetric aldol addition to produce chiral products under the catalysis of chiral Lewis acids.⁵ Several years later, Sakurai reported the allylation reaction of allyltrimethylsilane with aldehydes and ketones promoted by Lewis acids (now known as the Sakurai-Hosomi reaction).⁶ Furthermore, the allylation and crotylation of carbonyl groups using silane reagents have been well developed into efficient strategies for preparing homoallylic alcohols and amines with stereochemical control;⁷ these are versatile synthetic intermediates for organic synthesis. In the past decades, the allylation and crotylation methods have been expanded to ketone derivatives to prepare tertiary carbinols, usually promoted by Lewis acids or bases, either catalytically or stoichiometrically.⁸

In 1992 Myers reported noncatalyzed aldol addition reactions of enolic silacyclobutanes with aldehydes, which were accelerated by a four-membered silacyclobutane ring.⁹ In

contrast, acyclic enol silanes showed either no evidence of reaction or very sluggish reaction rates even at higher temperature for extended time (Scheme 1-1). One year later, Denmark independently released similar results on the reactivity of silacyclobutanes with aldehydes, in which he proposed a '*direct silicon group transfer*' process confirmed by a crossover experiment.¹⁰



Scheme 1-1. Noncatalyzed aldol addition reactions of silacyclobutanes with aldehydes.^{9,10a}

In 1994 Utimoto reported noncatalyzed allylation reactions of allylic silacyclobutanes with aldehydes and α -hydroxyketones, which were also accelerated by the four-

membered silacyclobutane ring (Scheme 1-2).¹¹ Allylic silanes without this fourmembered ring could not react with carbonyl compounds in the absence of catalysts, although this reaction could be promoted by Lewis acids.⁷ Utimoto also showed that the cyclohexyloxy silacyclobutane was even more reactive toward aldehydes as the Lewis acidity of its silicon was enhanced by the oxygen atom. The evidence indicated that it could react with α -hydroxyketones smoothly but other allylic silacyclobutanes could not.



Scheme 1-2. Noncatalyzed allylation reactions of silacyclobutanes with aldehydes and ketones.¹¹

Such high reactivity of silacyclobutanes described above has been explained as *'strain-release Lewis acidity'* by Denmark.^{10b,12,13} As can be seen in Figure 1-2, acyclic silane has a normal C-Si-C bond angle of 109°. When it is attacked by a nucleophile, an

Part I

intermediate of penta-coordinate trigonal bipyramidal (*tbp*) geometry will be formed in the transition state, in which the C-Si-C bond angle is 90°. In contrast, silacyclobutane has a C-Si-C bond angle of approximately 79° as C-Si bond length (0.193 nm) is longer than that of C-C (0.159 nm),¹⁴ which indicates that there is a ring strain in this molecule. In the corresponding *tbp* transition state, its ring strain is released. Therefore its reactivity with nucleophiles is greatly increased. Thus the silicon atom shows enhanced Lewis acidity when it is constrained in a small ring compared with the acyclic silane.



Figure 1-2. Lewis acidity of silicon is enhanced by the strained silacyclobutane ring.¹³

Silacyclobutane is not the only cyclic silane that exhibits high reactivity because of *strain-release Lewis acidity*. In 1997 Kira described the noncatalyzed allylation reactions of allylic silacycles with aldehydes, promoted by a five-membered 1,3-dioxa-2-silacyclopentane ring derived from a tartrate ester (Scheme 1-3).¹⁵ Not long ago very similar reactions utilizing this ring were first reported by Wang, in which triethylamine or *N,N*-dimethylformamide was used as a promoter to form a '*chiral penta-coordinate transition state*'.¹⁶ Earlier, Shanmuganathan also described allylation reactions of allylic 1,3-dioxa-2-silacyclopentane with aldehydes in the presence of a Lewis acid.¹⁷ Although

Wang had assumed the activation from an external promoter enhanced the reactivity of the 1,3-dioxa-2-silacyclopentane ring, no external promoters were employed in Kira's results. Kira suggested that the allylation proceeded via a cyclic transition state in which one of the carbonyl groups in the tartrate coordinated with the silicon atom.



Scheme 1-3. Noncatalyzed allylation reactions of allylic silacycles with aldehyde reported by Kira.¹⁵

Xiaolun Wang in Prof. Leighton's group found that allylchlorosilane 1, with a strained 1,3-dioxa-2-silacyclopentane ring, reacted with benzaldehyde smoothly in toluene at room temperature to produce a homoallylic alcohol in the absence of catalyst (Scheme 1-4).¹⁸ In contrast, neither silane 2 with a six-membered ring nor acyclic silane 3 could react with benzaldehyde even at the elevated temperature of 50 °C. These experiments confirmed that the high reactivity of 1 resulted from the five-membered 1,3-dioxa-2-

silacyclopentane ring due to *strain-release Lewis acidity*. Although this five-membered ring has smaller ring strain than the four-membered silacyclobutane (O-Si-O bond angle is 96°),¹⁹ actually its Lewis acidity is further enhanced by its two electron-withdrawing oxygen atoms.



Scheme 1-4. Noncatalyzed allylation reaction of an allylic silane containing a strained 1,3-dioxa-2-silacyclopentane ring with benzaldehyde by Xiaolun Wang.¹⁸

Denmark classified allylation reactions of carbonyl compounds into three types on the basis of their mechanism.²⁰ As shown in Figure 1-3, *type I* allylation proceeds via a cyclic transition state with a six-membered chair-like conformation, in which the carbonyl group is activated by the internal metal via the six-membered ring (M = B, Al, Sn). By comparison, *type II* allylation in general involves an acyclic transition state, in which the carbonyl group is activated by an external Lewis acid (M = Si, Sn). In *type III* allylation,

the (Z)-crotyl group is isomerized to the more stable (E)- via Lewis acid catalysis before the addition proceeds by a cyclic *type I* mechanism (M = Ti, Cr, Zr). Therefore the product stereochemistry is independent of the geometry of the starting material. The stereochemistry of *type I* allylation products can be predicted as follows: (E)-crotyl produces an *anti*-methyl product and (Z)-crotyl produces a *syn*-methyl product. Therefore the *type I* allylation reaction will potentially tend to give higher diastereoselectivity to the addition products than *type II* due to its tight transition state. The noncatalyzed allylation reactions reported by Utimoto and Kira both belong to *type I*, in which the carbonyl group was activated by the silicon with enhanced Lewis acidity.



Figure 1-3. Classification of allylation reactions proposed by Denmark.²⁰

Based on the discussion above, it can be seen that the silicon atom is a promising carrier of functional groups for both aldol addition and allylation reactions due to its high reactivity under diverse conditions. It is possible that both functional groups can be combined in one molecule via a silicon atom and therefore the aldol and allylation reactions could proceed in a consecutive manner so that a target 1,3-diol can be obtained conveniently in one step. In this part of the thesis, tandem aldol-allylation reactions using aldehyde- and ketonederived allylenolsilanes with aldehyde will be described and a discussion of the reaction results will be presented.

References

- [1] a) Khosla, C.; Gokhale, R. S.; Jacobsen, J. R.; Cane, D. E. Annu. Rev. Biochem. 1999, 68, 219-253. b) Khosla, C. Chem. Rev. 1997, 97(7), 2577-2590.
- [2] a) Oishi, T.; Nakata, T. Synthesis 1990, (8), 635-645. b) Bode, S. E.; Wolberg, M.; Mueller, M. Synthesis 2006, (4), 557-588.
- [3] Zacuto, M. J.; O'Malley, S. J.; Leighton, J. L. *Tetrahedron* 2003, 59(45), 8889-8900. b)
 Zacuto, M. J.; O'Malley, S. J.; Leighton, J. L. J. Am. Chem. Soc. 2002, 124(27), 7890-7891.
 c) O'Malley, S. J.; Leighton, J. L. Angew. Chem. Int. Ed. 2001, 40(15), 2915-2917.
- [4] a) Mukaiyama, T.; Banno, K.; Narasaka, K. J. Am. Chem. Soc. 1973, 96(24), 7503-7509.
 b) Mukaiyama, T.; Banno, K.; Narasaka, K. Chem. Lett. 1973, (9), 1011-1014.
- [5] a) Narasaka, K. Synthesis 1991, (1), 1-11. b) Kobayashi, S.; Uchiro, H.; Fujishita, Y.; Shiina, I.; Mukaiyama, T. J. Am. Chem. Soc. 1991, 113(11), 4247-4252. c) Furuta, K.; Maruyama, T.; Yamamoto, H. J. Am. Chem. Soc. 1991, 113(3), 1041-1042. d) Parmee, E. R.; Tempkin, O.; Masamune, S.; Abiko, A. J. Am. Chem. Soc. 1991, 113(24), 9365-9366. e) Kiyooka, S.; Kaneko, Y.; Komura, M.; Matsuo, H.; Nakano, M. J. Org. Chem. 1991, 56(7), 2276-2278.
- [6] Hosomi, A.; Sakurai, H. Tetrahedron 1976, 17(16), 1295-1298.
- [7] Yamamoto, Y.; Asao, N. Chem. Rev. 1993, 93(6), 2207-2293.
- [8] For examples of allylation of ketones promoted by Lewis acids or bases, see: a) Yamasaki, S.;
 Fujii, K.; Wada, R.; Kanai, M.; Shibasaki, M. J. Am. Chem. Soc. 2002, 124(23), 6536-6537.
 b) Cossy, J.; Lutz, F.; Alauze, V.; Meyer, C. Synlett 2002, (1), 45-48. c) Ishihara, K.;
 Hiraiwa, Y.; Yamamoto, H. Synlett 2001, (12), 1851-1854. d) Kira, M.; Sato, K.; Sekimoto,
 K.; Gewald, R.; Sakurai, H. Chem. Lett. 1995, (4), 281-282. e) Coppi, L.; Mordini, A.;
 Taddei, M. Tetrahedron Lett. 1987, 28(9), 969-972.
- [9] Myers, A. G.; Kephart, S. E.; Chen, H. J. Am. Chem. Soc. 1992, 114(20), 7922-7923.
- [10] a) Denmark, S. E.; Griedel, B. D.; Coe, D. M. J. Org. Chem. 1993, 58(5), 988-990. b)
 Denmark, S. E.; Griedel, B. D.; Coe, D. M.; Schnute, M. E. J. Am. Chem. Soc. 1994, 116(16),
 7026-7043. c) Denmark, S. E.; Griedel, B. D. J. Org. Chem. 1994, 59(18), 5136-5138.

- [11] Matsumoto, K.; Oshima, K.; Utimoto, K. J. Org. Chem. 1994, 59(23), 7152-7155.
- [12] Denmark, S. E.; Jacobs, R. T.; Dai-Ho, G.; Wilson, S. Organometallics 1990, 9(12), 3015-3019.
- [13] Zhang, X.; Houk, K. N.; Leighton, J. L. Angew. Chem. Int. Ed. 2005, 44(6), 938-941.
- [14] Gordon, M. S.; Barton, T. J.; Nakano, H. J. Am. Chem. Soc. 1997, 119(49), 11966-11973.
- [15] Zhang, L. C.; Sakurai, H.; Kira, M. Chem. Lett. 1997, (2), 129-130.
- [16] a) Wang, Z.; Wang, D.; Sui, X. Chem. Commun. 1996, (19), 2261-2262. b) Wang, D.;
 Wang, Z. G.; Wang, M. W.; Chen, Y. J.; Liu, L.; Zhu, Y. Tetrahedron: Asymm. 1999, 10(2), 327-338.
- [17] Shanmuganathan, K.; French, L. G.; Jensen, B. L. Tetrahedron: Asymm. 1994, 5(5), 797-800.
- [18] Kinnaird, J. W. A.; Ng, P. Y.; Kubota, K.; Wang, X.; Leighton, J. L. J. Am. Chem. Soc. 2002, 124(27), 7920-7921.
- [19] Ohshita, J.; Sakurai, H.; Masaoka, S.; Tamai, M.; Kunai, A.; Ishikawa, M. J. Organometallic Chem. 2001, 633(1-2), 131-136.
- [20] Denmark, S. E.; Weber, E. J. Hel. Chim. Acta 1983, 66(6), 1655-1660.

Chapter 2. Tandem Aldol-Allylation Reactions Promoted by Strained 1, 3-Dioxa-2-silacyclopentane Ring

From the introduction in Chapter 1, it can be seen that the five-membered 1,3-dioxa-2-silacyclopentane ring effectively activates the reactivity of silane reagents by enhancing its Lewis acidity. One advantage of using this 1,3-dioxa-2-silacyclopentane is that it can be easily prepared from chloro silanes and 1,2-diols in the presence of base. With chiral 1,2-diols incorporated in the silane molecules, absolute stereocontrol is possible using this approach.

In 1999 Berrisford reported the tandem aldol-allylation reactions using an aldehydederived allylenolsilane with aldehyde dimethyl acetals promoted by Lewis acids (Scheme 2-1).¹ Although only moderate diastereoselectivities were obtained in his results, this tandem aldol-allylation strategy offered an alternative method to prepare homoallylic 1,3-diol fragment for organic synthesis in a single step.



Scheme 2-1. Tandem aldol-allylation reactions of allylenolsilanes with dimethylacetals promoted by Lewis acids.¹

2.1 Tandem Aldol-Allylation Reactions Using Aldehyde-Derived Allylenolsilanes

This study followed from the work of Xiaolun Wang in Prof. Leighton's group which is described briefly in the following two paragraphs.

Based on the high reactivity of 1 with aldehyde (Scheme 1-4), it was imagined that allylenolsilane 4 containing both an enol and an allyl group, when treated with an aldehyde, would generate an intermolecular aldol addition reaction with the aldehyde followed by an intramolecular allylation reaction on the intermediate aldehyde. As a result, an allylic 1,3-diol **5** would be generated as final product. Compared with using aldehyde-derived enolates directly in aldol reaction, polymerization can be avoided effectively by this tandem aldol-allylation strategy. This new tandem aldol-allylation approach will create 1,3-diol fragments with stereoselectivity conveniently and rapidly and with higher efficiency than the traditional stepwise transformations.

For this purpose, 4 was prepared via reaction of 1 with lithium enolate of acetaldehyde and then purified by distillation. The results of the reaction of 4 with aldehydes supported our assumption. As shown in Scheme 2-2, 1,3-diol 5 was obtained as a $11\sim12:1$ mixture of diastereomers in an overall yield of $59\sim60\%$.² At the same time allylic alcohol 6 was also isolated as a side product due to direct allylation of 4 with aldehyde. This indicates that the allyl group in 4 is less reactive with aldehyde than the enol group, although its reaction with aldehyde is still competitive. It is noteworthy that allylenolsilane 7 with a six-membered ring did not react with aldehyde under the same reaction conditions as above. This observation once again supported our conclusion on the contribution of 1,3-dioxa-2-silacyclopentane to the high reactivity of the silanes with this ring structure.



Scheme 2-2. Noncatalyzed tandem aldol-allylation reactions of allylenolsilane 4 containing a strained 1,3-dioxa-2-silacyclopentane ring with aldehydes by Xiaolun Wang.²

My effort in this part of the project was to expand the scope of substrates for these tandem aldol-allylation allylenolsilanes reactions and investigate the effects of substituents on the enol group. Allyl-(Z)-enolsilane 8 and allyl-(E)-enolsilane 9 were synthesized respectively using the same method as that of 4 (Scheme 2-3). For this purpose, propenyloxytrimethylsilane 10 was first prepared as a mixture of (E)- and

(Z)-isomers (1:2.6) from propionaldehyde according to a known procedure.⁴ Then the (E)- and (Z)-isomers were carefully enriched by repeated fractional distillation using Vigreux distilling columns, based on the difference of their boiling points (E: 100-103 °C; Z: 96-97 °C)^{4d}. 8 and 9 were then prepared from 1 with lithium enolate of (Z)- and (E)-10 respectively and then purified by distillation (70% and 71 % yields).²

8 reacted with cyclohexanecarboxaldehyde in toluene at 40 °C and gave four tandem products (65:18:12:5 dr) in 59% overall yield as well as 15% yield of 6 as a direct allylation product, among which the major tandem product 11 was isolated in 38% yield. On the contrary, reaction of 9 with aldehyde under the same conditions essentially gave two tandem products (2:1 dr) in 30% overall yield as well as 30% yield of 6. These results prove that the (Z)-enol group is more reactive with aldehydes than the (E)-. The stereochemistries of 11 and 12 were characterized by transforming them into their corresponding acetonides 13 and 14 using 2,2-dimethoxypropane and a catalytic amount of (+)-camphorsulfonic acid (CSA) in CH₂Cl₂ (Scheme 2-3). ¹³C NMR spectra of **13** and 14 revealed that their configurations were both syn according to Rychnovsky's method to determine the stereochemistries of 1,3-diol acetonides,⁵ and therefore both 11 and 12 were assigned as syn-1,3-diols. The configurations of the methyl groups in 13 and 14 were established using Karplus equations by observing the coupling constants of the two adjacent hydrogen atoms between H_A and H_B , H_B and H_C in ¹H NMR spectra. In 13, the coupling constants between H_A and H_B , H_B and H_C are both about 10 Hz, which indicates



that H_A and H_B are *anti*, as are H_B and H_C . Therefore the configuration of the methyl groups in 13 (and 11 in turn) were assigned as *anti* relative to the 1,3-diol. In 14, the

Scheme 2-3. Synthesis of allylenolsilanes 8 and 9 containing a strained 1,3-dioxa-2silacyclopentane ring, their noncatalyzed tandem aldol-allylation reactions with aldehyde, and stereochemical confirmation of the major tandem products 11 and 12.²
coupling constants between H_A and H_B , H_B and H_C are both about 2 Hz, which indicates that H_A , H_B and H_C are all *syn*. Therefore the configuration of methyl groups in 14 (and 12 in turn) were assigned as *syn* relative to the 1,3-diol.

Rychnovsky et al investigated over 220 1,3-diol acetonide compounds and established an empirical method to determine their stereochemistries using ¹³C NMR spectral data.⁵ As shown in Figure 2-1, syn-1,3-diol acetonides unquestionably prefer to exist in a stable chair-like conformation so that both the R_1 and R_2 groups are situated at equatorial positions along with one of the two acetal methyl groups. The other acetal methyl group has an axial position. The equatorial acetal methyl group usually appears at about 30 ppm in ¹³C NMR spectra while the axial one is at about 19 ppm. The chemical shift of the acetal carbon is usually to about 98.5 ppm. Conversely, anti-1,3-diol acetonides do not prefer to exist in either chair conformation because of the repulsive interaction between one of the acetal methyl groups and R_1 or R_2 unless they are small enough (such as hydrogen), but adopt a twist-boat conformation to avoid such repulsive interaction. As a result, both acetal methyl groups of the anti-1,3-diol acetonides appear at about 25 ppm in ¹³C NMR spectra while the chemical shift of acetal carbon is usually to about 100.5 ppm. Rychnovsky's investigation indicates that chemical shifts of acetal methyl groups are reliable enough to confirm the stereochemistries of 1,3-diol acetonides as long as R_1 and R₂ are big enough so that the anti-1,3-diol acetonides adopt the twist-boat conformation as the most stable one. On the other hand, chemical shifts of acetal carbon

are less reliable because of significant data overlap between the syn- and anti-1,3-diol acetonides.



Figure 2-1. Rychnovsky's empirical method using ¹³C NMR spectra to assign the stereochemistries of *syn*- and *anti*-1,3-diol acetonides due to their different stable conformations.⁵

Other substituted allylenolsilanes 15-19 were also prepared by Xiaolun Wang and Andrew J. Nation to investigate the substituent effects on both enol and allyl groups.² (Z)-crotylenolsilane 15 and (E)-crotylenolsilane 16 were synthesized via reaction of (Z)-and (E)-crotylchlorosilane with lithium enolate of acetaldehyde respectively using the same method as that of 4. As can seen in Scheme 2-4, 15 reacted with cyclohexane-carboxaldehyde in toluene at 40 °C for 60 h and essentially gave two tandem products (10:1 dr), among which the major one 20 can be isolated in 71% yield. Reaction of 16

with aldehyde under the same conditions also gave two tandem products (8:1 dr), with the major one 21 isolated in 60% yield. In both cases, syn-1,3-diols were obtained as major tandem products and direct crotylation products were isolated in less than 5% yield, which indicates less competitive reactivity of the crotyl group with aldehyde compared with the enol group. The (Z)-crotyl 15 produced an anti-methyl product 20 while the (E)-crotyl 16 produced a syn-methyl product 21, which suggests that the intramolecular addition proceeds via a cyclic type I⁶ transition state. Crotylenolsilanes 17-19 were also prepared from crotylchlorosilanes with lithium enolates of propionaldehyde using the same method as that of 8 and 9 to further extend the scope of this study. All three of these crotylenolsilanes reacted with aldehyde to give 1,3-diols with up to four stereocenters. (Z)-Crotyl-(Z)-enolsilane 17 reacted with cyclohexanecarboxaldehyde in toluene at 40 $^{\circ}$ C for 80 h and gave four tandem products (66:23:8:3 dr), among which the major one 22 can be isolated in 52% yield. Reaction of (E)-crotyl-(Z)-enolsilane 18 with aldehyde under the same conditions also gave four tandem products (65:24:7:4 dr), with the major one 23 isolated in 51% yield. Although reaction of allyl-(E)-enolsilane 9 with aldehyde gave very poor diastereoselectivity in producing 1,3-diol products, (E)-crotyl-(E)enolsilane 19 reacted with cyclohexanecarboxaldehyde in benzene at 40 °C for 132 h and gave four tandem products with much better diastereoselectivity (86:11:3:1 dr), among which anti-1,3-diol 24 could be isolated in 60% yield. This is one of the FEW cases that an *anti*-1,3-diol was obtained as the major product in our tandem aldol-allylation reaction studies. The direct crotylation product was also obtained in 26% yield in this reaction.



Scheme 2-4. Xiaolun Wang and Andrew J. Nation's work on noncatalyzed tandem aldol-allylation reaction of substituted allylenolsilanes **15-19** with cyclohexanecarboxaldehyde.²

2.2 Tandem Aldol-Allylation Reactions Using Ketone-Derived Allylenolsilanes

Based on our encouraging and appealing results with tandem aldol-allylation reactions using aldehyde-derived allylenolsilanes to conveniently produce allylic 1,3-diol units with 2 to 4 stereocenters, extension of this chemistry to the synthesis of tertiary carbinols using ketone-derived allylenolsilanes attracted our attention since tertiary carbinols are also very important components in natural products. One of the examples is oleandomycin (Figure 1-1).

Based on our work with aldehyde-derived allylenolsilanes, it was natural to design a ketone-derived allylenolsilane (Figure 2-2). This silane molecule, when treated with an

aldehyde, would generate an intermolecular aldol addition reaction with the aldehyde followed by an intramolecular allylation reaction on the intermediate ketone. Therefore an allylic 1,3-diol would be obtained as tandem aldol-allylation product with diastereoselectivity.



Figure 2-2. Design of a ketone-derived allylenolsilane and the suggested tandem aldol-allylation reaction with aldehyde.

This part of the research began with the study of acetone-derived allylenolsilane 25, which was prepared via reaction of 1 with lithium enolate of $acetone^{4,7}$ and then purified by distillation (61% yield) (Scheme 2-5).⁸ 25 reacted with cyclohexanecarboxaldehyde in toluene at 60 °C for 48 h and gave 1,3-diol 27 as an 8:1 mixture of diastereomers in 82% overall yield as well as only 1.5% yield of 6 as a direct allylation product. It is noteworthy that β -hydroxyketone 28 was also obtained in 2.3% yield due to incomplete tandem reaction. Compared with acetaldehyde-derived allylenolsilane 4, direct allylation of 25 with aldehyde is not competitive at all due to the higher reactivity of the

acetone-derived enol group in 25 than the acetaldehyde-derived group in 4. Formation of 28 indicates that the aldol reaction of 25 with aldehyde is faster than the intramolecular allylation. To prove this, 25 was treated with cyclohexanecarboxaldehyde in toluene at 40 °C for 20 h only and 28 was produced in 56% yield as major product while 27 was obtained in only 27% yield.



Scheme 2-5. Synthesis of acetone-derived allylenolsilane 25, its noncatalyzed tandem aldol-allylation reaction with aldehyde, and stereochemical confirmation of tandem products 27.⁸

To confirm the stereochemistry of 27, both diastereomers were transformed into the corresponding acetonides 29 and 30 using the same method as in Scheme 2-3. ¹³C NMR data of 29 and 30 revealed that their stereochemistries could not be determined by Rychnovsky's method, as neither of them prefers to exist in a stable chair-like conformation due to the repulsive interaction of the two methyl groups. Instead, both of them adopt a twist-boat conformation to avoid such repulsive interactions. Therefore the stereochemistries of 29 and 30 could not be confirmed on the basis of their different stable conformations. An alternative method was employed to confirm the stereochemistry of 27, as shown in Scheme 2-5. Both diastereomers of 27 were treated

with mercury (II) acetate chloride in CH_2Cl_2 and the major one gave cyclic tetrahydropyran **31** in 88% yield while the minor one gave two diastereomeric products (1: 1) in an overall yield of 81%. Selective 1D NOESY enhancement of diastereomeric **31** and **32** revealed their relative configurations. The major tandem product of **27** was assigned as *syn* while the minor one was *anti*.

The tandem aldol-allylation reaction of acetone-derived allylenolsilane 25 with aldehyde was so efficient at creating a 1,3-diol with good diastereoselectivity that it urged us to expand the scope of this study to other allylenolsilanes with methyl-substituted allyl groups. Four more acetone-derived silanes, (Z)-crotylenolsilane 33, (E)-crotylenolsilane 34, methallylenolsilane 35 and angelylenolsilane 36, were synthesized using the same procedure as that of 25 from the corresponding methyl-substituted allyltrichlorosilanes 37,

39, 41 and **43** (Scheme 2-6).⁸ **37** and **43** were prepared from 1,3-butadiene and isoprene respectively in the presence of palladium catalysts following known methods.⁹ **39** and **41** were prepared from (*E*)-crotyl chloride¹⁰ and methallyl chloride respectively under the catalysis of copper (I) chloride using the same procedure as that for allyltrichlorosilane.⁹



Scheme 2-6. Synthesis of acetone-derived allylenolsilanes 33-36 with methyl-substituted allyl groups.⁸

The methyl-substituted allylenolsilanes **33-36** were treated with cyclohexanecarboxaldehyde to test the properties of these tandem aldol-allylation reactions (Scheme 2-7).⁸ (Z)-Crotylenolsilane 33 reacted with aldehyde in toluene at 110 °C for 50 h and gave 1,3-diol 45 as a 14:1 mixture of diastereomers in 58% overall yield as well as 4% yield of 28 due to incomplete tandem reaction. By comparison, reaction of (E)-crotylenolsilane 34 with aldehyde in benzene at 80 °C for 60 h gave 1,3-diol 46 as a 6:1 mixture of diastereomers in 88% overall yield as well as only 1.4% yield of 28. Here benzene was employed as solvent because it gave slightly better reaction yields than toluene. In both cases, no direct crotylation products were found. Just as with the aldehyde-derived allylenolsilanes 15 and 17, the (Z)-crotyl 35 produced an anti-methyl product 45 while the (E)-crotyl 38 produced a syn-methyl product 46. These results once again indicate that the intramolecular addition proceeds via a cyclic type I^{6} transition state. It can be seen here that the (E)-crotyl group is more reactive than the (Z)-crotyl group in the intramolecular addition on the generated intermediate ketone, but less reactive than the allyl group in 26. Methallylenolsilane 41 reacted with aldehyde in toluene at 40 °C for 30 h and gave 1,3-diol 47 as a >50:1 mixture of diastereomers in 54% overall yield as well as 4% yield of 28. This is the fastest tandem aldol-allylation reaction of allylenolsilanes with aldehyde we have ever observed, due to the electron-donating effect of the methyl substituent on the allyl group. This effect also promoted the formation of a direct methallylation product in 19% yield. It produced such a high diastereoselectivity on tandem products that the minor anti-product could barely be observed! This beneficial effect was further used to improve the lower reactivity of the (Z)-crotyl group. With this methyl substituent attached on the (Z)-crotyl group, angelylenolsilane 44 reacted with

aldehyde in toluene at 70 °C for 90 h and gave 1,3-diol **48** as a >30:1 mixture of diastereomers in 74% overall yield as well as 6% yield of **28** and 5% yield of direct angelylation product. Compared with (Z)-crotyl silane **35**, angelyl silane **44** gave both higher diastereoselectivity and chemical yield on tandem products. In summary, the reactivity order of methyl-substituted allyl groups in acetone-derived allylenolsilanes is: methallyl> allyl> angelyl> (E)-crotyl> (Z)-crotyl while the order of diastereoselectivity is: methallyl> angelyl> (Z)-crotyl> allyl> (E)-crotyl.



Scheme 2-7. Noncatalyzed tandem aldol-allylation reactions of allylenolsilanes 33-36 with cyclohexanecarboxaldehyde.⁸

The stereochemistries of 1,3-diols 45-48 were determined using the same method as that

of 27. As shown in Scheme 2-8, 45-48 were treated with mercury (II) acetate chloride in CH_2Cl_2 to generate cyclic tetrahydropyran derivatives. 45 and 46 produced 49 and 50 respectively as single products. Stereochemistries of 49 and 50 were confirmed by selective 1D NOESY enhancement and the relative configurations of the methyl groups were established using Karplus equations by observing the coupling constants of the adjacent hydrogen atoms between H_A and H_B in ¹H NMR spectra. 45 was assigned as a *syn*-1,3-diol with an *anti*-methyl group while 46 was a *syn*-1,3-diol with a *syn*-methyl group. 47 gave two diastereomeric products (1: 4.6) with an overall yield of 90%. By selective 1D NOESY enhancement of the minor product 51, 47 was also assigned as a *syn*-1,3-diol. 48 produced two diastereomeric products 52 and 53 (1: 3.0) with an overall yield of 43% as well as 51% yield of 28 via deangelylation by mercury (II). Selective 1D



Scheme 2-8. Stereochemical confirmation of the major tandem products 45-48.⁸

NOESY enhancement of 52 and 53 both confirmed 48 as a syn-1,3-diol with an *anti*-methyl group.

The high reactivity of the acetone-derived enol groups in the silanes described above encouraged us to expand the scope of this study to other ketone-derived allylenolsilanes. Cyclohexanone-derived allylenolsilane **54** was prepared using the same method as that of **25** via reaction of **1** with the lithium enolate of cyclohexanone^{7b,11} and then purified by distillation (56% yield) (Scheme 2-9). Cyclopentanone-derived **55** was also synthesized from **1** and the lithium enolate of cyclopentanone¹² (56% yield).

54 reacted with cyclohexanecarboxaldehyde in toluene at 80 °C for 48 h and gave four tandem products (62:21:8:5 dr) in 62% overall yield as well as 10% yield of 6 as a direct allylation product. It is noteworthy that no β -hydroxyketone due to incomplete tandem reaction was found, which indicates that the aldol addition of 54 with aldehyde is slower than the intramolecular allylation. This result proves that the cyclohexanone-derived enol group is less reactive than the acetone-derived one. On the contrary, reaction of 55 with cyclohexanecarboxaldehyde in toluene gave different results. At 60 °C for 20 h only one tandem product 59b was obtained in 31% yield as well as 56% yield of β -hydroxyketone 60 due to incomplete tandem reaction and 6% yield of direct allylation product 6. To complete the conversion of the ketone intermediate into tandem products, this reaction was heated at 60 °C for 15 h followed by 110°C for 140 h and finally gave four tandem Part I

products (55:33:10:2 dr) in 58% overall yield as well as 13% yield of 60 and 6% yield of 6. 59a was isolated as the major tandem product in 32% yield and 59b was the second highest in 19% yield. Here it can be seen that the cyclopentanone-derived enol group is more reactive than the cyclohexanone-derived one so that its aldol addition with aldehyde is very fast even at 60 °C.



Scheme 2-9. Synthesis of allylenolsilanes 54 and 55, their noncatalyzed tandem aldol-allylation reactions with aldehyde, and stereochemical confirmation of tandem products 58 and 59.

The stereochemistries of **58**, **59a** and **59b** were characterized by their transformation into the corresponding acetonides **61**, **62a** and **62b** using the same method as in Scheme 2-3. Their ¹³C NMR spectral data revealed that Rychnovsky's empirical method could not be used here to determine their stereochemistries, as they are not simple mono-cyclic acetonides. Instead, selective 1D NOESY enhancement revealed their relative configurations, as shown in Scheme 2-9. **58** and **59a** were confirmed to have the same stereochemistry while **59b** was different.



Scheme 2-10. Reaction pathways of 55 with aldehyde to form 59a, 59b and 60.

The unusual tandem aldol-allylation reaction results of **55** with aldehyde can be explained as shown in Scheme 2-10. **55** reacted with aldehyde very fast at 60 °C to generate both *syn-* and *anti-* intermediate ketones (\approx 3:2). Intramolecular allylation of the *anti-*ketone proceeded smoothly at 60 °C to form the 1,3-diol **59b** as major tandem product with high diastereoselectivity. Conversely, intramolecular allylation of the *syn*-ketone could not proceed at 60 °C and only **60** was obtained (stereochemistry of **60** was assigned as *syn* by comparing its ¹H and ¹³C NMR spectra with reported data¹³). At the elevated temperature of 110°C for longer time the intramolecular allylation of the *syn*-ketone eventually gave **59a** as the major tandem product.

2.3 Discussions on Tandem Aldol-Allylation Reactions of Allylenolsilanes

Based on the tandem aldol-allylation reaction results of allylenolsilanes 26, 33, 34 and 36 with aldehyde to form 1,3-diols, a transition state model was proposed to explain their stereoselectivity and reactivity (Figure 2-3). The essential process involves an intermolecular aldol addition followed by an intramolecular allylation reaction. Both steps proceed via a six-membered cyclic transition state, in which the silicon forms a penta-coordinate complex to release the strain in the 1,3-dioxa-2-silacyclopentane ring. As a result, the reactivity of both enol and allyl groups in allylenolsilanes is enhanced by *strain-release Lewis acidity* and therefore the tandem aldol-allylation reaction proceeded smoothly in the absence of catalysis. The latter process is a *type I* allylation⁶ producing a chair-like conformation prior to allyl transfer, a well-accepted mechanism described by many chemists.^{8,9a-b,14} This mechanism also explains the production by (*Z*)-crotyl **35** of an *anti-*methyl product **45** while (*E*)-crotyl **38** produced a *syn-* methyl product **46**. The

diastereoselectivity of the inflexible intramolecular allylation is potentially higher than the flexible intermolecular aldol addition.



Figure 2-3. Proposed tandem aldol-allylation transition state of allylenolsilanes 25, 33-36 with aldehyde to produce *syn*-1,3-diols as major products.

A transition state model to explain the reaction stereoselectivity of allylenolsilane 54 with aldehyde to generate tandem product 58 is also given in Figure 2-4, A (four methyl groups of pinacol are omitted for simplicity). In this model, the *type I* intramolecular allylation transition state explains why 58 is formed as major tandem product. It is reasonable to propose a similar transition state for the reaction of allylenolsilane 55 with aldehyde to produce 59a while the formation of 59b definitely passes through a different

type I transition state (Figure 2-4, B). Therefore, **59a** and **59b** were generated as preferred tandem products via different intermediate ketones resulting in their different stereochemistries.



Figure 2-4. Proposed tandem aldol-allylation transition state of allylenolsilanes 54 and 55 with aldehyde to produce 1,3-diols 58 and 59 as preferred products.

Our tandem aldol-allylation reactions of ketone-derived allylenolsilanes revealed the first

evidence of noncatalyzed intramolecular allylation (crotylation) on intermediate ketones (*type I*), as such reactions have not been previously reported. K. Utimoto described noncatalyzed intermolecular allylation of α -hydroxyketones (see Chaper 1)^{14a}, which is a rare example of noncatalyzed *type I* ketone allylation. On the other hand, intermolecular *type I* allylation of ketones promoted by Lewis acids or bases has been frequently reported.^{9b,14b-f,15} In general their substrates are limited to adjacently functionalized ketones such as α - or β -hydroxyketones, enolizable α -diketones, or α -oxocarboxylic acids abetting their formation of a cyclic *type I* transition state with allylsilane reagents. By comparison, intermolecular *type II* allylation of ketones are regular aliphatic and aromatic ketones.¹⁶

The tandem aldol-allylation reactions of allylenolsilanes rely heavily on the identity of 1,2-diols that constitute the five-membered 1,3-dioxa-2-silacyclopentane ring. As distillation is the only efficient method to purify the allylenolsilanes according to our experiments, the possible 1,2-diol candidates were limited to those with moderate molecular weights so that distillation under vacuum is practical. Further work by my colleagues and me on optimizing the 1,2-diol candidates for allylenolsilanes excluded secondary 1,2-diols. Synthesis of allylchlorosilanes from secondary 1,2-diols and allyltrichloro- silane under different conditions proved to be too difficult. For example, my attempted preparation of allylchlorosilanes **63-67** failed to produce any desired

products probably due to their high boiling points or the specific properties of the 1,2-diols (Scheme 2-11). Among them the synthesis of **63** was reported before,¹⁷ but it was used directly for the next reaction steps without further purification. An acetaldehyde-derived allylenolsilane **68** was synthesized using the same method as that of **4** from 1,2-diol **69** and **69** was prepared from cyclopentanone via a pinacol-coupling reaction¹⁸. **68** was then treated with benaldehyde and cyclohexanecarboxaldehyde to test their tandem aldol-allylation reaction properties. Both reactions were found to be similar to that of pinacol-derived allylenolsilane **4**. However, the large-scale requirement for the 1,2-diol to prepare the allylenolsilane excluded **69** as an ideal choice.



Scheme 2-11. Attempted synthesis of 63-68 and noncatalyzed tandem aldol-allylation reactions of68 with aldehydes.

Xiaolun Wang *et al* also synthesized allylenolsilanes derived from N-alkyl-2aminoalcohols and N,N-dialkyl-1,2-diamines and tested their tandem aldol-allylation reaction properties. Unfortunately, when treated with aldehydes, all of these silanes produced little or no 1,3-diols as tandem products. This might result from the basicity of the nitrogen atom in the allylenolsilanes. To prove this, an acetone-derived allylenolsilane 71 with a less basic nitrogen was prepared as a mixture of 3:1 diastereomers (Scheme 2-12).⁸ 71 was treated with benaldehyde and gave normal tandem aldol-allylation product, which is very similar to the situation with pinacol-derived allylenolsilane 25 with regard to both diastereoselectivity and chemical yield. However, the major product was found to be an *anti*-1,3-diol.



Scheme 2-12. Synthesis of an acetone-derived allylenolsilane 71 and its noncatalyzed tandem aldol-allylation reaction with benaldehyde by Xiaolun Wang *et al.*⁸

After our efforts on utilizing a broad range of 1,2-diols, pinacol was finally chosen as the ideal 1,2-diol to construct the strained 1,3-dioxa-2-silacyclopentane ring to initiate the tandem aldol-allylation reactions due to its commercial availability and relatively smaller molecular weight.

2.4 Summary and Outlook

Our studies on the tandem aldol-allylation reactions have proven that the strained allylenolsilanes can react with aldehydes smoothly in the absence of catalysts to afford homoallylic 1,3-diols with moderate to high diastereoselectivity. Pinacol was employed to constitute the strained 1,3-dioxa-2-silacyclopentane ring as an effective platform to activate the reactivity of both enol and allyl groups due to the effect of *strain-release Lewis acidity*. As a result, no further activation from an external promoter is needed and thus the reaction conditions are greatly simplified. The allylenolsilane reagents can be prepared easily in several steps with good yields and are stable enough to be stored for months. The tandem aldol-allylation reactions proceed via a cyclic transition state with a six-membered chair-like conformation (*type I* allylation) and high diastereoselectivity in the tandem 1,3-diol products is generated.

The tandem aldol-allylation reactions are established on the basis that enol group is more reactive with aldehyde than allyl group so that aldol addition is initiated followed by the intramolecular allylation reaction. Homoallylic 1,3-diols with up to four stereocenters can be generated using our tandem aldol-allylation reactions in a single step conveniently and rapidly, which has become an important methodology to construct 1,3-diol fragments for the total synthesis of natural products. With chiral 1,2-diols incorporated in the silane molecules, absolute stereocontrol will hopefully be achieved using this approach.

2.5 Experimental Section

General Information. All reactions were carried out under an atmosphere of nitrogen in flame- or oven-dried glassware with magnetic stirring unless otherwise indicated. Degassed solvents were purified by passage through an activated alumina column. ¹H NMR spectra were recorded on a Bruker DPX-400 (400 MHz) or a Bruker DPX-300 (300 MHz) spectrometer and are reported in ppm from CDCl₃ or C₆D₆ internal standards (7.26 or 7.15 ppm respectively). Data are reported as follows: (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublets; coupling constant(s) in Hz; integration; assignment). Proton decoupled ¹³C NMR spectra were recorded on a Bruker DPX-400 (100 MHz) or a Bruker DPX-300 (75 MHz) and are reported in ppm from CDCl₃ or C₆D₆ internal standards (77.0 or 128.0 ppm respectively). Infrared spectra were recorded on a Perkin Elmer Paragon 1000 FT-IR spectrometer.

Propenyloxytrimethylsilane (10).^{4d} To a stirred solution of triethylamine (230 mL; 1.65 mol) and propionaldehyde (50 mL; 0.75 mol) in anhydrous DMF (280 mL), chlorotrimethylsilane (104 mL; 0.82 mol) was added slowly at 0 °C and then the reaction mixture was heated at reflux for 4 h. Pentane (500 mL) was added and the white solid formed was removed by filtration. The pentane layer was separated from DMF and washed with cold saturated aqueous NaHCO₃ twice, HCl (1 M) once and then NaHCO₃ once, and finally dried over anhydrous MgSO₄. After most solvent was removed by

distillation, the crude product was distilled to give **10** as a colorless liquid (b.p. 95-103 °C; E/Z= 1:2.6; 73.6 g; 75% yield). Then the (*E*)- and (*Z*)-isomers were carefully enriched by repeated fractional distillation using Vigreux distilling columns: (*Z*)-10. b.p. 96-97 °C; E/Z< 1:50; ¹H NMR (400 MHz, CDCl₃) δ 6.16 (dq, 1H, *J* = 1.7, 5.9 Hz, CH₃CH=CHO), 4.55 (m, 1H, CH₃CH=CHO), 1.57 (dd, 3H, *J* = 1.7, 6.7 Hz, CH₃CH), 0.18 (s, 9H, Si(CH₃)₃). (*E*)-10. b.p. 100-103 °C; E/Z> 12:1; ¹H NMR (300 MHz, CDCl₃) δ 6.19 (dq, 1H, *J* = 1.6, 12.0 Hz, CH₃CH=CHO), 4.99 (dq, 1H, *J* = 6.8, 12.0 Hz, CH₃CH=CHO), 1.52 (dd, 3H, *J* = 1.7, 6.8 Hz, CH₃CH), 0.17 (s, 9H, Si(CH₃)₃).

(Isopropenyloxy)trimethylsilane (26).⁷ To a stirred solution of anhydrous acetone (20.0 mL; 0.272 mol) and triethylamine (34.4 mL; 0.247 mol), chlorotrimethylsilane (31.4 mL; 0.247 mol) was added slowly and white solid appeared. A solution of NaI (37.36 g; 0.249 mol) in anhydrous acetonitrile (230 mL) was added slowly and more solid was formed during this process. The reaction mixture was stirred at room temperature for 2 h. Pentane (200 mL) was added and the solid was removed by filtration. The acetonitrile layer was separated from pentane and extracted with pentane (100mL×2). The combined pentane layers were washed with cold saturated aqueous NaHCO₃, NH₄Cl and then NaHCO₃, and finally dried over anhydrous K₂CO₃. After most solvent was removed by distillation, the crude product was distilled to give **26** as a colorless liquid: b.p. 88-89 °C; 19.0 g; 59% yield; ¹H NMR (300 MHz, C₆D₆) δ 4.17 (s, 1H, one of CH₂=C), 4.09 (s, 1H, one of CH₂=C), 1.72 (s, 3H, CH₂=CCH₃), 0.15 (s, 9H, Si(CH₃)₃).

(Cyclohexenyloxy)trimethylsilane (56).¹¹ The procedure is the same as that of 25 to yield 56 as a colorless liquid: b.p. 74-77 °C/ 20 mmHg; 74% yield; ¹H NMR (300 MHz, C_6D_6) δ 4.95 (m, 1H, CH=CO), 1.93-2.08 (m, 4H, CH₂CH₂CH₂CH₂CH₂), 1.35-1.57 (m, 4H, CH₂CH₂CH₂CH₂), 0.17 (s, 9H, Si(CH₃)₃).

(Cyclopentenyloxy)trimethylsilane (57).¹² The procedure is the same as that of 10 to yield 57 as a colorless liquid: b.p. 47-48 °C/ 14 mmHg; 70% yield; ¹H NMR (400 MHz, C_6D_6) δ 4.67 (m, 1H, CH=CO), 2.22-2.35 (m, 4H, CH₂CH₂CH₂), 1.69-1.78 (m, 2H, CH₂CH₂CH₂), 0.16 (s, 9H, Si(CH₃)₃).

General synthetic procedure for allyltrichlorosilane.⁷ To a stirred suspension of triethylamine (55.6 mL; 0.400 mol) and CuCl (2.02 g; 20.4 mmol) in anhydrous ether (250 mL), a mixture of allyl chloride (0.400 mol) and trichlorosilane (44.4 mL; 0.440 mol) was added slowly at 0 °C. The reaction mixture was stirred at room temperature for 10 h. The solid formed was removed by filtration and the solution was concentrated by distillation. The residue was further distilled to give the final product as a colorless liquid:

Allyltrichlorosilane. b.p. 117-118 °C; 82% yield; ¹H NMR (300 MHz, C₆D₆) δ 5.69-5.85 (m, 1H, CH=CH₂), 5.16-5.24 (m, 2H, CH=CH₂), 2.35 (dt, 2H, J = 1.1, 7.7 Hz, CH₂Si).

(*E*)-Crotyltrichlorosilane (39).^{9a,9b} b.p. 144-146 °C; 55% yield; E/Z = 28:1; ¹H NMR (400 MHz, C₆D₆) δ 5.16-5.28 (m, 1H, one of CH=CH), 5.05-5.16 (m, 1H, one of CH=CH), 1.72 (d, 2H, J = 7.4 Hz, CH₂Si), 1.41 (dd, 3H, J = 1.2, 6.3 Hz, CH₃CH). (*E*)-Crotyl chloride (E/Z = 26:1) was prepared from commercial (*E*)-crotyl alcohol following the same procedure as that in *Ref. 9*.

Methallyltrichlorosilane (41).^{9f} b.p. 137-140 °C; 58% yield; ¹H NMR (300 MHz, C₆D₆) δ 4.74 (s, 1H, one of CH₂=C), 4.63 (s, 1H, one of CH₂=C), 1.84 (s, 2H, CH₂Si), 1.55 (s, 3H, CH₂=CCH₃).

(Z)-Crotyltrichlorosilane (37).^{9a,9b} To a 350 mL pressure vessel was added tetrakis(triphenylphosphine)palladium(0) (0.91 g; 0.79 mmol). Cooled to -78 °C, 1,3-butadiene (34.4 mL; 0.40 mol) and trichlorosilane (41.7 mL; 0.41 mol) were added successively. After stirring at -78 °C for 15 h, the reaction mixture was warmed to room temperature and distilled directly to afford **37** as a colorless liquid: b.p. 146-147 °C; 72.0 g; 95% yield; E/Z < 1:100; ¹H NMR (300 MHz, C₆D₆) δ 5.36-5.49 (m, 1H, one of CH=CH), 5.10-5.22 (m, 1H, one of CH=CH), 1.85 (dd, 2H, J = 0.6, 8.2 Hz, CH₂Si), 1.31 (dd, 3H, J = 0.8, 6.9 Hz, CH₃CH).

Angelyltrichlorolsilane (43).^{9e,9f} To a 350 mL pressure vessel were added

bis(benzonitrile)palladium (II) chloride (83 mg; 0.22 mmol), triphenylphosphine (131 mg; 0.50 mmol), isoprene (10.7 mL; 0.11 mol) and trichlorosilane (10.1 mL; 0.10 mol) successively. After stirring at 70 °C for 8 h, the reaction mixture was distilled under vacuum directly to afford **43** as a colorless liquid: b.p. 57-58 °C/ 17.5 mmHg; 19.4 g; 95% yield; E/Z < 1:41; ¹H NMR (400 MHz, C₆D₆) δ 5.18 (q, 1H, J = 6.1 Hz, CH₃CH=C), 1.95 (s, 2H, CH₂Si), 1.60 (m, 3H, CH=CCH₃), 1.61 (dm, 3H, J = 6.1 Hz, CH₃CH=C).

General synthetic procedure for allylchlorosilanes.^{2,8} To a stirred solution of allyltrichlorosilane (0.106 mol) in anhydrous CH_2Cl_2 (200 mL), 1,8-diazabicyclo-[5.4.0]undec-7-ene (DBU) (39.6 mL; 0.265 mol) was added at 0 °C. Then a solution of pinacol (12.53 g; 0.106 mol) in anhydrous CH_2Cl_2 (100 mL) was added slowly. The solution was stirred at room temperature for 14 h. Most CH_2Cl_2 was removed by distillation. Anhydrous ether (250 mL) was added to the residue and stirred at room temperature for 2 h until a white solid formed. The solid was removed by filtration and the solution was concentrated by distillation. The residue was further distilled under vacuum to give the final product as a colorless oil:

Allylchlorosilane (1).^{1,2} b.p. 63-65 °C/ 5mmHg; 61% yield; ¹H NMR (400 MHz, C₆D₆) δ 5.64-5.77 (m, 1H, CH=CH₂), 4.90-5.00 (m, 2H, CH=CH₂), 1.80 (dt, 2H, J = 1.2, 7.7 Hz, CH₂Si), 1.17 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.03 (s, 6H, six of (CH₃)₂CC(CH₃)₂). (Z)-Crotylchlorosilane (38).⁸ b.p. 65-66 °C/ 1.8 mmHg; 49% yield; E/Z < 1:100; ¹H NMR (300 MHz, C₆D₆) δ 5.35-5.57 (m, 2H, CH₃CH=CH), 1.84 (dm, 2H, J = 6.1 Hz, CH₂Si), 1.48 (dm, 3H, J = 4.1 Hz, CH₃CH=CH), 1.16 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.01 (s, 6H, six of (CH₃)₂CC(CH₃)₂).

(*E*)-Crotylchlorosilane (40).⁸ b.p. 47-48 °C/ 0.6 mmHg; 36% yield; E/Z > 18:1; ¹H NMR (400 MHz, C₆D₆) δ 5.29-5.42 (m, 2H, CH₃CH=CH), 1.78 (dm, 2H, J = 6.9 Hz, CH₂Si), 1.48 (dm, 3H, J = 4.7 Hz, CH₃CH=CH), 1.17 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.03 (s, 6H, six of (CH₃)₂CC(CH₃)₂).

Methallylchlorosilane (42).⁸ b.p. 48-50 °C/ 0.9 mmHg; 46% yield; ¹H NMR (300 MHz, C₆D₆) δ 4.77 (m, 2H, CH₂=C), 1.86 (s, 2H, CH₂Si), 1.74 (m, 3H, CH₂=CCH₃), 1.18 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.04 (s, 6H, six of (CH₃)₂CC(CH₃)₂).

Angelylchlorolsilane (44).⁸ b.p. 53-55 °C/ 0.5 mmHg; 59% yield; E/Z < 1:100; ¹H NMR (400 MHz, C₆D₆) δ 5.20 (q, 1H, J = 6.7 Hz, CH₃CH=C), 1.88 (s, 2H, CH₂Si), 1.75 (m, 3H, CH=CCH₃), 1.53 (dm, 3H, J = 6.7 Hz, CH₃CH=C), 1.17 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.03 (s, 6H, six of (CH₃)₂CC(CH₃)₂).

General synthetic procedure for allylenolsilanes.^{2,8} To a stirred solution of CH₃Li in ether (1.4 M; 22.9 mL; 32.1 mmol), a solution of vinyl trimethylsilyl ether (31.5 mmol)

in anhydrous ether (7.0 mL) was added slowly at 0 °C. After 30 min, the solution was warmed to room temperature and stirred for 1 hour. Cooled to -78 °C, allylchlorosilane (31.5 mmol) was added and the solution was warmed to room temperature and stirred for 6 hours. Pentane (30 mL) was added and a white solid appeared. The solid was removed by filtration and most of the solvent was removed under vacuum. The residue was further

distilled under vacuum to give the final product as a colorless oil:

Allyl-(*Z*)-enolsilane (8).² b.p. 49-51 °C/ 0.6 mmHg; 70% yield; E/Z = 1:31; ¹H NMR (400 MHz, C₆D₆) δ 6.37-6.50 (dq, 1H, J = 1.8, 5.9 Hz, CH=CHO), 5.82-5.92 (m, 1H, CH₂=CH), 5.01-5.07 (m, 1H, one of CH₂=CH), 4.94-4.97 (m, 1H, one of CH₂=CH), 4.57 (dq, 1H, J = 6.0, 6.7 Hz, CH=CHO), 1.76 (dt, 2H, J = 1.2, 7.8 Hz, CH₂Si), 1.66 (dd, 3H, J = 1.8, 6.8 Hz, CH₃CH), 1.16 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.09 (s, 6H, six of (CH₃)₂CC(CH₃)₂); ¹³C NMR (100 MHz, C₆D₆) δ 137.9, 131.6, 115.9, 107.1, 81.6, 25.8, 25.7, 19.2, 9.3; ²⁹Si NMR (59 MHz, CDCl₃) δ -32.7; IR (film) 3081, 3048, 2980, 2932, 1664, 1634, 1461, 1447, 1391, 1369, 1260, 1128, 1074, 1008, 974, 928, 881, 818, 614 cm⁻¹.

Allyl-(*E*)-enolsilane (9).² b.p. 56-57 °C/ 0.5 mmHg; 71% yield; E/Z = 18:1; ¹H NMR (400 MHz, C₆D₆) δ 6.32-6.37 (dq, 1H, J = 1.6, 12.0 Hz, CH=CHO), 5.82-5.90 (m, 1H, CH₂=CH), 5.24 (dq, 1H, J = 6.8, 12.0 Hz, CH=CHO), 5.01-5.06 (m, 1H, one of CH₂=CH), 4.94-4.97 (m, 1H, one of CH₂=CH), 1.77 (dt, 2H, J = 1.2, 7.8 Hz, CH₂Si), 1.35 (dd, 3H, *J* = 1.7, 6.8 Hz, *CH*₃CH), 1.17 (s, 6H, six of (*CH*₃)₂CC(*CH*₃)₂), 1.09 (s, 6H, six of (*CH*₃)₂CC(*CH*₃)₂); ¹³C NMR (100 MHz, C₆D₆) δ 139.0, 131.7, 115.9, 108.1, 81.6, 25.8, 19.1, 12.0; ²⁹Si NMR (59 MHz, CDCl₃) δ -32.7; IR (film) 3082, 3035, 2981, 2927, 2886, 1669, 1635, 1460, 1392, 1372, 1270, 1176, 1142, 1007, 973, 932, 878, 817, 608 cm⁻¹.

Allylenolsilane (25).⁸ b.p. 50-53 °C/ 0.9 mmHg; 61% yield; ¹H NMR (400 MHz, C₆D₆) δ 5.82-5.93 (m, 1H, CH₂=C*H*), 5.01-5.08 (m, 1H, one of C*H*₂=CH), 4.94-4.98 (m, 1H, one of C*H*₂=CH), 4.57 (s, 1H, one of C*H*₂=CCH₃), 4.15 (s, 1H, one of C*H*₂=CCH₃), 1.83 (d, 2H, *J* = 7.8 Hz, C*H*₂Si), 1.74 (s, 3H, CH₂=CC*H*₃), 1.19 (s, 6H, six of (C*H*₃)₂CC(C*H*₃)₂), 1.10 (s, 6H, six of (C*H*₃)₂CC(C*H*₃)₂); ¹³C NMR (100 MHz, C₆D₆) δ 154.4, 131.9, 115.8, 94.0, 81.7, 25.8, 25.7, 22.1, 19.0; ²⁹Si NMR (59 MHz, CDCl₃) δ -33.9; IR (film) 3114, 3081, 2980, 2927, 1644, 1445, 1390, 1372, 1270, 1145, 1064, 1008, 971, 933, 902, 881, 836, 815, 782, 695 cm⁻¹.

(Z)-Crotylenolsilane (33).⁸ b.p. 58-59 °C/ 0.7 mmHg; 36% yield; E/Z < 1:100; ¹H NMR (400 MHz, C₆D₆) δ 5.55-5.64 (m, 1H, CH₃CH=CH), 5.42-5.51 (m, 1H, CH₃CH=CH), 4.58 (s, 1H, one of CH₂=CCH₃), 4.15 (m, 1H, one of CH₂=CCH₃), 1.83 (dm, 2H, J = 8.3Hz, CH₂Si), 1.75 (d, 3H, J = 0.6 Hz, CH₂=CCH₃), 1.57 (dm, 3H, J = 6.6 Hz, CH₃CH=CH), 1.20 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.12 (s, 6H, six of (CH₃)₂CC(CH₃)₂); ¹³C NMR (100 MHz, CDCl₃) δ 153.6, 124.2, 122.2, 93.6, 81.6, 25.8, 25.7, 22.2, 12.8, 12.4; ²⁹Si NMR (59 MHz, CDCl₃) δ -33.6; IR (film) 3022, 2980, 2927, 1648, 1445, 1390, 1370, 1271, 1146, 1064, 1008, 972, 933, 880, 818, 697, 670, 639 cm⁻¹.

(*E*)-Crotylenolsilane (34).⁸ b.p. 45-46 °C/ 0.3 mmHg; 54% yield; E/Z = 28:1; ¹H NMR (400 MHz, C₆D₆) δ 5.47-5.57 (m, 1H, CH₃CH=CH), 5.36-5.46 (m, 1H, CH₃CH=C*H*), 4.59 (s, 1H, one of CH₂=CCH₃), 4.16 (s, 1H, one of CH₂=CCH₃), 1.80 (dq, 2H, J = 1.2, 7.4 Hz, CH₂Si), 1.75 (s, 3H, CH₂=CCH₃), 1.55 (dt, 3H, J = 1.3, 6.2 Hz, CH₃CH=CH), 1.21 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.12 (s, 6H, six of (CH₃)₂CC(CH₃)₂); ¹³C NMR (100 MHz, C₆D₆) δ 154.3, 126.2, 124.0, 93.9, 81.8, 26.2, 26.1, 22.5, 18.4, 17.6; ²⁹Si NMR (59 MHz, CDCl₃) δ -33.3; IR (film) 3114, 3020, 2980, 2936, 2887, 2858, 1645, 1445, 1390, 1370, 1271, 1182, 1162, 1146, 1063, 1007, 972, 933, 903, 879, 836, 817, 759, 642 cm⁻¹.

Methallylenolsilane (35).⁸ b.p. 40-44 °C/ 0.2 mmHg; 67% yield; ¹H NMR (400 MHz, C₆D₆) δ 4.82 (m, 1H, one of CH₂=CCH₂), 4.79 (m, 1H, one of CH₂=CCH₂), 4.59 (s, 1H, one of CH₂=CO), 4.15 (s, 1H, one of CH₂=CO), 1.86 (d, 2H, J = 0.8 Hz, CH₂Si), 1.83 (s, 3H, CH₃CO), 1.75 (m, 3H, CH₃CCH₂), 1.19 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.11 (s, 6H, six of (CH₃)₂CC(CH₃)₂); ¹³C NMR (100 MHz, C₆D₆) δ 154.3, 140.2, 111.6, 93.9, 81.8, 26.1 (four CH₃), 25.2, 23.6, 22.5; ²⁹Si NMR (59 MHz, CDCl₃) δ -33.7; IR (film) 3078, 2980, 2936, 1646, 1446, 1390, 1371, 1269, 1147, 1064, 1008, 972, 933, 880, 842, 814 cm⁻¹.

Angelylenolsilane (36).⁸ b.p. 55-59 °C/ 0.3 mmHg; 38% yield; E/Z < 1:100; ¹H NMR (400 MHz, C₆D₆) δ 5.22 (qm, 1H, J = 6.8 Hz, CH₃CH=CCH₂), 4.60 (s, 1H, one of CH₂=CO), 4.15 (s, 1H, one of CH₂=CO), 1.84-1.86 (m, 5H, CH₃CCH₂Si), 1.75 (s, 3H, CH₃CO), 1.63 (dm, 3H, J = 6.8 Hz, CH₃CH=CCH₂), 1.20 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.12 (s, 6H, six of (CH₃)₂CC(CH₃)₂); ¹³C NMR (100 MHz, C₆D₆) δ 154.3, 130.4, 119.0, 93.7, 81.6, 26.3, 26.1, 26.0, 25.2, 23.6, 22.5, 17.7, 14.3; ²⁹Si NMR (59 MHz, CDCl₃) δ -33.3; IR (film) 3115, 2980, 2931, 1647, 1445, 1373, 1271, 1146, 1062, 1007, 972, 931, 880, 821, 712, 643 cm⁻¹.

Allylenolsilane (54). b.p. 80-85 °C/ 0.2 mmHg; 56% yield; ¹H NMR (400 MHz, C₆D₆) δ 5.86-5.98 (m, 1H, CH₂=CH), 5.29 (m, 1H, OC=CH), 5.02-5.09 (m, 1H, one of CH₂=CH), 4.94-4.99 (m, 1H, one of CH₂=CH), 2.08-2.15 (m, 2H, two of CH₂CH₂CH₂CH₂), 1.86-1.93 (m, 2H, two of CH₂CH₂CH₂CH₂CH₂), 1.85 (dt, 2H, J = 1.3, 7.8 Hz, CH₂Si), 1.43-1.51 (m, 2H, two of CH₂CH₂CH₂CH₂), 1.30-1.39 (m, 2H, two of CH₂CH₂CH₂CH₂), 1.21 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.12 (s, 6H, six of (CH₃)₂CC(CH₃)₂); ¹³C NMR (100 MHz, C₆D₆) δ 149.1, 132.2, 115.6, 106.3, 81.7, 29.9, 26.3, 26.2, 24.3, 23.6, 22.8, 19.7; ²⁹Si NMR (59 MHz, CDCl₃) δ -33.7; IR (film) 3079, 2979, 2930, 2859, 2841, 1676, 1634, 1444, 1389, 1369, 1266, 1184, 1144, 1008, 971, 936, 899, 879, 817 cm⁻¹.

Allylenolsilane (55). b.p. 57-58 °C/ 0.2 mmHg; 56% yield; ¹H NMR (400 MHz, C₆D₆) δ 5.81-5.93 (m, 1H, CH₂=CH), 5.00-5.08 (m, 2H, OC=CH and one of CH₂=CH), 4.92-4.99 (m, 1H, one of CH₂=CH), 2.30-2.37 (m, 2H, two of CH₂CH₂CH₂), 2.15-2.21 (m, 2H, two of CH₂CH₂CH₂), 1.83 (dt, 2H, J = 1.3, 7.8 Hz, CH₂Si), 1.65-1.75 (m, 2H, CH₂CH₂CH₂), 1.20 (s, 6H, six of (CH₃)₂CC(CH₃)₂), 1.11 (s, 6H, six of (CH₃)₂CC(CH₃)₂); ¹³C NMR (100 MHz, C₆D₆) δ 153.3, 131.8, 115.8, 105.2, 81.9, 33.6, 29.2, 26.2, 26.1, 22.0, 19.4; ²⁹Si NMR (59 MHz, CDCl₃) δ -32.9; IR (film) 3080, 2979, 2937, 2853, 1652, 1449, 1390, 1377, 1340, 1243, 1182, 1144, 1032, 1008, 973, 898, 882, 817, 611 cm⁻¹.

General procedure for the tandem aldol-allylation reactions.^{2,8} To a stirred solution of the allyl(crotyl)enolsilane (0.60 mmol) in toluene (0.4 mL) was added cyclohexanecarboxaldehyde (48.5 μ L; 0.40 mmol). The solution was stirred at the temperature indicated in the appropriate scheme for the indicated time. After the solution was cooled, ethyl acetate (20 mL) and aqueous HCl (1M; 5 mL) were added. The mixture was stirred at room temperature for 1 h. The organic layer was separated and washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel to afford the 1,3-diol product typically as a clear colorless oil.

General synthetic procedure for generating acetonides from 1,3-diols.^{2,8} To a stirred solution of 1,3-diol (1.00 mmol) in CH_2Cl_2 (3.0 mL) were added 2,2-dimethoxypropane

(0.90 mL; 7.17 mmol) and (+)-camphorsulfonic acid (CSA) (27 mg; 0.12 mmol). The reaction was stirred at room temperature for 15 h. Ethyl acetate (40 mL) and saturated aqueous NaHCO₃ (20 mL) were added. The organic layer was separated and washed with saturated aqueous NaCl, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel to afford the acetonide product typically as a colorless oil.

General synthetic procedure for generating cyclic tetrahydropyrans from 1,3-diols.⁸

To a stirred solution of 1,3-diol (0.50 mmol) in CH_2Cl_2 (5.0 mL) was added mercury (II) acetate chloride (236 mg; 0.80 mmol). The reaction was stirred at room temperature for 15 h. Ethyl acetate (40 mL) was added and the solution was washed with saturated aqueous NaCl, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel to afford the tetrahydropyran product as an oil or solid.

Compound (11).^{2 1}H NMR (400 MHz, CDCl₃) δ 5.83-5.93 (m, 1H, CH=CH₂), 5.14-5.18 (m, 2H, CH=CH₂), 3.69 (dt, 1H, J = 3.0, 8.1 Hz, CHOHCH₂), 3.39 (dd, 1H, J = 1.9, 8.8 Hz, c-C₆H₁₁CH), 3.20-3.24 (br s, 1H, OH), 3.13-3.18 (br s, 1H, OH), 2.44-2.49 (m, 1H, one of CH₂CH=CH₂), 2.14-2.20 (m, 1H, one of CH₂CH=CH₂), 1.62-1.80 (m, 5H, CHCH₃ and four of c-C₆H₁₁), 1.48-1.54 (m, 2H, two of c-C₆H₁₁), 1.12-1.41 (m, 5H, five of c-C₆H₁₁), 0.81 (d, 3H, J = 6.9 Hz, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ 134.8, 118.4,

80.8, 75.4, 40.4, 40.0, 39.4, 30.6, 26.7, 26.5, 26.3, 24.5, 13.1; IR (KBr) 3321, 3078, 2978, 2930, 2853, 1645, 1446, 1338, 1266, 1111, 1046, 982, 912, 861, 702 cm⁻¹; LRMS (FAB+) calcd for C₁₃H₂₄O₂: 212, found (M+H)⁺: 213.

Compound (13).^{2 1}H NMR (400 MHz, CDCl₃) δ 5.86-5.96 (m, 1H, CH=CH₂), 5.01-5.08 (m, 2H, CH=CH₂), 3.50 (m, 1H, OCHCH₂), 3.28 (d, 1H, J = 10.3 Hz, c-C₆H₁₁CH), 2.34-2.41 (m, 1H, one of CH₂CH=CH₂), 2.13-2.21 (m, 1H, one of CH₂CH=CH₂), 1.71-1.77 (m, 2H, two of c-C₆H₁₁), 1.37 (s, 3H, three of C(CH₃)₂), 1.34 (s, 3H, three of C(CH₃)₂), 1.11-1.65 (m, 10H, CHCH₃ and nine of c-C₆H₁₁), 0.74 (d, 3H, J = 6.6 Hz, CHCH₃); ¹³C NMR (100 MHz, CDCl₃) δ 135.3, 116.1, 97.7, 77.7, 74.2, 38.4, 37.6, 34.2, 30.4, 30.1, 26.9, 26.6, 26.5, 24.8, 19.5, 11.8; IR (film) 3075, 2991, 2931, 2853, 1641, 1451, 1379, 1264, 1203, 1176, 1131, 1051, 1015, 987, 935, 910 cm⁻¹.

Compound (12).^{2 1}H NMR (400 MHz, CDCl₃) δ 5.85-5.86 (m, 1H, CH=CH₂), 5.10-5.17 (m, 2H, CH=CH₂), 3.84-3.88 (m, 1H, CHOHCH₂), 3.45 (dd, 1H, J = 1.9, 9.2 Hz, c-C₆H₁₁CH), 2.78 (br s, 2H, two OH), 2.27-2.35 (m, 1H, one of CH₂CH=CH₂), 2.18-2.25 (m, 1H, one of CH₂CH=CH₂), 2.01-2.08 (m, 1H, one of c-C₆H₁₁), 1.56-1.81 (m, 5H, CHCH₃ and four of c-C₆H₁₁), 1.36-1.46 (m, 1H, one of c-C₆H₁₁), 1.08-1.30 (m, 3H, three of c-C₆H₁₁), 0.90 (d, 3H, J = 7.1 Hz, CH₃), 0.79-0.98 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 135.0, 117.8, 81.8, 76.2, 40.8, 39.9, 36.8, 29.7, 28.9, 26.3, 26.0, 25.8, 4.3.

Compound (14).^{2 1}H NMR (400 MHz, CDCl₃) δ 5.71-5.82 (m, 1H, CH=CH₂), 5.08-5.14 (m, 1H, one of CH=CH₂), 5.03-5.07 (m, 1H, one of CH=CH₂), 3.87 (dt, 1H, J = 2.2, 7.1 Hz, OCHCH₂), 3.42 (dd, 1H, J = 2.0, 9.6 Hz, c-C₆H₁₁CH), 2.26-2.34 (m, 1H, one of CH₂CH=CH₂), 2.10-2.18 (m, 1H, one of CH₂CH=CH₂), 2.02-2.09 (m, 1H, one of c-C₆H₁₁), 1.57-1.75 (m, 4H, four of c-C₆H₁₁), 1.44-1.50 (tq, 1H, J = 2.1, 6.8 Hz, CHCH₃), 1.37-1.45 (m, 1H, one of c-C₆H₁₁), 1.39 (s, 6H, C(CH₃)₂), 1.08-1.29 (m, 3H, three of c-C₆H₁₁), 0.84 (d, 3H, J = 6.8 Hz, CHCH₃), 0.73-0.86 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 134.6, 116.8, 98.8, 78.0, 73.2, 38.3, 37.3, 31.8, 30.1, 30.0, 27.4, 26.6, 25.9, 25.6, 19.6, 4.5.

Compound (*syn*-27).⁸ ¹H NMR (400 MHz, CDCl₃) δ 5.82-5.93 (m, 1H, CH=CH₂), 5.08-5.17 (m, 2H, CH=CH₂), 3.78 (ddd, 1H, J = 1.9, 5.6, 10.7 Hz, c-C₆H₁₁CH), 3.11 (br s, 2H, two OH), 2.25 (d, 2H, J = 7.4 Hz, CH₂CH=CH₂), 1.72-1.85 (m, 3H, three of c-C₆H₁₁), 1.57-1.71 (m, 3H, one of c-C₆H₁₁CHCH₂ and two of c-C₆H₁₁), 1.47-1.52 (dd, 1H, J = 1.9, 14.5 Hz, one of c-C₆H₁₁CHCH₂), 1.26 (s, 3H, CH₃), 0.95-1.36 (m, 6H, six of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 133.6, 118.9, 73.5, 73.0, 49.1, 44.3, 42.9, 28.8, 28.0, 26.5, 26.3, 26.2, 25.6; IR (film) 3354, 3077, 2926, 2854, 1641, 1450, 1376, 1130, 1064, 998, 915, 880, 842 cm⁻¹; LRMS (FAB+) calcd for C₁₃H₂₄O₂: 212, found (M+H)⁺: 213. **Compound (29).** ¹H NMR (400 MHz, CDCl₃) δ 5.79-5.90 (m, 1H, C*H*=CH₂), 5.01-5.09 (m, 2H, CH=C*H*₂), 3.59 (ddd, 1H, *J* = 2.7, 7.3, 11.3 Hz, c-C₆H₁₁C*H*), 2.14-2.28 (m, 2H, C*H*₂CH=CH₂), 1.93-1.99 (m, 1H, one of c-C₆*H*₁₁), 1.63-1.76 (m, 4H, four of c-C₆*H*₁₁), 1.42 (s, 3H, three of C(C*H*₃)₂), 1.36 (s, 3H, three of C(C*H*₃)₂), 1.28 (s, 3H, CH₂CC*H*₃), 1.09-1.45 (m, 6H, c-C₆H₁₁CHC*H*₂ and four of c-C₆*H*₁₁), 0.87-0.99 (m, 2H, two of c-C₆*H*₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 134.4, 117.5, 98.3, 72.6, 69.7, 50.1, 42.7, 36.3, 31.7, 29.0, 28.2, 26.6, 26.5, 26.1, 25.9, 25.0; IR (film) 3075, 2986, 2926, 2854, 1641, 1450, 1374, 1286, 1254, 1086, 998, 913,856 cm⁻¹.

Compound (31).⁸ 88% yield; ¹H NMR (400 MHz, CDCl₃) δ 3.98-4.05 (m, 1H, CHCH₂Hg), 3.47 (ddd, 1H, J = 2.0, 6.5, 11.6 Hz, c-C₆H₁₁CH), 2.29 (dd, 1H, J = 5.0, 11.8 Hz, one of CH₂Hg), 2.02 (dd, 1H, J = 7.1, 11.8 Hz, one of CH₂Hg), 1.85-1.93 (m, 1H, one of c-C₆H₁₁), 1.49-1.79 (m, 6H, OH, one of CH₂CCH₂ and four of c-C₆H₁₁), 1.26 (s, 3H, CH₃), 0.92-1.37 (m, 9H, three of CH₂CCH₂ and six of c-C₆H₁₁); ¹³C NMR (100 MHz, C₆D₆) δ 77.1, 71.4, 68.0, 48.2, 42.9, 41.0, 38.0, 31.7, 29.3, 29.1, 27.0, 26.6, 26.5; IR (film) 3418, 2924, 2852, 1449, 1373, 1325, 1269, 1166, 1069, 1019, 914, 811, 733 cm⁻¹; LRMS (FAB+) calcd for C₁₃H₂₃O₂HgCl: 448, found M⁺: 448.

Compound (*anti*-27).⁸ ¹H NMR (400 MHz, CDCl₃) δ 5.75-5.86 (m, 1H, CH=CH₂), 5.14 (m, 1H, one of CH=CH₂), 5.09-5.12 (m, 1H, one of CH=CH₂), 3.78 (dt, 1H, J = 5.2, 7.8 Hz, c-C₆H₁₁CH), 2.99 (br s, 2H, two OH), 2.31-2.43 (m, 2H, CH₂CH=CH₂), 1.72-1.85 (m,
3H, three of c-C₆H₁₁), 1.64-1.71 (m, 2H, two of c-C₆H₁₁), 1.59-1.62 (m, 2H, c-C₆H₁₁CHCH₂), 1.21 (s, 3H, CH₃), 1.08-1.37 (m, 4H, four of c-C₆H₁₁), 0.94-1.07 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 134.1, 118.6, 73.3, 73.0, 44.9, 44.4, 42.8, 29.1, 28.7, 28.1, 26.5, 26.2, 26.1; IR (film) 3342, 3077, 2926, 2854, 1641, 1450, 1374, 1309, 1257, 1130, 1064, 997, 914, 843 cm⁻¹.

Compound (30). ¹H NMR (400 MHz, CDCl₃) δ 5.78-5.89 (m, 1H, CH=CH₂), 5.01-5.10 (m, 2H, CH=CH₂), 3.56 (ddd, 1H, J = 3.0, 7.3, 11.3 Hz, c-C₆H₁₁CH), 2.44-2.51 (dd, 1H, J = 6.5, 14.0 Hz, one of CH₂CH=CH₂), 2.21-2.28 (dd, 1H, J = 7.9, 14.0 Hz, one of CH₂CH=CH₂), 1.92-1.98 (m, 1H, one of c-C₆H₁₁), 1.63-1.76 (m, 4H, four of c-C₆H₁₁), 1.58-1.62 (dd, 1H, J = 3.0, 13.1 Hz, one of c-C₆H₁₁CHCH₂), 1.41 (s, 3H, three of C(CH₃)₂), 1.38 (s, 3H, three of C(CH₃)₂), 1.20 (s, 3H, CH₂CCH₃), 1.09-1.43 (m, 5H, one of c-C₆H₁₁CHCH₂ and four of c-C₆H₁₁), 0.88-1.00 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 134.8, 117.6, 98.6, 72.8, 69.9, 45.3, 42.7, 37.2, 31.7, 30.4, 29.0, 28.1, 26.6, 26.1, 25.9, 25.4; IR (film) 3074, 2982, 2927, 2854, 1640, 1450, 1373, 1232, 1196, 1152, 1072, 1021, 994, 913, 858 cm⁻¹.

Compound (32).⁸ 81% yield; ¹H NMR (400 MHz, CDCl₃) δ 3.65-3.73 (m, 1H, CHCH₂Hg), 3.13 (ddd, 1H, J = 1.7, 6.3, 11.6 Hz, c-C₆H₁₁CH), 2.31 (dd, 1H, J = 4.9, 11.9 Hz, one of CH₂Hg), 2.06 (dd, 1H, J = 7.0, 11.9 Hz, one of CH₂Hg), 1.84-1.92 (m, 1H, one of c-C₆H₁₁), 1.60-1.82 (m, 5H, one of CH₂CCH₂ and four of c-C₆H₁₁), 1.29 (s, 3H,

CH₃), 1.11-1.47 (m, 8H, OH, three of CH₂CCH₂ and four of c-C₆H₁₁), 0.85-1.07 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, C₆D₆) δ 79.3, 73.2, 68.8, 50.2, 43.1, 43.0, 38.0, 29.3, 29.0, 26.9, 26.5, 26.4, 26.0; IR (KBr) 3427, 2924, 2844, 1444, 1378, 1325, 1268, 1144, 1100, 1083, 1017, 922, 867 cm⁻¹; LRMS (FAB+) calcd for C₁₃H₂₃O₂HgCl: 448, found M⁺: 448.

Compound (45).⁸ ¹H NMR (400 MHz, CDCl₃) δ 5.75-5.85 (ddd, 1H, J = 8.6, 11.1, 16.3 Hz, CH=CH₂), 5.12 (m, 1H, one of CH=C H_2), 5.07-5.10 (m, 1H, one of CH=C H_2), 3.76 (ddd, 1H, J = 2.6, 5.6, 9.7 Hz, c-C₆H₁₁CH), 3.27 (br s, 2H, two OH), 2.17-2.25 (m, 1H, CHCH₃), 1.63-1.86 (m, 5H, five of c-C₆ H_{11}), 1.48-1.61 (m, 2H, c-C₆ H_{11} CHC H_2), 1.20 (s, 3H, C H_3 COH), 1.12-1.37 (m, 4H, four of c-C₆ H_{11}), 1.03 (d, 3H, J = 6.9 Hz, CHC H_3), 0.95-1.11 (m, 2H, two of c-C₆ H_{11}); ¹³C NMR (100 MHz, CDCl₃) δ 139.6, 116.5, 74.8, 73.1, 50.3, 44.4, 40.6, 29.0, 28.2, 26.7, 26.4, 26.3, 23.1, 14.9; IR (film) 3350, 3077, 2974, 2924, 2853, 1637, 1450, 1420, 1375, 1280, 1188, 1128, 1063, 1002, 939, 913, 838 cm⁻¹; LRMS (FAB+) calcd for C₁₄ H_{26} O₂: 226, found (M+H)⁺: 227.

Compound (49).⁸ 62% yield; ¹H NMR (400 MHz, CDCl₃) δ 3.67 (ddd, 1H, J = 4.5, 7.7, 9.5 Hz, CHCH₂Hg), 3.48 (ddd, 1H, J = 2.0, 6.6, 11.7 Hz, c-C₆H₁₁CH), 2.37 (dd, 1H, J = 4.5, 11.8 Hz, one of CH₂Hg), 2.10 (dd, 1H, J = 7.8, 11.8 Hz, one of CH₂Hg), 1.87-1.94 (m, 1H, one of c-C₆H₁₁), 1.59-1.79 (m, 5H, one of c-C₆H₁₁CHCH₂ and four of c-C₆H₁₁), 1.39 (dd, 1H, J = 11.8, 13.6 Hz, one of c-C₆H₁₁CHCH₂), 1.22 (s, 3H, CH₃COH),

1.07-1.33 (m, 6H, O*H*, C*H*CH₃ and four of c-C₆*H*₁₁), 0.90 (d, 3H, *J* = 6.9 Hz, CHC*H*₃), 0.88-1.06 (m, 2H, two of c-C₆*H*₁₁); ¹³C NMR (100 MHz, C₆D₆) δ 76.8, 76.7, 69.8, 48.7, 43.5, 43.0, 37.7, 29.7, 29.4, 29.3, 27.3, 26.9, 26.8, 10.4; IR (film) 3575, 2968, 2929, 2849, 1447, 1373, 1327, 1284, 1164, 1120, 1098, 1074, 1033, 1016, 938, 892, 851, 735 cm⁻¹; LRMS (FAB+) calcd for C₁₄H₂₅O₂HgCl: 462, found M⁺: 462.

Compound (46).⁸ ¹H NMR (400 MHz, CDCl₃) δ 5.76-5.86 (m, 1H, C*H*=CH₂), 5.12 (m, 1H, one of CH=C*H*₂), 5.07-5.10 (m, 1H, one of CH=C*H*₂), 3.76 (m, 1H, c-C₆H₁₁C*H*), 3.31 (br s, 1H, O*H*), 2.80 (br s, 1H, O*H*), 2.18-2.26 (m, 1H, C*H*CH₃), 1.72-1.87 (m, 3H, three of c-C₆H₁₁), 1.63-1.70 (m, 2H, two of c-C₆H₁₁), 1.46-1.57 (m, 2H, c-C₆H₁₁CHC*H*₂), 1.24 (s, 3H, C*H*₃COH), 1.11-1.35 (m, 4H, four of c-C₆H₁₁), 1.04 (d, 3H, *J* = 6.9 Hz, CHC*H*₃), 0.95-1.10 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 139.9, 116.2, 74.9, 73.6, 50.4, 44.4, 40.1, 29.0, 28.2, 26.7, 26.4, 26.3, 23.2, 14.6; IR (film) 3350, 3077, 2974, 2924, 2852, 1637, 1450, 1419, 1378, 1309, 1102, 1062, 1038, 998, 912, 835 cm⁻¹; LRMS (FAB+) calcd for C₁₄H₂₆O₂: 226, found (M+H)⁺: 227.

Compound (50).⁸ 89% yield; ¹H NMR (400 MHz, CDCl₃) δ 4.27 (ddd, 1H, J = 2.2, 6.2, 8.8 Hz, CHCH₂Hg), 3.47 (ddd, 1H, J = 2.9, 6.6, 11.3 Hz, c-C₆H₁₁CH), 2.04-2.16 (m, 2H, CH₂Hg), 1.86-1.95 (m, 1H, one of c-C₆H₁₁), 1.68-1.79 (m, 2H, two of c-C₆H₁₁), 1.59-1.68 (m, 2H, two of c-C₆H₁₁), 1.24 (s, 3H, CH₃COH), 1.08-1.45 (m, 8H, c-C₆H₁₁CHCH₂, OH, CHCH₃ and four of c-C₆H₁₁), 0.91-1.05 (m, 2H, two of c-C₆H₁₁),

0.89 (d, 3H, *J* = 7.1 Hz, CHC*H*₃); ¹³C NMR (100 MHz, CDCl₃) δ 77.8, 72.8, 72.3, 45.1, 42.6, 36.5, 35.9, 30.0, 29.2, 28.9, 26.7, 26.3, 26.2, 8.9; IR (film) 3450, 2926, 2853, 1450, 1376, 1332, 1200, 1121, 1100, 1066, 1038, 1006, 970, 951, 924, 890, 834 cm⁻¹; LRMS (FAB+) calcd for C₁₄H₂₅O₂HgCl: 462, found M⁺: 462.

Compound (47).⁸ ¹H NMR (400 MHz, CDCl₃) δ 4.94 (m, 1H, one of CH₂C=CH₂), 4.75 (m, 1H, one of CH₂C=CH₂), 3.78 (dd, 1H, J = 5.5, 9.8 Hz, c-C₆H₁₁CH), 3.27 (br s, 1H, OH), 2.89 (br s, 1H, OH), 2.18-2.26 (m, 2H, CH₂C=CH₂), 1.85 (s, 3H, CH₃C=CH₂), 1.72-1.86 (m, 3H, three of c-C₆H₁₁), 1.60-1.71 (m, 3H, one of c-C₆H₁₁CHCH₂ and two of c-C₆H₁₁), 1.47-1.52 (dd, 1H, J = 1.8, 14.5 Hz, one of c-C₆H₁₁CHCH₂), 1.28 (s, 3H, CH₃COH), 1.11-1.36 (m, 4H, four of c-C₆H₁₁), 0.95-1.10 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 142.1, 115.0, 73.4, 73.2, 52.2, 44.4, 43.3, 29.0, 28.3, 26.7, 26.4, 26.3, 26.2, 25.2; IR (film) 3350, 3073, 2924, 2853, 1643, 1450, 1376, 1318, 1260, 1186, 1130, 1039, 981, 892, 860, 735 cm⁻¹; LRMS (FAB+) calcd for C₁₄H₂₆O₂: 226, found (M+H)⁺: 227.

Compound (51).⁸ 16% yield; ¹H NMR (400 MHz, CDCl₃) δ 3.63 (ddd, 1H, J = 1.8, 6.2, 11.6 Hz, c-C₆H₁₁CH), 2.29 (d, 1H, J = 11.5 Hz, one of CH₂Hg), 2.18 (d, 1H, J = 11.6 Hz, one of CH₂Hg), 1.82-1.90 (m, 1H, one of c-C₆H₁₁), 1.60-1.79 (m, 5H, one of CH₂CCH₂ and four of c-C₆H₁₁), 1.46 (s, 3H, CH₃CCH₂Hg), 1.42-1.57 (m, 1H, one of c-C₆H₁₁), 1.35 (d, 1H, J = 13.9 Hz, one of CH₂CCH₂), 1.25 (s, 3H, CH₃COH), 1.10-1.34 (m, 6H, OH,

two of CH_2CCH_2 and three of c-C₆ H_{11}), 0.91-1.10 (m, 2H, two of c-C₆ H_{11}); ¹³C NMR (100 MHz, CDCl₃) δ 74.4, 71.3, 69.4, 51.6, 49.6, 42.5, 41.1, 33.4, 29.1, 29.0, 26.8, 26.7, 26.4, 26.3; IR (film) 3498, 2921, 2852, 1450, 1371, 1275, 1233, 1107, 1022, 950, 913, 889, 815, 734 cm⁻¹; LRMS (FAB+) calcd for C₁₄H₂₅O₂HgCl: 462, found M⁺: 462.

Compound (48).⁸ ¹H NMR (400 MHz, CDCl₃) δ 4.91 (m, 1H, one of CH₃C=CH₂), 4.78 (m, 1H, one of CH₃C=CH₂), 3.73-3.78 (m, 1H, c-C₆H₁₁CH), 3.46 (br s, 1H, OH), 2.75 (br s, 1H, OH), 2.25 (q, 1H, J = 7.2 Hz, CHCH₃), 1.79 (s, 3H, CH₃C=CH₂), 1.63-1.85 (m, 5H, five of c-C₆H₁₁), 1.50-1.62 (m, 2H, c-C₆H₁₁CHCH₂), 1.22 (s, 3H, CH₃COH), 1.11-1.38 (m, 4H, four of c-C₆H₁₁), 1.08 (d, 3H, J = 7.2 Hz, CHCH₃), 0.95-1.10 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 147.2, 113.8, 74.9, 73.2, 52.4, 44.4, 41.8, 29.1, 28.3, 26.7, 26.5, 26.4, 23.6, 22.1, 14.4; IR (film) 3368, 3072, 2974, 2925, 2853, 1640, 1450, 1375, 1133, 1066, 981, 937, 892, 863, 841 cm⁻¹; LRMS (FAB+) calcd for C₁₅H₂₈O₂: 240, found (M+H)⁺: 241.

Compound (52).⁸ 11% yield; ¹H NMR (400 MHz, CDCl₃) δ 3.57 (ddd, 1H, J = 2.9, 6.3, 11.3 Hz, c-C₆H₁₁CH), 2.34 (d, 1H, J = 11.5 Hz, one of CH₂Hg), 2.25 (d, 1H, J = 11.5 Hz, one of CH₂Hg), 1.85-1.92 (m, 1H, one of c-C₆H₁₁), 1.69-1.79 (m, 2H, two of c-C₆H₁₁), 1.60-1.68 (m, 2H, two of c-C₆H₁₁), 1.57 (d, 1H, J = 13.5 Hz, one of c-C₆H₁₁CHCH₂), 1.34 (s, 3H, CH₃CCH₂Hg), 1.21 (s, 3H, CH₃COH), 1.11-1.41 (m, 7H, OH, CHCH₃, one of c-C₆H₁₁CHCH₂ and four of c-C₆H₁₁), 0.99 (d, 3H, J = 7.1 Hz, CHCH₃), 0.92-1.08 (m,

Part I

2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, C₆D₆) δ 78.2, 71.0, 70.4, 49.9, 49.8, 43.6, 43.0, 30.8, 29.7, 29.3, 27.4, 27.0, 26.8, 22.6, 10.5; IR (film) 3611, 3536, 2963, 2933, 2847, 1449, 1372, 1226, 1191, 1102, 1047, 1025, 973, 943, 863, 760 cm⁻¹; LRMS (FAB+) calcd for C₁₅H₂₇O₂HgCl: 476, found M⁺: 476.

Compound (53).⁸ 32% yield; ¹H NMR (400 MHz, CDCl₃) δ 3.56 (ddd, 1H, J = 2.9, 6.3, 11.3 Hz, c-C₆H₁₁CH), 3.21 (d, 1H, J = 12.4 Hz, one of CH₂Hg), 2.14 (d, 1H, J = 12.4 Hz, one of CH₂Hg), 1.96-2.03 (m, 1H, one of c-C₆H₁₁), 1.64 (dd, 1H, J = 1.7, 13.6 Hz, one of c-C₆H₁₁CHCH₂), 1.60-1.81 (m, 4H, four of c-C₆H₁₁), 1.53 (q, 1H, J = 7.1 Hz, CHCH₃), 1.38 (dd, 1H, J = 11.6, 13.4 Hz, one of c-C₆H₁₁CHCH₂), 1.30 (s, 3H, CH₃CCH₂Hg), 1.22 (s, 3H, CH₃COH), 1.12-1.32 (m, 4H, four of c-C₆H₁₁), 0.98 (d, 3H, J = 7.1 Hz, CHCH₃), 0.94-1.09 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, C₆D₆) δ 77.8, 70.4, 70.3, 48.4, 43.6, 43.3, 40.0, 34.5, 30.7, 30.0, 29.9, 27.3, 27.0, 26.9, 10.4; IR (film) 3522, 3448, 2929, 2851, 1451, 1423, 1375, 1294, 1267, 1178, 1156, 1099, 1067, 1048, 1010, 949, 899, 860, 790 cm⁻¹; LRMS (FAB+) calcd for C₁₅H₂₇O₂HgCl: 476, found M⁺: 476.

Compound (58). ¹H NMR (400 MHz, CDCl₃) δ 5.73-5.85 (m, 1H, CH=CH₂), 5.07-5.13 (m, 2H, CH=CH₂), 3.77 (d, 1H, J = 9.4 Hz, c-C₆H₁₁CH), 2.98 (br s, 1H, OH), 2.93 (br s, 1H, OH), 2.34-2.46 (m, 2H, CH₂CH=CH₂), 2.04-2.11 (m, 1H, one of c-C₆H₁₁), 1.08-1.82 (m, 17H, CH(CH₂)₄COH and eight of c-C₆H₁₁), 0.86-0.98 (m, 1H, one of c-C₆H₁₁), 0.72-0.83 (m, 1H, one of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 133.7, 118.3, 75.7,

74.5, 46.0, 42.4, 40.3, 37.6, 30.3, 28.7, 26.5, 26.1, 26.0, 25.9, 21.5, 19.4; IR (film) 3338, 3076, 2921, 2851, 1640, 1446, 1317, 1262, 1202, 1142, 1097, 1065, 1045, 980, 958, 914, 831, 737 cm⁻¹; LRMS (FAB+) calcd for C₁₆H₂₈O₂: 252, found (M+H)⁺: 253.

Compound (61). ¹H NMR (400 MHz, CDCl₃) δ 5.78-5.89 (m, 1H, C*H*=CH₂), 5.00-5.10 (m, 2H, CH=CH₂), 3.68 (d, 1H, *J* = 9.6 Hz, c-C₆H₁₁C*H*), 2.70-2.77 (dd, 1H, *J* = 5.9, 14.4 Hz, one of CH₂CH=CH₂), 2.21-2.28 (dd, 1H, *J* = 8.4, 14.4 Hz, one of CH₂CH=CH₂), 2.04-2.11 (m, 1H, one of c-C₆H₁₁), 1.45 (s, 3H, three of C(CH₃)₂), 1.41 (s, 3H, three of C(CH₃)₂), 1.03-1.78 (m, 17H, c-C₆H₁₁CHC*H*(CH₂)₄ and eight of c-C₆H₁₁), 0.74-0.89 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 134.6, 117.4, 98.7, 74.7, 72.4, 44.4, 38.3, 37.5, 36.4, 31.8, 30.5, 27.6, 26.8, 26.1, 25.8, 25.7, 25.4, 21.5, 20.7; IR (film) 3075, 2989, 2927, 2852, 1639, 1448, 1377, 1367, 1255,1234, 1198, 1166, 1135, 1013, 984, 911,846 cm⁻¹.

Compound (59a). ¹H NMR (400 MHz, CDCl₃) δ 5.81-5.93 (m, 1H, CH=CH₂), 5.17 (m, 1H, one of CH=CH₂), 5.11-5.15 (dm, 1H, J = 7.1 Hz, one of CH=CH₂), 3.65 (d, 1H, J = 8.7 Hz, c-C₆H₁₁CH), 3.31 (br s, 1H, OH), 2.68 (br s, 1H, OH), 2.44-2.50 (dd, 1H, J = 8.1, 13.5 Hz, one of CH₂CH=CH₂), 2.21-2.27 (dd, 1H, J = 6.8, 13.5 Hz, one of CH₂CH=CH₂), 2.03-2.10 (m, 1H, one of c-C₆H₁₁), 1.91-2.01 (m, 1H, one of (CH₂)₃COH), 1.54-1.83 (m, 10H, c-C₆H₁₁CHCH, five of (CH₂)₃COH and four of c-C₆H₁₁), 1.08-1.39 (m, 4H, four of c-C₆H₁₁), 0.81-0.99 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 133.9, 118.7,

83.0, 74.4, 47.3, 44.4, 42.4, 39.8, 30.0, 29.3, 26.6, 26.3, 26.1, 22.2, 21.5; IR (film) 3338, 3076, 2923, 2852, 1640, 1448, 1317, 1188, 1111, 1014, 914, 892 cm⁻¹; LRMS (FAB+) calcd for C₁₅H₂₆O₂: 238, found (M+H)⁺: 239.

Compound (62a). ¹H NMR (400 MHz, CDCl₃) δ 5.84-5.95 (m, 1H, CH=CH₂), 5.08 (m, 1H, one of CH=CH₂), 5.02-5.06 (m, 1H, one of CH=CH₂), 3.53 (dd, 1H, *J* = 1.8, 9.5 Hz, c-C₆H₁₁CH), 2.64-2.71 (dd, 1H, *J* = 6.9, 14.2 Hz, one of CH₂CH=CH₂), 2.15-2.22 (dd, 1H, *J* = 7.4, 14.2 Hz, one of CH₂CH=CH₂), 2.05-2.12 (m, 1H, one of c-C₆H₁₁), 1.85-1.95 (m, 1H, one of OC(CH₂)₃), 1.41 (s, 3H, three of C(CH₃)₂), 1.36 (s, 3H, three of C(CH₃)₂), 1.10-1.81 (m, 14H, c-C₆H₁₁CHCH, five of OC(CH₂)₃ and eight of c-C₆H₁₁), 0.80-0.91 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 134.9, 116.9, 98.4, 83.5, 71.5, 43.7, 41.9, 40.8, 39.8, 31.6, 30.2, 28.2, 26.8, 26.1, 25.8, 24.8, 23.5, 22.6; IR (film) 3075, 2925, 2853, 1640, 1450, 1377, 1257, 1240, 1224, 1199, 1160, 1069, 1039, 996, 977, 912, 891, 850 cm⁻¹.

Compound (59b). ¹H NMR (400 MHz, CDCl₃) δ 5.86-5.97 (m, 1H, CH=CH₂), 5.08-5.16 (m, 2H, CH=CH₂), 3.57 (d, 1H, J = 9.3 Hz, c-C₆H₁₁CH), 2.63 (br s, 1H, OH), 2.48-2.55 (dd, 1H, J = 7.8, 13.7 Hz, one of CH₂CH=CH₂), 2.29-2.35 (dd, 1H, J = 7.0, 13.7 Hz, one of CH₂CH=CH₂), 2.27 (br s, 1H, OH), 1.55-1.92 (m, 10H, c-C₆H₁₁CHCH, four of (CH₂)₃COH and five of c-C₆H₁₁), 1.05-1.48 (m, 8H, two of (CH₂)₃COH and six of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 134.5, 118.3, 82.2, 77.7, 48.5, 48.0, 41.2, 40.4,

30.9, 29.6, 26.8, 26.7, 26.5, 25.4, 23.3; IR (KBr) 3420, 3074, 2930, 2854, 1639, 1448, 1367, 1321, 1266, 1188, 1102, 1045, 1001, 975, 941, 908, 854, 619 cm⁻¹; LRMS (FAB+) calcd for C₁₅H₂₆O₂: 238, found (M+H)⁺: 239.

Compound (62b). ¹H NMR (400 MHz, CDCl₃) δ 5.85-5.96 (m, 1H, CH=CH₂), 5.09 (m, 1H, one of CH=CH₂), 5.03-5.07 (dm, 1H, J = 7.0 Hz, one of CH=CH₂), 3.20 (dd, 1H, J = 5.5, 9.0 Hz, c-C₆H₁₁CH), 2.72-2.79 (dd, 1H, J = 5.8, 14.3 Hz, one of CH₂CH=CH₂), 2.14-2.21 (dd, 1H, J = 8.1, 14.3 Hz, one of CH₂CH=CH₂), 1.60-1.90 (m, 10H, c-C₆H₁₁CHCH, four of OC(CH₂)₃ and five of c-C₆H₁₁), 1.44 (s, 3H, three of C(CH₃)₂), 1.31 (s, 3H, three of C(CH₃)₂), 0.95-1.45 (m, 8H, two of OC(CH₂)₃ and six of c-C₆H₁₁); 1³C NMR (100 MHz, CDCl₃) δ 135.0, 117.1, 99.0, 84.1, 76.9, 46.2, 46.1, 42.3, 40.2, 31.9, 30.1, 29.5, 28.3, 26.8, 26.6, 26.4, 25.7, 24.1; IR (film) 3075, 2986, 2928, 2854, 1639, 1450, 1377, 1224, 1161, 1069, 1029, 994, 911, 892 cm⁻¹.

Compound (60). ¹H NMR (400 MHz, CDCl₃) δ 3.82 (m, 1H, c-C₆H₁₁CH), 2.24-2.36 (m, 2H, two of CHCH₂CH₂CH₂), 1,91-2.16 (m, 5H, one of c-C₆H₁₁ and four of CHCH₂CH₂CH₂), 1.55-1.81 (m, 6H, OH, four of c-C₆H₁₁ and one of CHCH₂CH₂CH₂), 1.08-1.40 (m, 4H, four of c-C₆H₁₁), 0.86-1.00 (m, 2H, two of c-C₆H₁₁); ¹³C NMR (100 MHz, CDCl₃) δ 222.0, 73.9, 52.2, 41.3, 39.2, 29.7, 29.1, 26.3, 26.1, 25.8, 22.5, 20.7.

References

- [1] Frost, L. M.; Smith, J. D.; Berrisford, D. J. Tetrahedron Lett. 1999, 40(11), 2183-2186.
- [2] Wang, X.; Meng, Q.; Nation, A. J.; Leighton, J. L. J. Am. Chem. Soc. 2002, 124(36), 10672-10673.
- [3] Denmark, S. E.; Jacobs, R. T.; Dai-Ho, G.; Wilson, S. Organometallics 1990, 9(12), 3015-3019.
- [4] a) Cazeau, P.; Duboudin, F.; Moulines, F.; Babot, O.; Dunogues, J. Tetrahedron 1987, 43(9), 2075-2088. b) House, H. O.; Czuba, L. J.; Gall, M.; Olmstead, H. D. J. Org. Chem. 1969, 34(8), 2324-2336. c) Tino, J. A.; Lewis, M. D.; Kishi, Y. Heterocycles 1987, 25(1), 97-104.
 d) Heathcock, C. H.; Buse, C. T.; Kleschick, W. A.; Pirrung, M. C.; Sohn, J. E.; Lampe, J. J. Org. Chem. 1980, 45(6), 1066-1081.
- [5] a) Rychnovsky, S. D.; Skalitzky, D. J. Tetrahedron Lett. 1990, 31(7), 945-948. b)
 Rychnovsky, S. D.; Rogers, B.; Yang, G. J. Org. Chem. 1993, 58(13), 3511-3515.
- [6] Denmark, S. E.; Weber, E. J. Hel. Chim. Acta 1983, 66(6), 1655-1660.
- [7] a) Dybowski, P.; Skowrońska, A. Synthesis 1990, (7), 609-612. b) Cazeau, P.; Duboudin, F.;
 Moulines, F.; Babot, O.; Dunogues, J. Tetrahedron 1987, 43(9), 2075-2088.
- [8] Wang, X.; Meng, Q.; Perl, N. R.; Xu, Y.; Leighton, J. L. J. Am. Chem. Soc. 2005, 127(37), 12806-12807.
- [9] a) Kira, M.; Hino, T.; Sakurai, H. Tetrahedron Lett. 1989, 30(9), 1099-1102. b) Iseki, K.; Kuroki, Y.; Takahashi, M.; Kishimoto, S.; Kobayashi, Y. Tetrahedron 1997, 53(10), 3513-3526. c) Tsuji, J.; Hara, M.; Ohno, K. Tetrahedron 1974, 30(14), 2143-2146. d) Coleman, R. S.; Gurrala, S. R.; Mitra, S.; Raao, A. J. Org. Chem. 2005, 70(22), 8932-8941.
 e) Fleming, I.; Lee, D. J. Chem. Soc., Perkin Trans. 1 1998, (17), 2701-2709. f) Benkeser, R. A.; Merritt, F. M., II; Roche, R. T. J. Organometallic Chem. 1978, 156(1), 235-244.
- [10] For preparation of (E)-crotyl chloride, see: a) Magid, R. M.; Fruchey, O. S.; Johnson, W. L. *Tetrahedron Lett.* 1977, 18(35), 2999-3002. b) Magid, R. M.; Fruchey, O. S.; Johnson, W. L.; Allen, T. G. J. Org. Chem. 1979, 44(3), 359-363.

- [11] Duhamel, P.; Hennequin, L.; Poirier, J. M.; Tavel, G.; Vottero, C. Tetrahedron 1986, 42(17), 4777-4786.
- [12] House, H. O.; Czuba, L. J.; Gall, M.; Olmstead, H. D. J. Org. Chem. 1969, 34(8), 2324-2336.
- [13] For ¹H and ¹³C NMR spectral data of *anti*-hydroxyketone **60**, see: a) Enholm, E. J.; Whitley,
 P. E.; Xie, Y. J. Org. Chem. **1996**, *61(16)*, 5384-90. b) Enholm, E. J.; Whitley, P. E. *Tetrahedron Lett.* **1995**, *36(50)*, 9157-9160.
- [14] a) Matsumoto, K.; Oshima, K.; Utimoto, K. J. Org. Chem. 1994, 59(23), 7152-7155. b)
 Chemler, S. R.; Roush, W. R. J. Org. Chem. 2003, 68(4), 1319-1333. c) Sato, K.; Kira, M.;
 Sakurai, H. J. Am. Chem. Soc. 1989, 111(16), 6429-6431. d) Kira, M.; Sato, K.; Sekimoto,
 K.; Gewald, R.; Sakurai, H. Chem. Lett. 1995, (4), 281-282. e) Chemler, S. R.; Roush, W. R.
 J. Org. Chem. 1998, 63(12), 3800-3801. f) Chemler, S. R.; Roush, W. R. Tetrahedron Lett.
 1999, 40(25), 4643-4647.
- [15] a) Gewald, R.; Kira, M.; Sakurai, H. Synthesis 1996, (1), 111-115. b) Wang, Z.; Xu, G.;
 Wang, D.; Pierce, M. E.; Confalone, P. N. Tetrahedron Lett. 2000, 41(23), 4523-4526.
- [16] For examples of intermolecular type II allylation on ketones, see: a) Yamasaki, S.; Fujii, K.;
 Wada, R.; Kanai, M.; Shibasaki, M. J. Am. Chem. Soc. 2002, 124(23), 6536-6537. b)
 Cossy, J.; Lutz, F.; Alauze, V.; Meyer, C. Synlett 2002, (1), 45-48. c) Ishihara, K.;
 Hiraiwa, Y.; Yamamoto, H. Synlett 2001, (12), 1851-1854. d) Coppi, L.; Mordini, A.; Taddei,
 M. Tetrahedron Lett. 1987, 28(9), 969-972.
- [17] a) Wang, Z.; Wang, D.; Sui, X. Chem. Commun. 1996, (19), 2261-2262. b) Zhang, L. C.;
 Sakurai, H.; Kira, M. Chem. Lett. 1997, (2), 129-130.
- [18] a) Mundy, B. P.; Srinivasa, R.; Kim, Y.; Dolph, T.; Warnet, R. J. J. Org. Chem. 1982, 47(9), 1657-1661. b) Naro, P. A.; Dixon, J. A. J. Am. Chem. Soc. 1959, 81(7), 1681-1684.

Part II. Design and Synthesis of Modified Fluorescent Nucleotides for DNA Sequencing by Synthesis

Chapter 1. Introduction to DNA Sequencing by Synthesis

Genomics has become one of the most attractive and important research fields in modern biological engineering, to which many scientists are devoting their enthusiasm. This branch of science involves the cooperation and interaction of biology, chemistry, engineering, computer programming and medical sciences. DNA sequencing is a fundamental technique for genomics, biological research and medical applications. High-throughput and rapid DNA sequencing is more demanding but a requirement for pharmacogenomics, disease gene discovery and gene function studies.¹ Modern sequencing research is mainly focused on developing new and fast DNA sequencing techniques, which will reduce the sequencing cost dramatically without compromise in sequencing accuracy. The hope is that 'Personal Genomics' will become the norm in medical diagnostics and health care in the near future. Current DNA sequencing techniques still confront such limitations as sequencing cost, data accuracy and read length. More efficient new sequencing techniques are in high demand to broaden the application of genomics to modern medical research and health care. A long-term aim is to develop a fast and accurate sequencing method that will cost as little as \$1,000 per human genome. Sanger dideoxy sequencing is a frequently used sequencing method in modern genomics. To overcome the drawback of this most popular current sequencing technique, new sequencing methods, such as DNA sequencing by synthesis and nanopore DNA sequencing, are being developed.



1.1 Structures and Properties of DNA, Nucleotides and Nucleosides



Deoxyribonucleic acid (DNA) is the molecular carrier of genetic information in the cells of all organisms; it controls the cells' division, growth and death, determines the differentiated properties of cells, and directs the biosynthesis and properties of proteins and enzymes. DNA is basically a biopolymer of nucleotides, which come together to form an extremely long strand. There are four natural nucleotides, as shown in Figure 1-1 (purine numbering is used throughout this thesis), which are constituted of nucleosides and a phosphate group attached at the 5' position. Each nucleoside is composed of a 2'-deoxyribose moiety and a nucleobase. There are four nucleobases in DNA molecules: two purines (Adenine and Guanine) and two pyrimidines (Cytosine and Thymine). The structures of DNA and the genetic information it carries all depend on the sequence of individual nucleotides, namely the sequence of the nucleobases: A, G, C and T.



Figure 1-2. Watson-Crick base pairing of double-stranded DNA.

Nucleotide molecules are bound together in DNA via a phosphate ester bond between a 5'-phosphate of one nucleotide and a 3'-hydroxyl group of another nucleotide (often referred to as phosphodiester bonds). Thus a long chain of single-strand DNA is established, in which one end has a free hydroxyl group at the 3' carbon of the terminal deoxyribose (3' end) and the other end has a phosphate at the 5' carbon of the terminal deoxyribose (5' end). DNA consists of double strands of polynucleotide, which coil around each other to form a double helix in three-dimensions. The two complementary

strands extend in opposite directions and are held together by hydrogen bonds between specific pairs of nucleobases, referred to as Watson-Crick base pairing: A and T form two hydrogen bonds and C and G form three hydrogen bonds (Figure 1-2).

The DNA polymerase reaction is the fundamental step in DNA sequencing, which involves the process whereby an individual nucleotide is added to the 3' end of a growing primer (a small piece of DNA) in the presence of the DNA polymerase enzyme. In this process a nucleoside 5'-triphosphate molecule (dNTP), which is complementary to the corresponding nucleotide on the template, is added to the 3' end of the primer strand. Thus hydrogen bonds are formed between the incoming dNTP and the complementary nucleotide on the template. In the presence of DNA polymerase, the 3'-hydroxyl group of the primer end then attacks the α -phosphorus of the nucleoside triphosphate and a diphosphate anion is released. In this way a new phosphate ester bond is generated and the incoming nucleotide is incorporated at the 3' end of the primer strand, which still has a free hydroxyl group on the 3' end for the next cycle of the polymerase reaction (Scheme 1-1). In this way the primer strand grows one nucleotide at a time in the 5' to 3' direction.

It can be seen that the free hydroxyl group on the 3' end of the primer is necessary for the incorporation of the next nucleotide. Absence of the free 3'-hydroxyl group will result in the termination of the DNA polymerase reaction.



Scheme 1-1. The process of the DNA polymerase reaction: the hydroxyl group on the 3' end of the primer is necessary for the incorporation of the next nucleotide.

1.2 Sanger Dideoxy DNA Sequencing²

The Sanger DNA sequencing technique has been the most frequently used sequencing method since its debut in 1977. The current automated version employs four regular 2'-deoxyribonucleoside triphosphate (dNTP) molecules and four fluorescent dye-labeled 2',3'-dideoxyribonucleoside triphosphate (ddNTP) molecules for DNA sequencing. Each ddNTP molecule is labeled with a fluorescent dye of a distinct color. The four ddNTP molecules are used as terminators in the polymerase reaction.

As shown in Scheme 1-2, when DNA polymerase is added to the mixture of four dNTP's, primer, template to be sequenced and a very small quantity of the four ddNTP's, the primer begins to grow in the 5' to 3' direction with a sequence complementary to the template. In most cases the regular dNTP's are incorporated into the primer strand and the next cycle of nucleotide incorporation will occur, but occasionally the ddNTP's are incorporated and the polymerase reaction stops at this point due to the absence of the 3'-hydroxyl group in the ddNTP molecule. When the polymerase reaction is complete, a mixture of primer strands with each different possible length will be generated, all of which have a fluorescent dye-labeled ddNTP on the 3' end. The mixture of primer strands are separated by capillary gel electrophoresis according to their length so that a sequence ladder is obtained. Finally, identity of the fluorescent dye-labeled ddNTP's on each primer's 3' end is decoded by the color of each fluorescent dye via laser-induced

fluorescence detection. As the length of the primers differs from each other, the total sequence of the primer strand can be identified. As a result, the sequence of the complementary template will be obtained.



Scheme 1-2. Schematic illustration of Sanger dideoxy DNA sequencing by four dNTP's and four fluorescent dye-labeled ddNTP's to obtain the sequence of the primer.

Sanger dideoxy DNA sequencing is an efficient and fast sequencing method with excellent accuracy. Thanks to many advancements over the years, such as solid phase DNA synthesis, capillary gel electrophoresis, laser-induced fluorescence detection and computer programming, Sanger DNA sequencing has been improved to the point where it is able to determine unknown DNA sequences as long as 1,000 base pairs. A well-known application of this method was the human and other organism genome sequences produced by the Human Genome Project.

1.3 DNA Sequencing by Synthesis

DNA sequencing by synthesis (SBS) is an efficient approach to overcome many of the drawbacks of the current sequencing techniques, especially cost and throughput.

In Sanger DNA sequencing, the identity of each fluorescent dye-labeled ddNTP molecule is recognized after all the polymerase reactions terminate. In SBS, the identity of each nucleotide is detected immediately after it is incorporated into a growing DNA strand in a polymerase reaction. To do so, each nucleotide addition must be detected either by directly observing a reporter group attached to the nucleotide, or by observing a secondary reporter from the polymerase reaction product. Both solutions require the reporter group to be removed before the next nucleotide is incorporated so that the detection of each new addition is not affected by previous ones. The nucleotides can be added stepwise in which case only the one incorporated will produce a signal, or the four nucleotides can be added together and the incorporated one distinguished from the other three. Of course the latter is much more efficient.

As can be seen in Scheme 1-1, there are two possible methods to detect the polymerase reaction: either by identifying the incorporated nucleotide in the polymerase reaction, or by measuring the released pyrophosphate. Hyman first revealed the concept of DNA sequencing by synthesis in 1988.³ In his alternative sequencing method, the pyrophosphate released in the DNA polymerase reaction was precisely measured so that whether and how many cycles of the polymerase reactions occurred could be detected. In the polymerase reaction, the pyrophosphate anion is generated only when there is a nucleoside triphosphate molecule has been incorporated. The pyrophosphate will then be recognized via interaction with a series of enzymes and finally a light signal will be generated for detection (see Chapter 1.4).

Another approach in SBS is to attach fluorescent dyes on individual nucleoside 5'-triphosphate molecules as reporter groups. As four fluorescent dye-labeled nucleotide molecules are added in the polymerase reaction, the fluorescent dye moiety has to be cleaved after it is identified to avoid signal overlap. Ideally, the four fluorescent dye-labeled nucleotide molecules should act as temporary terminators. Each time a nucleotide molecule is incorporated into the growing DNA strand, the polymerase

reaction will terminate temporarily to allow fluorescent dye detection. After that the fluorescent dye moiety is cleaved and the polymerase reaction reinitiates.

Possible labeling sites for the fluorescent dyes are on the γ -phosphorus of the 5'-triphosphate, on the 3' hydroxyl, or on the nucleobase of the nucleotide. The advantage of attaching the fluorescent dye to the γ -phosphorus of the 5'-triphosphate (Figure 1-3, A)⁴ is that after the dye-labeled nucleotide molecule is incorporated, the fluorescent dye moiety will be released with the pyrophosphate anion and therefore a regular nucleotide is left in the DNA strand. So there will be no negative influence on Watson-Crick base pairing and stability of the double helix. Another labeling position is the 3' hydroxyl group of the ribose (Figure 1-3, B).⁵ Here the labeled fluorescent dve will act as a 3'-hydroxyl capping group as well. Therefore the polymerase reaction stops after it is incorporated into the DNA strand. Unfortunately, successful reports on such 3' modifications are rare due to the high sensitivity of the polymerase enzyme toward the size of the 3' capping group.⁶ Fluorescent dyes can also be labeled on a specific site of the nucleobase (Figure 1-3, C). It has been shown that there is little influence on base pairing and stability of the DNA double helix if a fluorescent dye moiety is attached to the 5 position of pyrimidines (cytosine or uracil) or the 7 position of purines (adenine or guanine), despite the fact that most fluorescent dyes are quite bulky.⁷ The fluorescent dves need to be attached to the nucleobase via a linker molecule, so that they can be removed conveniently after being detected to allow the next cycle of incorporation and

detection. Thus an ideal fluorescent dye-labeled nucleotide molecule is one in which there is a fluorescent dye with a unique emission wavelength attached to a specific site of the nucleobase via a cleavable linker and a small capping group at the 3'-hydroxyl position so that the modified nucleotide can be recognized by the polymerase.



Figure 1-3. Possible labeling sites of fluorescent dyes on the nucleoside triphosphate molecule.

One problem to be solved in DNA sequencing using fluorescent dye-labeled nucleotides is how to suppress undesired background noise and raise the signal / noise ratio.

1.4 DNA Pyrosequencing⁸

In the DNA polymerase reaction, the pyrophosphate anion (PPi) is generated only when a

nucleoside triphosphate molecule is incorporated (see Scheme 1-1). The DNA

pyrosequencing technique is based on measuring by bioluminescence the pyrophosphate anion released in the polymerase reaction.

DNA pyrosequencing is predicated on a four-enzyme chain reaction, in which four enzymes (DNA polymerase, ATP sulfurylase, luciferase and apyrase) are added to a mixture of template, primer, D-luciferin, and adenosine 5' phosphosulfate (APS). As shown in Scheme 1-3, in each set of four sequencing cycles, each of the four nucleoside 5' triphosphates (dNTP) will be added into the reaction mixture in turn. Only the one complementary to the first free base on the template will be incorporated into the primer strand and a pyrophosphate anion will be released as a result. Catalyzed by ATP sulfurylase, the pyrophosphate will react with APS to generate ATP. Then, D-luciferin will be converted into oxyluciferin by ATP in the presence of luciferase. In this process, a quantitative amount of light will be released and detected. The light strength is proportional to the ATP concentration and therefore the quantity of dNTP incorporated each time. Excess dNTP and ATP will be degraded by apyrase to form nucleoside monophosphate (dNMP and AMP) and phosphoric acid (Pi) so that the residual light signal is quenched. To identify each base on the template, four dNTP's (dATP S, dGTP, dCTP and dTTP) are added and detected iteratively. In the case that no dNTP is incorporated, no light signal will be generated and all of the dNTP will be simply degraded by apyrase. Here dATP S (deoxyadenosine α -thio triphosphate) is used instead

of dATP because the latter is a substrate of luciferase but the former and the other three dNTP's are not.



Scheme 1-3. Schematic process of DNA pyrosequencing in the presence of enzymes.

Commercial DNA pyrosequencing requires the following to ensure its high efficiency: 1) an optimized concentration of substrates; 2) a high quality of apyrase to hydrolyze completely all the dNTP's and ATP; 3) a decomposition rate of dNTP's that is slower than that of their incorporation into the primer so that the incorporation will be complete; and 4) an ATP generation rate that is faster than its decomposition rate by apyrase so that the light strength is proportional to the quantity of dNTP incorporated each time. For instance, if the next two nucleotides on the template strand are both dA's, twice as much dTTP will be incorporated with the production of twice as many photons.

The current DNA pyrosequencing technique has been further developed to enable sequencing of short DNA strands up to 30 base pairs with excellent data accuracy and convenience. But some difficulties have restricted its further applications. For example, in pyrosequencing four nucleotides need to be added and detected iteratively in each sequencing cycle; the increased work and time make it unsuitable for large-scale sequencing. Additionally, it is difficult to accurately identify homopolymeric regions in DNA strands with this method. It is hard to ensure a linear response of light when more than three nucleotides are incorporated in one sequencing cycle. And as the read length increases, more and more substances (such as nucleotides) will be accumulated in the reaction mixture and the sequencing accuracy will be decreased. All these difficulties call for improvements in pyrosequencing to increase its sequencing accuracy and read length.

In the last year, 454 Life Sciences Corp. has combined the pyrosequencing approach with template amplification on individual beads that are arrayed on chips by synthesis.⁹ With this commercial platform, routine 100 base sequences have been determined in a

massively parallel way. However, all the negative aspects of pyrosequencing still exist, most especially the decreased ability to accurately determine the length of homopolymeric runs.

1.5 DNA Sequencing Using Mass Spectrometry ¹⁰

In Sanger dideoxy sequencing and DNA SBS, one drawback of the use of laser-induced fluorescence detection is the overlap of fluorescence emission spectra, which usually leads to difficulty in heterozygote (namely a different version of gene on each of the two corresponding chromosomes) detection. DNA sequencing by synthesis using mass spectrometry, as an alternative approach, can overcome this problem and produce sequencing data with high resolution and accuracy. As mass resolution can be as good as one Dalton or less, the MS sequencing approach could produce sequencing data with high resolution by detecting small molecular tags rather than long DNA fragments. The high resolution allows accurate detection of single nucleotide polymorphisms (SNPs), mutations and heterozygotes.

In 2002 Prof. Ju described the DNA sequencing by mass spectrometry using mass-tagged solid phase capturable dideoxynucleotides having cleavable linker.^{10b} A mass tag for this purpose is any stable molecular moiety that can be attached to a modified nucleotide without adversely affecting its selectivity and efficiency during the polymerase reaction

and can be conveniently cleaved from the nucleotide and detected on a mass spectrometer, thus the nucleotide to which it was originally attached can be identified. In this method, a mass tag is attached on the specific site of the nucleobase via a cleavable linker, as shown in Scheme 1-4. The mass tag moiety can be cleaved chemically or photochemically. The latter is more desirable as it is fast, mild, easy to control, and can be conducted without the addition of other reagents that need subsequent purification. A mass tag-labeled nucleotide, after being incorporated into a DNA strand, can be cleaved by photolysis to generate a DNA strand residue and a tag moiety. The latter can be detected by electrospray ionization (ESI) mass spectrometry with high resolution and accuracy.



Detected by ESI Mass Spectrometry

Scheme 1-4. Schematic illustration of photocleavable mass-tagged nucleotides for DNA sequencing using mass spectrometry.

When four modified nucleotides each labeled with a reversible capping group on the 3' hydroxyl and a mass tag of unique molecular weight on the nucleobase are prepared, they can be used for DNA sequencing and analysis. After the nucleotide is incorporated into a DNA strand in the DNA polymerase reaction, the identity of the nucleotide will be confirmed precisely by observing the exact mass data by MS. Therefore the sequence of the DNA strand can be obtained accordingly. Later, the capping group on the 3' hydroxyl can be removed by a chemical method so that the next cycle of nucleotide incorporation will be initiated.

1.6 DNA Sequencing in Nanopore

DNA sequencing via nanopore (a small ion channel) is another approach to sequence a DNA strand. In the past few years, ion channels have been explored to detect individual DNA or RNA strands, with the nanopore approach being a possible candidate for very high rate DNA sequencing and analysis.

In 1996, Kasianowicz *et al* first demonstrated that the α -hemolysin channel could be used to detect nucleic acids at the single molecule level.¹¹ The monomeric polypeptide self-assembles in a lipid bilayer membrane to form a heptameric pore, with a 2.6 nm-diameter vestibule and 1.5 nm-diameter limiting aperture (the narrowest point of the pore).¹²⁻¹⁴ In an aqueous ionic salt solution, the pore formed by the α -hemolysin channel conducts a sufficiently strong and steady ionic current when an appropriate electronic voltage is applied across the membrane. The limiting aperture of the nanopore allows only linear single-stranded but not double-stranded nucleic acid molecules (diameter \approx 2.0 nm) to pass through. The polyanionic nucleic acids are driven through the pore by the applied electric field, which as a result blocks or reduces the ionic current that would be otherwise unimpeded. This process of passage generates an electronic signature (Figure 1-4).^{14,15} A particular nucleic acid molecule, when it enters and passes through the nanopore, will generate a characteristic signature that distinguishes it from other nucleic acid molecules. The duration of the blockade is proportional to the length of the nucleic acid, and its signal strength is relative to the steric and electronic properties of the nucleotides sequenced, namely the identity of the four nucleobases. Thus a specific event diagram, which is the plot of translocation time versus blockade current, is obtained and used to distinguish the length and the composition of polynucleotides by single-channel recording techniques.^{14,15}



Figure 1-4. A nucleic acid stretch passes through a nanopore by the applied electric field (left) and electronic signatures are generated as a result (right).

Although the nanopore approach shows promise as a DNA detection method, the more demanding goal of accurate base-to-base sequencing has not yet been obtained, partly due to the similar molecular sizes of purines (A and G) or pyrimidines (C and T). Further research effort is needed to make this approach more feasible.

Other methods to analyze DNA strands are also available nowadays, such as mass spectrometry.¹⁶ Matrix-assisted laser desorption / ionization time-of-flight mass spectrometry (MALDI-TOF MS) is an efficient tool to analyze DNA fragments conveniently with accuracy based on the different molecular weights of polynucleotide fragments with different strand lengths and base composition; of course, the range of sizes that can be detected in the mass spectrometer limits the throughput of this method.

In this part of the thesis the design and synthesis of fluorescent nucleotides as reversible terminators will be described and the application of such molecules in DNA sequencing by synthesis will also be discussed.

References

- [1] a) Roses, A.D. Nature 2000, 405, 857-865. b) Collins, F. S.; Green, E. D.; Guttmacher, A. E.; Guyer, M. S. Nature 2003, 422, 835-847. c) Friedman, L. S.; Ostermeyer, E. A.; Szabo, C. I.; Dowd, P.; Lynch, E. D.; Rowell, S. E.; King, M. C. Nat. Genet. 1994, 8, 399-404.
- [2] a) Sanger, F.; Nicklen, S.; Coulson, A. R. Proc. Natl. Acad. Sci. USA 1977, 74(12), 5463-5467.
 b) Sanger, F. Science 1981, 214(4526), 1205-1210. c) Sanger, F. Angew. Chem. 1981, 93(11), 937-944.
- [3] Hyman, E. D. Anal. Biochem. 1988, 174, 423-436.
- [4] a) Hardin, S. H.; Briggs, J. M.; Tu, S.; Gao, X.; Willson, R. WO Patent 2002 0204680. b)
 Levene, M. J.; Korlach, J.; Turner, S. W.; Foquet, M.; Graighead, H. G.; Webb, W. W. Science 2003, 299(5607), 682-686.
- [5] a) Welch, M. B.; Burgess, K. Nucleosides Nucleotides 1999, 18(2), 197-201. b) Metzker, M. L.; Raghavachari, R.; Richards, S.; Jacutin, S. E.; Civitello, A.; Burgess, K.; Gibbs, R. A. Nucleic Acids Res. 1994, 22(20), 4259-4267.
- [6] Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J. Science 1994, 264(5167), 1891-1903.
- [7] a) Li, Z.; Bai, X.; Ruparel, H.; Kim, S.; Turro, N. J.; Ju, J. Proc. Natl. Acad. Sci. USA 2003, 100(2), 414-419. b) Seo, T. S.; Bai, X.; Kim, D. H.; Meng, Q.; Shi, S.; Ruparel, H.; Li, Z.; Turro, N. J.; Ju, J. Proc. Natl. Acad. Sci. USA 2005, 102(17), 5926-5931. c) Ruparel, H.; Bi, L.; Li, Z.; Bai, X.; Kim, D. H.; Turro, N. J.; Ju, J. Proc. Natl. Acad. Sci. USA 2005, 102(17), 5932-5937.
- [8] For selected reviews on DNA pyrosequencing, see: a) Ahmadian, A.; Ehn, M.; Hober, S. Clinica Chimica Acta 2006, 363, 83-94. b) Ronaghi, M.; Karamohamed, S.; Pettersson, B.; Uhlén, M.; Nyrén, P. Anal. Biochem. 1993, 208(1), 171-175. c) Nyrén, P.; Pettersson, B.; Uhlén, M. Anal. Biochem. 1996, 242(1), 84-89. d) Ronaghi, M.; Uhlen, M.; Nyren, P. Science 1998, 281(5375), 363-365.
- [9] Margulies, M. et al Nature 2005, 437(7057), 376-380.

- [10] a) Liu, Z.; Bai, X.; Meng, Q.; Ruparel, H.; Turro, N. J.; Ju, J. *Nucleic Acids Res.*, submitted.
 b) Ju, J.; Edwards, J. R.; Li, Z. WO 2002079.
- [11] Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. Proc. Natl. Acad. Sci. USA 1996, 93(24), 13770-13773.
- [12] Meller, A.; Nivon, L.; Brandin, E.; Golovchenko, J.; Branton, D. Proc. Natl. Acad. Sci. USA
 2000, 97(3), 1079-1084.
- [13] Akeson, M.; Branton, D.; Kasianowicz, J. J.; Brandin, E.; Deamer, D. W. Biophys. J. 1999, 77(6), 3227-3233.
- [14] Deamer, D. W.; Branton, D. Acc. Chem. Res. 2002, 35(10), 817-825.
- [15] Vercoutere, W.; Winters-Hilt, S.; Olsen, H.; Deamer, D.; Haussler, D.; Akeson, M. Nature Biotech. 2001, 19(3), 248-252.
- [16] Edwards, J. R.; Ruparel, H.; Ju, J. Mutation Res. 2005, 573(1-2), 3-12.

Chapter 2. Design and Synthesis of Modified Fluorescent Nucleotides for DNA Sequencing by Synthesis

2.1 Background

As discussed in Chapter 1, DNA sequencing by synthesis using laser-induced fluorescence detection is an efficient approach to overcome some of the drawbacks of the current sequencing technique. With this approach no gel electrophoresis for DNA strand separation is needed and therefore the sequencing can be finished rapidly with reduced cost and improved throughput.

In 2002 Prof. Ju described his sequencing technique using dye-labeled nucleotides with 3'-hydroxyl groups protected by a small labile moiety.¹ It was proposed that only nucleotides with small capping group (such as allyl group) on the 3'-hydroxyl could be recognized by DNA polymerase enzymes.

H. Pelletier *et al* established a molecular model for the polymerase reaction based on the crystal structure of complex of rat DNA polymerase, a DNA template-primer and a 2',3'-dideoxycytidine triphosphate (ddCTP) molecule, in which the 3' position of ddCTP is very close to the amino acid residues of the polymerase (Figure 2-1).² This results in the high sensitivity of DNA polymerase to modifications on the 3' end of nucleoside

triphosphates. Only nucleoside triphosphate molecules with very small capping groups on their 3' end may be recognized by the polymerase enzyme. On the other hand, there is ample space between the 5 position of ddCTP and other groups in the complex. That makes possible the modification with even bulky groups on the 5 position of pyrimidines (C or U) or by extension, the 7 position of purines (A or G).



Figure 2-1. A transition state complex of DNA template, primer, DNA polymerase and a 2',3'-dideoxycytidine triphosphate (ddCTP) in the DNA polymerase reaction established by H. Pelletier *et al.* The left figure illustrates the incorporation process of ddCTP onto the 3' end of the primer catalyzed by Mg^{2+} and polymerase active residues. The right figure shows the steric interaction of polymerase active residues with ddCTP in the presence of primer and template.²

A series of 3' modified thymidine-5'-triphosphates were tested in polymerase reactions using different commercially available polymerase enzymes by my colleagues in Prof. Ju's group, as shown in Table 2-1. It can be seen that methyl, the smallest capping group, as well as D4T, can be recognized by quite a few polymerase enzymes, while the ethyl group is recognized by only one enzyme (HIV RT). The propyl cap was rejected by all the enzymes tested. These results further confirmed Pelletier's incorporation model of the polymerase reaction: only nucleoside triphosphate molecules with small enough capping groups on the 3' end tend to be recognized by polymerase enzymes. Although the methyl



Table 2-1. Incorporation analysis of modified nucleotides by DNA polymerases.*

R Enzyme	—CH ₃	-CH ₂ CH ₃	-CH ₂ CH=CH ₂	-CH ₂ CH ₂ CH ₃	D4T
Sequenase®	-		-	-	+
Therminator	+	-	-		+
Vent _D [®] (exo-)	-	-	-	-	+
M-MuLV-RT	+	-	-	-	
AMV-RT	-	-	-		
RAV 2RT	+	-	-	-	
HIV RT	+	+	-	-	-
Klenow Fragment (exo-)	-	-	-	-	
Deep Vent _D [®] (exo-)	-		-		
T7 DNA polymerase	-		-		
9° N _m DNA polymerase			+		

* +: recognized by polymerase; -: not recognized by polymerase; blank : not tested.

* The DNA polymerase enzymes were purchased from New England BioLabs Inc.
group is recognized by more enzymes, its removal usually requires harsh reaction conditions, not suitable for a DNA sequencing system. It is equally problematic to transform D4T into a nucleotide with a free 3'-hydroxyl group. Fortunately, we have found that the allyl group can be recognized by 9° N_m polymerase and it is relatively easy to cleave under mild conditions. This makes it a very important potential capping group for our designed fluorescent nucleotides to be used in DNA SBS.

In our molecular design of fluorescent nucleotides, the fluorescent dye is attached on the nucleobase via a linker molecule so that it can be removed conveniently after being detected to permit the next cycle of incorporation and detection. In this way an ideal fluorescent dye-labeled nucleotide molecule is constructed by attaching a fluorescent dye with a unique emission wavelength to a specific site on the nucleobase via a cleavable linker and a small capping group to the 3'-hydroxyl so that the modified nucleotide can be recognized by a polymerase, as shown in Scheme 2-1. With these four modified fluorescent nucleotides, an entire sequencing cycle, which includes incorporation of a nucleotide by polymerase, fluorescence detection, cleavage of the fluorophore and removal of the 3' capping group, will hopefully sequence an unknown DNA strand efficiently and accurately.

Prof. Ju described the design of a massively parallel DNA sequencing chip system, based on the fluorescent nucleotide design above (Scheme 2-2).¹ In this system, a chip is



Scheme 2-1. Designed molecular model of fluorescent nucleotide and proposed sequencing cycle using these modified nucleotides.

constructed with a DNA template immobilized via its 5' end, which is able to self prime to initiate the polymerase reaction. With a free hydroxyl group on the 3' end, the designed fluorescent nucleotides can be incorporated into the growing primer strand on the chip during the polymerase reaction. When four fluorescent nucleotide molecules are added together, only the one complementary to the corresponding nucleotide on the template can be incorporated. As the 3' hydroxyl group of the designed fluorescent nucleotides is blocked by the capping group, the polymerase reaction will stop before the next cycle of incorporation. This will allow the detection of the incorporated nucleotide by the unique emission wavelength of the fluorophore on each spot of the chip. After that, the



Scheme 2-2. Massively parallel DNA sequencing chip using modified fluorescent nucleotides.¹

fluorescent dye moiety will be removed by photocleavage and the 3' capping group will also be removed to regenerate the free 3' hydroxyl group. At that point, the polymerase reaction will be reinitiated and the next nucleotide incorporated. With this approach, although only one base will be identified in each sequencing cycle, the parallel processing capacity of the chip can sequence a large number of bases at the same time. The development of current surface techniques will further enhance the potential of this approach as a powerful tool for DNA sequencing.

In 2005 Prof. Ju reported four-color DNA sequencing by synthesis on a chip using four photocleavable fluorescent nucleotides (Figure 2-2).³ In this paper a photocleavable 2-nitrobenzyl linker was used to attach a fluorescent dye to the 5 position of pyrimidines or 7 position of purines. These nucleotides could be accurately incorporated into a growing DNA strand in solution or on a chip in the presence of Thermo Sequenase. The fluorescent dye moiety could be cleaved efficiently by laser irradiation (\approx 355 nm). Each SBS cycle included incorporation of the fluorescent nucleotide, fluorophore detection and photocleavage of the fluorescent dye. The entire sequencing cycle was repeated to successfully identify 12 continuous bases on the DNA template.

In a companion paper, Prof. Ju reported the synthesis of 3'-O-allyl-dUTP-PC-Bodipy-FL-510 and its application as a reversible terminator for DNA sequencing by synthesis (Scheme 2-3).⁴ In this molecule, the 3'-hydroxyl was blocked by an allyl group and the



Figure 2-2. Structures of the four photocleavable fluorescent nucleotides and photocleavage of the 2-nitrobenzyl linker by near-UV irradiation.³

fluorophore Bodipy-FL-510 was attached to the 5 position of uracil via a photocleavable 2-nitrobenzyl linker. As the capping group is a small allyl, 3'-O-allyl-dUTP-PC-Bodipy-FL-510 could be recognized by 9° N_m DNA polymerase and was incorporated into a growing DNA strand in the polymerase reaction. After that, the fluorophore moiety was cleaved efficiently by near-UV laser irradiation. Next, the 3' allyl capping group was

Part II

removed in aqueous palladium (II)-phosphine solution to regenerate the free 3'-hydroxyl group with high efficiency. The polymerase reaction was thereby reinitiated and a photocleavable fluorescent dGTP-PC-Bodipy-FL-510 was added onto the 3'-hydroxyl of the modified uridine nucleotide. This set of reactions has established the feasibility of using this 3'-O-allyl- dUTP-PC-Bodipy-FL-510 as a reversible terminator for DNA SBS.



Scheme 2-3. An entire sequencing cycle using photocleavable fluorescent 3'-O-allyl-dUTP-PC-Bodipy-FL-510 nucleotide as a reversible terminator for DNA SBS.⁴

2.2 Design of Modified Fluorescent Nucleotides for DNA SBS

Based on the results introduced above, my work in this part of the thesis started with the design and synthesis of a full set of four photocleavable fluorescent nucleotides (3'-O-allyl-dNTP-PC-Dyes) so that a DNA strand composed of different bases hopefully could be identified successfully using these four molecules. Therefore four 3'-O-allyldNTP-PC-Dye nucleotides were designed, as shown in Scheme 2-4, in which a photocleavable linker was used to attach the fluorescent dyes to the pyrimidine or purine base of 3' allyl modified nucleoside triphosphates 1-4 to construct the target 3'-O-allyldNTP-PC-Dye molecules. Presumably these molecules would be incorporated into a growing DNA strand in the polymerase reaction in the presence of 9° N_m polymerase. After fluorescence detection, the fluorophore moiety would be removed by laser irradiation. Later, the allyl capping group would be removed in aqueous palladium (II)-phosphine solution to regenerate the free 3'-hydroxyl group. Finally, the polymerase reaction would be reinitiated and the next nucleotide incorporated, and the process repeated for multiple cycles. In this way these four modified fluorescent nucleotide molecules could be used together to sequence a DNA strand composed of different bases.

The fluorescent dyes used as reporters to label the nucleotide molecules for DNA SBS should be stable under the various conditions that are required during the sequencing processes such as photocleavage and deallylation. They can not interfere with each other



Scheme 2-4. Design of modified photocleavable fluorescent nucleotides for DNA SBS.

98

spectrally. That means the emission wavelengths between any two of them should be well enough separated from each other in the visible light range so that the fluorescence imager can recognize them unambiguously. The four fluorescent dyes selected to fulfill these criteria were Bodipy-FL-510, R6G, ROX and Bodipy-650 (Bodipy: 4,4-difluoro-4bora-3 α , 4 α -diaza-*s*-indacene; R6G: 6-carboxyrhodamine-6G; ROX: 6-carboxy-Xrhodamine). These fluorescent dyes have distinct emission wavelengths of 510, 550, 602 and 650 nm respectively. Their stability and spectral properties meet our requirements in DNA SBS based on our past work.³

Since an entire sequencing cycle includes incorporation of fluorescent nucleotide, fluorophore detection, cleavage of fluorescent dye and removal of 3' capping group, the overall efficiency of these steps in each sequencing cycle will determine the speed and potential read length of SBS, as can be seen in Table 2-2.⁵ In sequencing a long DNA strand, there will be numerous repetitions of the above steps. Therefore even minor losses in efficiency at each step will decrease the overall possible read length in SBS. Any improvement in efficiency for each step will obviously contribute to the overall SBS capacity.

Cycle efficiency (%)90.095.098.099.099.599.9Read length (bases)7143569149693

 Table 2-2.
 Influence of cycle efficiency on sequencing read length.⁵

The use of the disulfide group as a chemically cleavable linker to attach a fluorophore to a nucleotide molecule for DNA SBS and 2-mercaptoethanol to cleave the fluorophore from this nucleotide has been reported.⁶ However, its further application in DNA SBS was restricted as disulfide bonds can be reversed and destabilized under certain conditions.⁷ Since the allyl group has been used as a capping group and it could be recognized by 9° N_m polymerase and removed by palladium (II)-phosphine efficiently, a new chemically cleavable linker was synthesized in the Ju laboratory and used to link the fluorophore Bodipy-FL-510 to the 7 position of 3' allyl modified nucleoside triphosphate 2. This new fluorescent nucleotide 3'-O-allyl-dGTP-allyl-Bodipy-FL-510 has been applied in DNA SBS as a reversible terminator successfully (Scheme 2-5).⁸ After being incorporated into a DNA strand, its fluorophore moiety and 3' allyl capping group were removed together by palladium (II)-phosphine, efficiently and in one step. Thus, the original two steps of fluorophore cleavage and 3' capping group removal were combined in a single step, improving the efficiency of the entire sequencing cycle, and greatly increasing the speed and potential sequencing read length of SBS.

Based on this encouraging result, a full set of four fluorescent 3'-O-allyl-dNTP-allyl-Dye nucleotides was designed, as shown in Scheme 2-6. In these molecules, another new chemically cleavable allyl linker 5 would be used to attach the fluorescent dye to the pyrimidine or purine base of nucleoside triphosphates 1-4 to construct the target 3'-O-allyl-dNTP-allyl-Dye molecules. With a terminal carbon-carbon double bond in 5,



Scheme 2-5. Application of a chemically cleavable fluorescent 3'-O-allyl-dGTP-allyl-Bodipy-FL-510 nucleotide as a reversible terminator for DNA SBS.⁸

there is no harassing syn / anti problem that appeared in the case of the linear allyl linker. The allyl linker **5** can be prepared from 2-methylenepropane-1,3-diol conveniently. Presumably these 3'-*O*-allyl-dNTP-allyl-Dye molecules could be incorporated into a growing DNA strand in the polymerase reaction in the presence of 9° N_m polymerase. After fluorophore detection, it should be possible to remove efficiently both the fluorescent dye moiety and 3' allyl capping group using palladium (II)-phosphine and regenerate the free 3'-hydroxyl group. Thus, the polymerase reaction of the next cycle would be initiated and the next nucleotide identified. In this way these four fluorescent

nucleotides could be used together to sequence a DNA strand composed of different bases with improved sequencing efficiency and speed.



Scheme 2-6. Design of modified chemically cleavable fluorescent nucleotides for DNA SBS.

2.3 Synthesis of 3' Allyl Modified Nucleoside Triphosphates (1, 2, 3 and 4)

2.3.1 Synthesis of 3'-O-allyl-7-deaza-7-(3-aminoprop-1-ynyl)-2'-deoxyadenosine-5'triphosphate (3'-O-allyl-dATP-NH₂) 1

To the best of our knowledge, the synthesis of 1 was only described by Milton *et al* in their patent application, in which 7-deaza-7-iodo-2'-deoxyadenosine 6 was used as starting material (Scheme 2-7).⁹ The 5'-hydroxyl group of 6 was first protected by a *tert*-



Scheme 2-7. Milton's synthetic strategy of 1.⁹

butyldiphenylsilyl group, and then the 6-amino group on the nucleobase was protected in the form of an amidine by N,N-dimethylformamide dimethyl acetal to afford **8**. **8** was then transformed to **9** with an an equimolar amount of allyl bromide in the presence of sodium hydride in tetrahydrofuran. Next, **9** was deprotected to produce **11** via two steps. The aminoalkynyl side chain was introduced onto the nucleobase via the palladiumcatalyzed Sonogashira cross-coupling reaction with the terminal alkyne to furnish triphosphate precursor **12**. Finally **12** was transformed to 5'-triphosphate **1** by the standard triphosphate preparation procedure.

6, a commercially unavailable compound, could be prepared from 6-chloro-7deazapurine 13 and 3,5-di-O-(p-toluoyl)-2-deoxy-D-ribofuranosyl chloride 15 either by regiospecific iodination first and then stereoselective glycosylation and aminolysis (route A),^{10,11} or by



Scheme 2-8. Synthetic routes of 6 from 13.¹⁰⁻¹⁴

glycosylation first and then iodination and aminolysis (route B), as shown in Scheme 2-8.¹²⁻¹⁴ The overall yield of **12** from **13** is 12-14% in eight or nine steps.

The key step in the synthesis of triphosphate **1** is the allylation of the 3'-hydroxyl group, which requires prior protection of the more active 5'-hydroxyl group as well as the 6-amino group on the nucleobase due to its enhanced reactivity and less steric hindrance.

Our synthesis of 1 started from 13 via route A by regiospecific iodination with Niodosuccinimide first to afford a single 7-iodo isomer 14 in 82% yield (Scheme 2-9),
following the reported procedure.¹⁰ Then stereoselective nucleobase-anion glycosylation



Scheme 2-9. Our synthesis of 1 from 13.

of 14 with 15^{11} followed by one-pot aminolysis at room temperature furnished single product 18 (71% yield). Here the 6-chloro substituent was preserved intentionally in the process of aminolysis, so that there was no necessity to protect / deprotect the 6-amino group for the coming step of allylation on the 3'-hydroxyl group.

The 1'-configuration of **18** is identified as a β -anomer based on Robins' rule for confirming the absolute anomeric configuration of 2'-deoxy-D-ribofuranosylpurines via ¹H NMR spectra: the 1'-H of β -anomers is observed to be a 'doublet-doublet' or 'pseudo-triplet' with a peak width $(J_1 + J_2)$ of 13.7±0.5 Hz, while the 1'-H of α -anomers is observed to be a 'doublet-doublet' with a peak width of 10.4±0.4 Hz.¹⁵ This rather empirical method of anomeric configuration assignment appears to be general for 2'-deoxy-D-ribofuranosyl purines despite of variations in the structure of the nucleobases, or the temperature and solvents used in NMR experiments.

According to our ¹H NMR experiment (400 MHz, CD₃OD), the 1'-H of **18** is detected as a 'doublet-doublet' at 6.72 ppm with a peak width $(J_1 + J_2)$ of 13.8 Hz, which confirms that **18** is a β -anomer. Our ¹H NMR experiments in Figure 2-3 also indicate that all β -anomers of 2'-deoxy-D-ribofuranosylpurines, from either commercial products or our synthetic intermediates, display a peak width between 13.6 and 14.5 Hz. On the other hand, we find that α -anomers of 2'-deoxy-D-ribofuranosylpurines, as shown in Figure 2-4, show a peak width between 9.6 and 10.8 Hz, as reported by some references.^{15,16} All



these results support Robins' empirical method.





Figure 2-4. Coupling constants of selected α -anomers of 2'-deoxy-D-ribofuranosylpurines.^{15,16}

The 5'-hydroxyl group of 18 was then protected with a *tert*-butyldimethylsilyl group to

produce **19** following the reported procedure.¹⁷ The first attempted allylation of **19** with an an equimolar amount of allyl bromide in the presence of sodium hydride in tetrahydrofuran (Milton's method) failed to produce any allylation product, nor was any starting material recovered. Probably **19** was decomposed by sodium hydride during this process. Therefore, an alternative allylation method, which employed an excess of allyl bromide in heterogeneous methylene chloride and 40% aqueous sodium hydroxide solution in the presence of a catalytic amount of tetrabutylammonium bromide as phase-transfer catalyst, was developed and proved to be successful, giving 95% yield of pure 3'-*O*-allyl product **20**.

Next, **11** was generated in 84% yield via a one-pot procedure of deprotection of the 5' *tert*-butyldimethylsilyl group by tetrabutylammonium fluoride¹⁷ and nucleophilic aromatic substitution of the 6-chloro group by an amino group using methanolic ammonia in an autoclave¹⁸. Finally, a Sonogashira cross-coupling reaction of **11** with the terminal alkyne catalyzed by tetrakis(triphenylphosphine)palladium(0) and copper(I) iodide furnished triphosphate precursor **12**.¹⁹

In summary, we have developed a new and concise synthetic method to prepare 12 from 13 in only six steps with an overall yield of 36%. 12 was then transformed to the 5'-triphosphate 1 according to the usual triphosphate preparation procedure.²⁰

2.3.2 Synthesis of 3'-*O*-allyl-7-deaza-7-(3-aminoprop-1-ynyl)-2'-deoxyguanosine-5'triphosphate (3'-*O*-allyl-dGTP-NH₂) **2**²¹

The synthesis of 2 was only reported by Milton *et al* in their patent application, in which 7-deaza-7-iodo-2'-deoxyguanosine 21 was used as starting material (Scheme 2-10).⁹ The 5'-hydroxyl of 21 was first protected by a *tert*-butyldimethylsilyl group to afford 22. 22 was then transformed to 23 via allylation with an an equimolar amount of allyl bromide in the presence of sodium hydride and then cleavage of the 5' *tert*-butyldimethylsilyl group. Next, the aminoalkynyl side chain was introduced to the nucleobase via the palladium-catalyzed Sonogashira cross-coupling reaction with the terminal alkyne to give triphosphate precursor 24. Finally 24 was transformed to 5'-triphosphate 2 by a standard



Scheme 2-10. Milton's synthetic strategy of 2.9

triphosphate preparation procedure.

21, a commercially unavailable compound, could be prepared in four steps from 2-amino-6-methoxy-9-(β -D-2'-deoxyribofuranosyl)-7-deazapurine 25 (Scheme 2-11), in which the reactive NH of 7-deazaguanine is masked by a 6-methoxy group. The 2-amino group of 25 was first transformed to formamide by *N*,*N*-dimethylformamide diethyl acetal to furnish 26.²² Then both the 3'- and 5'-hydroxyl groups of 26 were acylated with an excess of isobutyric anhydride. With its 2-amino and 3'-, and 5'-hydroxyl groups fully protected, 27 was regiospecifically iodinated with *N*-iodosuccinimide to give 7-iodo product 28. Then cleavage of the 6-methoxy substituent along with the formyl and two isobutyryl groups of 28 afforded 21 in one pot by hydrolysis in refluxing aqueous sodium hydroxide solution.²³ The overall yield of 24 from 25 is 8.5% in eight steps.



Scheme 2-11. Synthetic route of 21 from 25.^{22,23}

According to Seela's explanation, the iodination site of 7-deazaguanine is strongly affected by the substituent on the purine.²³ For example, the 2-amino group on the 7-deazaguanine promotes an electrophilic attack on the 8-position, as the mesomeric σ -complex formed in the transition state is stabilized by the electron-donating effect of the 2-amino group (Figure 2-5, A). On the other hand, a 2-acylamino is less electron-donating. Therefore, an electrophilic attack on the 7-position is preferred, as the mesomeric σ -complex formed in the transition state is stabilized by the electron-donating effect of the pyrrole nitrogen (Figure 2-5, B).



Figure 2-5. Substituent effect on the iodination site of 7-deazaguanine.²³

Another important method to prepare 21, as shown in Scheme 2-12, started from 2-methylsulfanyl-7H-pyrrolo[2,3-d]pyrimidin-4-ol 29 in a few more steps. Chlorination of 29 with phosphorus oxychloride in the presence of N,N-dimethylaniline first generated

30.²⁴ Then stereoselective glycosylation of 30 with 15 furnished 31^{13} and 31 was regiospecifically iodinated by *N*-iodosuccinimide to afford 7-iodo product 32. Later, hydrolysis of 32 in the solution of *syn*-2-pyridine aldoxime and 1,1,3,3-tetramethylguanidine in dioxane and DMF furnished 33. Next, the methylthio group of 33 was converted to an amino group via oxidation by 3-chloroperoxybenzoic acid first, and then nucleophilic aromatic substitution by the amino group. 21 was finally obtained after the two toluoyl groups of 34 were cleaved with methanolic ammonia.^{13,14,25} The overall yield of 21 from 29 is 30% in six steps and therefore the overall yield of 24 from 29 is 4.1% in ten steps.



Scheme 2-12. Alternative synthetic route of 21 from 29.^{13,14,24,25}

The key steps in the synthesis of triphosphate 2 include: 1) regiospecific iodination of the 7-deazapurine at the 7-position and the protection of sensitive groups; 2) successful

allylation on the 3'-hydroxyl group; and 3) efficient cleavage of the 6-methoxy group to generate the corresponding 7-deazaguanine.

Our first attempted synthesis of 2 started from 25, which was first protected by isobutyryl chloride to form 35 quantitatively (Scheme 2-13).²² 35 was further iodinated with N-iodosuccinimide to afford a single 7-iodo product 36 (84% yield), as the 2-acylamino group on the nucleobase promotes the formation of the 7-substituted product on the basis of the discussion above.



Scheme 2-13. Our first attempted synthesis of 37 from 25.

Seela developed a simple but effective method to confirm the iodination position on the 7-deazapurine moiety by ¹³C NMR spectra: the significant upfield C(7) chemical shift of 7-deazapurine nucleosides after being 7-iodinated ($\Delta \delta$ = 49±3 ppm), with the iodinated

C(7) usually appearing between 49-57 ppm.²⁶ Our ¹³C NMR spectra revealed that the C(7) of **35** shows up at 100.7 ppm while in **36** it moves to 51.8 ppm. This result clearly indicates that **36** is a 7-iodo compound. This rule also applies to **18**, whose 52.6 ppm C(7) peak indicates that it too is a 7-iodo compound.

Next, **21** was obtained in 48% yield via cleavage of the 6-methoxy substituent and the three isobutyryl groups of **36** in one step by hydrolysis in refluxing aqueous sodium hydroxide solution. The 5'-hydroxyl of **21** was then protected by a *tert*-butyldimethylsilyl group to produce **22** following the established procedure.¹⁷ But unfortunately, our first attempted allylation of **22** with an equimolar amount of allyl bromide in the presence of sodium hydride in tetrafuran (Milton's method)⁹ failed to produce any allylation product, and only starting material was recovered. Other methods such as using an excess of allyl bromide in methylene chloride and 40% aqueous sodium hydroxide solution in the presence of a catalytic amount of tetrabutylammonium bromide gave none of the desired product, nor was any starting material recovered.

An alternative route to circumvent this difficulty was also investigated, as shown in Scheme 2-14. Two of the isobutyryl groups in 36 were cleaved by ammonia to produce 38 following a reported procedure.²² Here the isobutyryl group on the 2-amino was preserved due to its stability while the other two on the glycose were removed. The 5'-hydroxyl of 38 was then protected by a *tert*-butyldimethylsilyl group to give 39

according to the established procedure.¹⁷ To our surprise, allylation of **39** with an excess of allyl bromide in methylene chloride and 40% aqueous sodium hydroxide solution in the presence of catalytic amount of tetrabutylammonium bromide furnished none of our desired product, but instead N^2 -allyl compound **40** was produced with minor N^2 ,*O*-diallyl product. Other allylation methods gave similar results, which indicates that allylation on the 2-amino group is even faster than that on the 3'-hydroxyl. This is probably because the 2-amino group, after being acylated, is more acidic, and therefore it becomes more nucleophilic than the 3'-hydroxyl group upon allyl bromide reaction in the presence of base. This route proved to be inaccessible to the final target **2**.



Scheme 2-14. Attempted synthetic route toward 2 from 36.

Another route to 2 had to be investigated, as shown in Scheme 2-15. All three isobutyryl groups of 36 were cleaved to afford 41 in 94% yield in refluxing methanolic sodium

methoxide, since methanolic ammonia proved to be ineffective at the same temperature. Then the 5'-hydroxyl of **41** was again protected by a *tert*-butyldimethylsilyl group to give **42** in 88% yield by the established procedure.¹⁷ The 3'-hydroxyl of **42** was next allylated with an excess of allyl bromide in heterogeneous methylene chloride and 40% aqueous sodium hydroxide solution in the presence of a catalytic amount of tetrabutylammonium bromide to furnish **43** as single product (92% yield). Here it was unnecessary to protect the 2-amino of **42** for the allylation of 3'-hydroxyl group. This reaction is so selective and efficient that no N^2 -allyl product was observed at all, although this method is limited to only some nucleobase analogues.



Scheme 2-15. Another attempted synthetic route toward 2 from 36.

Next, the *tert*-butyldimethylsilyl group on the 5'-position of **43** was cleaved to produce **44** (93% yield). Then cleavage of the 6-methoxy group of **44** in alkaline medium such as

refluxing aqueous sodium hydroxide solution²⁷ was first investigated, but it turned out to be ineffective and only starting material was recovered. Another method using chlorotrimethylsilane and sodium iodide in acetonitrile²⁸ was also performed and no desired product **45** was obtained. Seela has proved that this method could not be applied for demethylation of iodo nucleosides as partial deiodination was observed in this process.^{26c}

Therefore, the synthetic route to 2 was modified still further (Scheme 2-16). From 43, the Sonogashira cross-coupling reaction with the terminal alkyne was performed first to produce 46 in 94% yield in the presence of catalytic tetrakis(triphenylphosphine)-palladium(0) and copper(I) iodide.¹⁹ Next, a one-pot procedure of cleavage of the 6-methoxy group in 46 by chlorotrimethylsilane and sodium iodide in acetonitrile²⁸ first and then deprotection of the 5' *tert*-butyldimethylsilyl group by tetrabutylammonium fluoride¹⁷ furnished the target 24 in a moderate 34% yield. It is obvious that other methods could not be used here due to the sensitivity of 46 to strong base or acid.^{27,29}



Scheme 2-16. Our final synthesis of 24 from 43.

Partial cleavage of the allyl group of 46 by iodotrimethylsilane is the probable reason for the low yield of this step.³⁰

In summary, we have developed a new and improved synthetic method to prepare 24 from 25 in only seven steps with an overall yield of 22%. Finally 24 was transformed to the corresponding 5'-triphosphate 2 according to the usual triphosphate preparation procedure.²⁰ The complete synthetic route is shown in Scheme 2-17.



Scheme 2-17. Our complete synthetic route of 2 from 25.

2.3.3 Synthesis of 3'-O-allyl-5-(3-aminoprop-1-ynyl)-2'-deoxycytidine-5'-triphosphate (3'-O-allyl-dCTP-NH₂) 3

The synthesis of **3** was also only reported by Milton *et al* in their patent application, in which 5-iodo-2'-deoxycytidine **47** was used as starting material (Scheme 2-18).⁹ The 5'-hydroxyl of **47** was first protected by a *tert*-butyldimethylsilyl group to afford **48**, which was then transformed to **49** with an equimolar amount of allyl bromide in the presence of sodium hydride. Next **50** was obtained via cleavage of the *tert*-butyldimethylsilyl group on the 5'-position. The aminoalkynyl side chain was subsequently introduced to the 5-position of cytosine via the palladium-catalyzed



Scheme 2-18. Milton's synthetic strategy of 3.⁹

Sonogashira cross-coupling reaction to give triphosphate precursor **51**. Finally **51** was transformed to 5'-triphosphate **3** by a standard triphosphate preparation procedure.

We started from 47 to prepare 5'-triphosphate 3 by following Milton's route exactly. 47 was first protected by a tert-butyldimethylsilyl group to produce 48 in 89% yield according to the established procedure.¹⁷ Then allylation of **48** with an equimolar amount of allyl bromide in the presence of sodium hydride in tetrahydrofuran produced O-allyl product 49 and N^4 -allyl product (7:1 ratio) in a total yield of 47%. Here it is unnecessary to protect the 4-amino of **48** for the allylation of the 3'-hydroxyl group, although **49** was only obtained in moderate yield and the N^4 -allyl product was also generated. Other allylation methods, such as using an excess of allyl bromide in methylene chloride and 40% aqueous sodium hydroxide solution in the presence of a catalytic amount of tetrabutylammonium bromide, only resulted in decomposition of the starting material. Reaction with an equimolar amount of allyl bromide using sodium hydroxide powder in tetrahydrofuran promoted by a catalytic amount of tris[2-(2-methoxyethoxy)ethyl]amine (TDA-1) also proved to be ineffective and only starting material was recovered. The mixture of N- and O-allyl products was used for the next step directly, since we could not separate them by column chromatography. Next, the *tert*-butyldimethylsilyl group on the 5'-position was cleaved to produce pure O-allyl product 50 in 71% yield. Later, a Sonogashira cross-coupling reaction of 50 with the terminal alkyne catalyzed by tetrakis-(triphenylphosphine)palladium(0) and copper(I) iodide furnished triphosphate precursor 51^{19} , which was transformed to the corresponding 5'-triphosphate 3 according to the standard triphosphate preparation procedure.²⁰

2.3.4 Synthesis of 3'-O-allyl-5-(3-aminoprop-1-ynyl)-2'-deoxyuridine-5'-triphosphate (3'-O-allyl-dUTP-NH₂) 4

Milton *et al* also reported the synthesis of **4** in their patent application, in which 5-iodo-2'- deoxyuridine **52** was used as starting material (Scheme 2-19).⁹ The synthetic route is similar to that of **3**.



Scheme 2-19. Milton's synthetic strategy of 4.⁹

In our previous work an alternative route was used for the synthesis of 4, which also started from 52 and used a 4,4'-dimethoxytrityl group to protect its 5'-hydroxyl (Scheme 2-20).⁴



Scheme 2-20. Synthetic strategy of 4 in our previous work.⁴

In both routes, preparation of **56** from **55** via the Sonogashira cross-coupling reaction was found to be problematic. The yield of this step was low and the reaction was very difficult to duplicate, even if Milton's procedure was followed exactly and carefully, although similar cross-coupling reactions, such as preparation of **12** from **11**, **46** from **43** and **51** from **50**, all gave good results and high yields. Most of the starting material **55** was decomposed in the reaction and there is no obvious explanation for this result. Therefore, a modified route was investigated, as shown in Scheme 2-21. **52** was first protected by a *tert*-butyldimethylsilyl group to produce **53** (90% yield).¹⁷ Then allylation of **53** with an equimolar amount of allyl bromide in the presence of sodium hydride in tetrahydrofuran

produced *O*-allyl product **54**. Here it is unnecessary to protect the N(3)-H of **53** for the allylation of the 3'-hydroxyl group, and **54** was obtained in 80% yield while no N^3 -allyl product was found. Next, with the 5'-position still protected, **54** was reacted with the terminal alkyne via the Sonogashira cross-coupling reaction catalyzed by tetrakis-(triphenylphosphine)palladium (0) and copper(I) iodide and **58** was obtained in a reasonable yield of 73%. This result could be repeated consistently. It is surprising that the 5' protection status should affect the results of the Sonogashira cross-coupling reaction so much here, since in most cases the hydroxyl group does not need to be protected. We have no exact explanation for this phenomenon. Next, the 5' *tert*-butyldimethylsilyl group of **58** was cleaved by tetrabutylammonium fluoride¹⁷ to



Scheme 2-21. Our modified synthetic route of 4 from 52.

produce **56** (49% yield). Finally **56** was transformed to the corresponding 5'-triphosphate **4** according to the typical triphosphate preparation procedure.²⁰

In summary, we have modified the previous synthetic method to prepare 4 from 52 with excellent reproducibility.

2.4 Synthesis of Cleavable Linkers

2.4.1 Synthesis of chemically cleavable allyl linker 5

The synthesis of the allyl linker **5** started from 2-methylenepropane-1,3-diol **59** (Scheme 2-22). One hydroxyl group of **59** was first protected with a trityl group in the presence of triethylamine to afford **60** in 62% yield.³¹ Then the other hydroxyl group of **60** was transformed into a bromo group employing carbon tetrabromide and triphenylphosphine in methylene chloride³² and **61** was obtained in 92% yield. Later, nucleophilic replacement of the bromo group by cyanide generated from trimethylsilyl cyanide and tetrabutylammonium fluoride in situ³³ furnished nitrile **62** in 64% yield. This reaction must be quenched in a short time (20 min) to obtain pure **62**, while longer reaction time (10 h) gave isomerized **64** as the main product. The cyano group of **62** was further reduced to an amino group by lithium aluminum hydride³⁴ to produce **63** in 65% yield. Finally the trityl group was removed by hydrochloric acid in methanol to give the allyl

linker 5 in 95% yield. This reaction must be performed in methanol instead of aqueous solution because 5 is very hydrophilic.



Scheme 2-22. Synthesis of allyl linker 5.

In summary, the allyl linker 5 was synthesized conveniently in five steps with an overall yield of 22%.

2.4.2 Synthesis of photocleavable linker 65

The photocleavable linker **65** was prepared using the procedure established by Dr. Xiaopeng Bai in our group³⁵ with minor modification. As shown in Scheme 2-23, commercial 5-methyl-2-nitrobenzoic acid **66** was first converted into the corresponding acetophenone **67** (90% yield). **67** was then brominated by *N*-bromosuccinimide to afford **68** in 43% yield. Reduction of **68** by sodium borohydride gave alcoholic **69** quantitatively.

Finally aminolysis of **69** produced photocleavable linker **65** as major product (72% yield).



Scheme 2-23. Synthesis of photocleavable linker 65.³⁵

2.5 Synthesis of Fluorescent Nucleotides (73, 77, 81, 85, 86, 87, 88, 89)

With the 3' allyl modified nucleoside triphosphates (1, 2, 3 and 4) and chemically cleavable allyl linker 5 prepared, the four chemically cleavable fluorescent nucleotides 3'-O-allyl-dCTP-allyl-Bodipy-FL-510 73, 3'-O-allyl-dUTP-allyl-R6G 77, 3'-O-allyl-dATP-allyl-ROX 81 and 3'-O-allyl-dGTP-allyl-Bodipy-650 85 were prepared using procedures similar to those reported previously by our group (Scheme 2-24).³ The allyl linker 5 was first linked to the commercial fluorescent dye *N*-hydroxysuccinimide (NHS) esters (70, 74, 78 and 82) in a solution of aqueous sodium bicarbonate and DMF to produce allyl-Dye (71, 75, 79 and 83). The allyl-Dyes were then transformed to corresponding NHS esters (72, 76, 80 and 84) by reaction with *N*.*N*²-disuccinimidyl


Scheme 2-24. Synthesis of 3'-O-allyl-dNTP-allyl-Dyes 73, 77, 81 and 85.



127

carbonate (DSC) and triethylamine in DMF. Finally the allyl-Dye NHS esters were coupled with the nucleoside triphosphates (1, 2, 3 and 4) to furnish our final target fluorescent nucleotides 73, 77, 81 and 85.



Scheme 2-25. Synthesis of 3'-O-allyl-dNTP-PC-Dyes 86-89.

The photocleavable fluorescent nucleotides 3'-O-allyl-dGTP-PC-Bodipy-FL-510 **86**, 3'-O-allyl-dUTP-PC-R6G **87**, 3'-O-allyl-dATP-PC-ROX **88** and 3'-O-allyl-dCTP-PC-Bodipy-650 **89** were also prepared using similar procedures to those already established by our group (Scheme 2-25).³ The photocleavable linker **65** was first linked to the commercial fluorescent dye NHS esters (**70**, **74**, **78** and **82**) to produce PC-Dyes. The PC-Dyes were then transformed to the corresponding NHS esters by reacting with DSC and triethylamine. Finally the PC-Dye NHS esters were coupled with nucleoside triphosphates **1-4** to afford fluorescent nucleotides **86-89**.

2.6 Confirmation of Allyl Position in the Modified Nucleoside Molecules: *O*-allyl or *N*-allyl.

Ambiguous results are obtained when trying to confirm the allyl position in the modified nucleoside molecules via ¹H NMR spectra, although it is sometimes possible using the chemical shifts of CH₂ in the allyl group, or 3'-H, or acidic protons (such as OH, NH, and NH₂). These methods are not effective all the time. Here we have established an empirical method to confirm the allyl position with ¹³C NMR spectra. This method usually gives clear and definitive results.

In the synthesis of 20 from 19, it is straightforward to confirm that 20 is an O-allyl product since this is the only possible allyl product. By comparing ¹³C NMR spectrum of

20 with 19, it is easy to conclude that the peak at 80.2 ppm belongs to the CH₂ of *O*-allyl (Figure 2-6). In the synthesis of 50 from 49, 50 is also assigned as an *O*-allyl product (a 79.6 ppm peak for the CH₂ of *O*-allyl). At the same time, a mono-allylic side product 50N was also separated, which is assigned as *N*-allyl (a 44.8 ppm peak is shown for the CH₂ of *N*-allyl). For the same reason, 43 is assigned as an *O*-allyl product (a 79.1 ppm peak for the CH₂ of *O*-allyl) while 40 is assigned as *N*-allyl (a 50.5 ppm peak for the CH₂ of *N*-allyl). Mono-allylic 91 and di-allylic 92 were also prepared and their structures also confirmed. Actually all the *O*-allyl compounds we have ever synthesized showed up as *O*-allyl CH₂ peaks between 78~81 ppm in their ¹³C NMR spectra.



Figure 2-6. ¹³C NMR chemical shifts of selected allyl compounds.

In summary, most O-allyl CH₂ peaks show up between 78~81 ppm in ¹³C NMR spectra while *N*-allyl CH₂ peaks usually show up between 43~51 ppm, whether they are attached on the exo- or endo- nitrogen of the nucleobase. This empirical method using ¹³C NMR chemical shifts to confirm the allyl position appears to be general despite of variations in the structure of the nucleotides, or the temperature and solvents used in NMR experiments.

2.7 Experimental Section

General Information. ¹H NMR spectra were recorded on a Bruker DPX-400 (400 MHz) spectrometer and are reported in ppm from CD₃OD or DMSO-d6 internal standards (3.31 or 2.50 ppm respectively). Data are reported as follows: (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublets, ddd = doublet of doublets of doublets; coupling constant(s) in Hz; integration; assignment). Proton decoupled ¹³C NMR spectra were recorded on a Bruker DPX-400 (100 MHz) spectrometer and are reported in ppm from CD₃OD, DMSO-d6, or CDCl₃ internal standards (49.0, 39.5, or 77.0 ppm respectively). Proton decoupled ³¹P NMR spectra were recorded on a Bruker DPX-300 (121.4 MHz) spectrometer without calibration. High Resolution Mass Spectra (HRMS) were obtained on a JEOL JMS HX 110A mass spectrometer. **13** and **25** were purchased from Berry & Associates. All Dye NHS esters were purchased from Molecular Probes. All other chemicals were purchased from Sigma-Aldrich.

6-Chloro-7-iodo-7-deazapurine (14). To a vigorously stirred solution of **13** (1.0 g; 6.51 mmol) in CH₂Cl₂ (55 mL) was added *N*-iodosuccinimide (1.70 g; 7.18 mmol). The reaction mixture was stirred at room temperature for 1 h, during which more precipitate appeared. The solid was filtered and recrystallized in hot methanol to afford **14** as slightly grey crystals (1.49 g; 82% yield): ¹H NMR (400 MHz, DMSO-d6) δ 12.96 (br s, 1H, N*H*), 8.59 (s, 1H, 2-H), 7.94 (s, 1H, 8-H); ¹³C NMR (100 MHz, DMSO-d6) δ 151.2, 150.4, 150.2, 133.6, 115.5, 51.7; HRMS (FAB+) calcd for C₆H₄N₃CII (M+H⁺): 279.9139, found: 279.9141.

6-Chloro-9-(β-D-2'-deoxyribofuranosyl)-7-iodo-7-deazapurine (18). To a stirred solution of 14 (707 mg; 2.53 mmol) in CH₃CN (43 mL) were added KOH powder (355 mg; 6.34 mmol) and tris[2-(2-methoxyethoxy)ethyl]amine (TDA-1) (52 μL, 0.165 mmol). The mixture was stirred at room temperature for 10 min and then 90% 3,5-di-*O*-(p-toluyl)-2-deoxy-D-ribofuranosyl chloride 15 (1.18 g; 2.73 mmol) was added. The reaction mixture was stirred vigorously at room temperature for 1 h, and the insoluble material was filtered and washed with hot acetone. The combined solution was evaporated and dissolved in 7M methanolic ammonia (86 mL). The solution was stirred at room temperature for 24 h. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (0~1:20) as the eluent to afford 19 as a white solid (711 mg; 71% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.57 (s, 1H, 2-H), 8.08 (s, 1H, 8-H), 6.72 (dd, *J* = 6.3, 7.5 Hz, 1H, 1'-H), 4.53 (m, 1H,

3'-H), 4.00 (m, 1H, 4'-H), 3.80 (dd, J = 3.6, 12.0 Hz, 1H, one of 5'-H), 3.74 (dd, J = 3.6, 12.0 Hz, 1H, one of 5'-H), 2.56-2.64 (ddd, J = 6.1, 7.5, 13.5 Hz, 1H, one of 2'-H), 2.36-2.43 (ddd, J = 3.3, 6.2, 13.5 Hz, 1H, one of 2'-H); ¹³C NMR (100 MHz, CD₃OD) δ 152.9, 151.7, 151.3, 134.7, 118.5, 89.0, 85.7, 72.6, 63.2, 52.6, 41.7; HRMS (FAB+) calcd for C₁₁H₁₂O₃N₃ClI (M+H⁺): 395.9612, found: 395.9607.

9-[\beta-D-5'-O-(tert-Butyldimethylsilyl)-2'-deoxyribofuranosyl]-6-chloro-7-iodo-7-

deazapurine (19). To a stirred solution of 18 (711 mg; 1.80 mmol) and imidazole (295 mg; 4.32 mmol) in anhydrous DMF (9.0 mL) was added tert-butyldimethylsilyl chloride (TBDMSCl) (325 mg; 2.09 mmol). The reaction mixture was stirred at room temperature for 20 h. Then most of the solvent was removed under vacuum, and the residue was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:3~2) as the eluent to afford 19 as a white solid (597 mg; 65% yield) and 18 (213 mg; 30% yield). Further recycling of 18 for this reaction gave an overall yield of 19 as 86%: ¹H NMR (400 MHz, CD₃OD) δ 8.56 (s, 1H, 2-H), 7.99 (s, 1H, 8-H), 6.73 ('t', J = 6.7 Hz, 1H, 1'-H), 4.52 (m, 1H, 3'-H), 4.02 (m, 1H, 4'-H), 3.92 (dd, J = 3.0, 11.4 Hz, 1H, one of 5'-H), 3.86 (dd, J = 3.1, 11.4 Hz, 1H, one of 5'-H), 2.47-2.55 (ddd, J = 5.8, 7.1, 13.4 Hz, 1H, one of 2'-H), 2.40-2.47 (ddd, J = 3.6, 6.3, 13.4 Hz, 1H, one of 2'-H), 0.94 (s, 9H, $C(CH_3)_3$, 0.14 (s, 3H, one of SiCH₃), 0.13 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CD₃OD) & 152.8, 151.5, 151.3, 133.8, 118.2, 88.9, 85.4, 72.5, 64.6, 52.6, 42.4, 26.7, 19.5, -4.9, -5.0; HRMS (FAB+) calcd for $C_{17}H_{26}O_3N_3ClSiI$ (M+H⁺): 510.0477, found:

510.0487.

9-[B-D-3'-O-Allvl-5'-O-(tert-butyldimethylsilyl)-2'-deoxyribofuranosyl]-6-chloro-7iodo-7-deazapurine (20). To a stirred solution of 19 (789 mg; 1.55 mmol) in CH₂Cl₂ (48 mL) were added tetrabutylammonium bromide (TBAB) (255 mg; 0.77 mmol), allyl bromide (0.69 mL, 7.74 mmol) and 40% agueous NaOH solution (24 mL). The reaction mixture was stirred at room temperature for 1h. Ethyl acetate (150 mL) was added and the organic layer was separated. The aqueous layer was extracted with ethyl acetate $(2 \times 50 \text{ mL})$. The combined organic layer was washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:6) as the eluent to afford 20 as a yellow oil (809 mg; 95% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.52 (s, 1H, 2-H), 7.94 (s, 1H, 8-H), 6.64 (dd, J = 6.1, 7.6Hz, 1H, 1'-H), 5.88-5.99 (m, 1H, $CH_2CH=CH_2$), 5.28-5.34 (dm, J = 17.3 Hz, 1H, one of $CH_2CH=CH_2$), 5.16-5.21 (dm, J = 10.4 Hz, 1H, one of $CH_2CH=CH_2$), 4.28 (m, 1H, 3'-H), 4.13 (m, 1H, 4'-H), 4.01-4.11 (m, 2H, $CH_2CH=CH_2$), 3.88 (dd, J = 3.6, 11.2 Hz, 1H, one of 5'-H), 3.80 (dd, J = 3.1, 11.3 Hz, 1H, one of 5'-H), 2.51-2.57 (ddd, J = 2.7, 6.0, 13.5 Hz, 1H, one of 2'-H), 2.42-2.50 (ddd, J = 5.7, 7.7, 13.5 Hz, 1H, one of 2'-H), 0.93 (s, 9H, C(CH₃)₃), 0.13 (s, 3H, one of SiCH₃), 0.12 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CD₃OD) & 152.8, 151.4, 151.3, 135.5, 133.6, 118.2, 117.2, 86.5, 85.6, 80.2, 71.0, 64.8, 52.8, 39.7, 26.7, 19.4, -4.8, -5.0; HRMS (FAB+) calcd for $C_{20}H_{30}O_{3}N_{3}ClSiI$ (M+H⁺):

3'-O-Allyl-7-deaza-7-iodo-2'-deoxyadenosine (11). To a stirred solution of 20 (809 mg; 1.47 mmol) in anhydrous THF (34 mL) was added 1 M TBAF in THF solution (1.62 mL; 1.62 mmol) and the reaction mixture was stirred at room temperature for 1 h. The solvent was evaporated and the residue was dissolved in 7M methanolic ammonia (24 mL). The solution was stirred in an autoclave at 115-120 °C for 17 h. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:20) as the eluent to afford 11 as a white solid (514 mg; 84% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.08 (s, 1H, 2-H), 7.56 (s, 1H, 8-H), 6.45 (dd, J = 5.8, 8.6Hz, 1H, 1'-H), 5.90-6.00 (m, 1H, CH₂CH=CH₂), 5.29-5.35 (dm, J = 17.2 Hz, 1H, one of $CH_2CH=CH_2$), 5.16-5.21 (dm, J = 10.5 Hz, 1H, one of $CH_2CH=CH_2$), 4.28 (m, 1H, 3'-H), 4.12 (m, 1H, 4'-H), 4.02-4.12 (m, 2H, $CH_2CH=CH_2$), 3.78 (dd, J = 3.7, 12.1 Hz, 1H, one of 5'-H), 3.70 (dd, J = 3.6, 12.1 Hz, 1H, one of 5'-H), 2.53-2.61 (ddd, J = 5.8, 8.6, 13.6 Hz, 1H, one of 2'-H), 2.41-2.47 (ddd, J = 2.0, 5.8, 13.5 Hz, 1H, one of 2'-H); ¹³C NMR (100 MHz, CD₃OD) δ 158.5, 152.3, 150.3, 135.7, 128.8, 117.0, 105.3, 86.8, 86.4, 80.7, 71.0, 63.7, 51.3, 38.8; HRMS (FAB+) calcd for $C_{14}H_{18}O_3N_4I$ (M+H⁺): 417.0424, found: 417.0438.

3'-O-Allyl-7-deaza-7-{3-[(trifluoroacetyl)amino]prop-1-ynyl}-2'-deoxyadenosine (12). To a stirred solution of 11 (514 mg; 1.23 mmol) in anhydrous DMF 5.5 mL) were

136

added tetrakis(triphenylphosphine)palladium(0) (142 mg; 0.123 mmol) and CuI (51 mg; 0.268 mmol). The solution was stirred at room temperature for 10 min. Then *N*-propargyltrifluoroacetamide (558 mg; 3.69 mmol) and Et₃N (343 µL; 2.46 mmol) were added. The reaction mixture was stirred at room temperature for 12 h with exclusion of air and light. Most DMF was removed under vacuum and the residue was dissolved in ethyl acetate (100 mL). The solution was washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:1~0) as the eluent to afford 12 as a yellow solid (486 mg; 90% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.08 (s, 1H, 2-H), 7.60 (s, 1H, 8-H), 6.41 (dd, J = 5.8, 8.6Hz, 1H, 1'-H), 5.89-6.00 (m, 1H, CH₂CH=CH₂), 5.29-5.35 (dm, J = 17.3 Hz, 1H, one of $CH_2CH=CH_2$), 5.16-5.21 (dm, J = 10.4 Hz, 1H, one of $CH_2CH=CH_2$), 4.31 (s, 2H, C=CCH₂), 4.29 (m, 1H, 3'-H), 4.13 (m, 1H, 4'-H), 4.01-4.11 (m, 2H, CH₂CH=CH₂), 3.79 (dd, J = 3.6, 12.1 Hz, 1H, one of 5'-H), 3.71 (dd, J = 3.5, 12.1 Hz, 1H, one of 5'-H),2.54-2.62 (ddd, J = 5.8, 8.6, 13.6 Hz, 1H, one of 2'-H), 2.42-2.48 (ddd, J = 1.9, 5.8, 13.6 Hz, 1H, one of 2'-H); ¹³C NMR (100 MHz, CD₃OD) δ 158.8, 158.6 (q, J = 38 Hz, COCF₃), 152.9, 149.6, 135.6, 128.1, 117.1 (q, *J* = 284 Hz, COCF₃), 117.0, 104.5, 96.3, 87.3, 86.9, 86.8, 80.7, 77.0, 71.0, 63.8, 38.7, 31.1; HRMS (FAB+) calcd for $C_{19}H_{21}O_4N_5F_3$ (M+H⁺): 440.1546, found: 440.1544.

12 (140 mg; 0.319mmol) and proton sponge (83.6 mg; 0.386 mmol) were dried in a vacuum desiccator over P_2O_5 overnight before dissolved in trimethylphosphate (0.60 mL). Freshly distilled POCl₃ (35.8 µL; 0.383 mmol) was added dropwise at 0 °C and the mixture was stirred for 3 h. Then the solution of tributylammonium pyrophosphate (607 mg) and tributylamine (0.61 mL; 2.56 mmol) in anhydrous DMF (2.56 mL) was added in one portion at room temperature and the reaction was stirred for 30 min. After that triethylammonium bicarbonate solution (TEAB) (0.1 M; 16 mL) was added and the mixture was stirred for 1 h. Aqueous ammonia (28%; 16 mL) was added and the reaction mixture was stirred for 12 h. After most liquid was removed under vacuum, the residue was redissolved in water and filtered. The aqueous solution was purified by DEAE Sephadex A25 ion exchange column chromatography using aqueous TEAB solution gradient (from 0.1 M to 1.0 M) as to elute 1 as a colorless syrup: ¹H NMR (300 MHz, D_2O) δ 8.02 (s, 1H, 2-H), 7.89 (s, 1H, 8-H), 6.54 (t, J = 6.6 Hz, 1H, 1'-H), 5.89-6.02 (m, 1H, CH₂CH=CH₂), 5.30-5.39 (dm, J = 17.3 Hz, 1H, one of CH₂CH=CH₂), 5.20-5.27 (dm, J = 10.4 Hz, 1H, one of CH₂CH=CH₂), 4.48 (s, 2H, C=CCH₂), 4.35 (m, 1H, 3'-H), 4.05-4.17 (m, 4H, CH₂CH=CH₂ and 5'-H), 3.99 (m, 1H, 4'-H), 2.50-2.59 (m, 2H, 2'-H); ³¹P NMR (121.4 MHz, D₂O) δ -6.1 (d, J = 21.1 Hz, 1P, y-P), -10.8 (d, J = 18.8 Hz, 1P, α -P), -21.9 (t, J = 19.9 Hz, 1P, β -P).

6-Methoxy-2-(2-methylpropanoyl)amino-9-[3',5'-bis-O-(2-methylpropanoyl)-β-D-2'deoxyribofuranosyl]-7-deazapurine (35). To a stirred suspension of 25 (1.00 g; 3.57

mmol) in anhydrous pyridine (35 mL) was added slowly isobutyryl chloride (3.40 mL; 32.2 mmol) at 0 °C. The reaction mixture was stirred at 0 °C for 1 h. Methanol (2 mL) was added and the mixture was stirred for another 10 min. Then most solvent was removed under vacuum. Ethyl acetate (200 mL) and saturated aqueous NaHCO₃ (50 mL) were added. The organic layer was separated and washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using ethyl acetate-hexane $(1:3\sim2)$ as the eluent to afford 35 as a white foam (1.75 g); 99% yield): ¹H NMR (400 MHz, CD₃OD) δ 7.28 (d, J = 3.7 Hz, 1H, 8-H), 6.66 (dd, J = 5.9, 8.6 Hz, 1H, 1'-H), 6.51 (d, J = 3.7 Hz, 1H, 7-H), 5.41 (m, 1H, 3'-H), 4.33-4.36 (m, 2H, 5'-H), 4.22 (m, 1H, 4'-H), 4.08 (s, 3H, 6-OCH₃), 2.83-2.96 (m, 2H, one of CH(CH₃)₂ and one of 2'-H), 2.54-2.70 (m, 2H, two of $CH(CH_3)_2$), 2.48-2.54 (ddd, J = 2.0, 5.9, 14.2Hz, one of 2'-H), 1.15-1.23 (m, 18H, CH(CH₃)₂); ¹³C NMR (100 MHz, CD₃OD) δ 178.2, 177.7, 177.4, 164.2, 153.4, 152.5, 123.4, 103.5, 100.7, 85.2, 83.0, 75.9, 65.0, 54.4, 37.9, 36.6, 35.0, 34.9, 19.9 (two CH₃), 19.3-19.4 (four peaks for four CH₃); HRMS (FAB+) calcd for $C_{24}H_{35}O_7N_4$ (M+H⁺): 491.2506, found: 491.2503.

7-Iodo-6-methoxy-2-(2-methylpropanoyl)amino-9-[3',5'-bis-O-(2-methylpropanoyl)β-D-2'-deoxyribofuranosyl]-7-deazapurine (36). To a vigorously stirred solution of 35
(1.75 g; 3.57 mmol) in anhydrous DMF (27 mL) was added N-iodosuccinimide (866 mg;
3.66 mmol). The reaction mixture was stirred at room temperature for 22 h, and then

139

most solvent was removed under vacuum. Diethyl ether (200 mL) and saturated aqueous NaHCO₃ (50 mL) were added. The organic layer was separated and washed with saturated aqueous NaCl, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:3) as the eluent to afford **36** as a white solid (1.98 g; 90% yield): ¹H NMR (400 MHz, CD₃OD) δ 7.43 (s, 1H, 8-H), 6.63 (dd, *J* = 6.0, 8.2 Hz, 1H, 1'-H), 5.41 (m, 1H, 3'-H), 4.33-4.36 (m, 2H, 5'-H), 4.23 (m, 1H, 4'-H), 4.09 (s, 3H, 6-OCH₃), 2.78-2.94 (m, 2H, one of CH(CH₃)₂ and one of 2'-H), 2.57-2.70 (m, 2H, two of CH(CH₃)₂); ¹³C NMR (100 MHz, CD₃OD) δ 178.3, 177.8, 177.5, 164.3, 153.3, 152.8, 128.6, 105.2, 85.3, 83.3, 75.8, 65.0, 54.4, 51.8, 38.2, 36.8, 35.2, 35.1, 19.9 (two CH₃), 19.3-19.5 (four peaks for four CH₃); HRMS (FAB+) calcd for C₂₄H₃₄O₇N₄I (M+H⁺): 617.1472, found: 617.1464.

2-Amino-9-(β -D-2'-deoxyribofuranosyl)-7-iodo-6-methoxy-7-deazapurine (41). 36 (1.98 g; 3.21 mmol) was dissolved in 0.5 M methanolic CH₃ONa (50 mL) and stirred at 65 °C for 12 h. Saturated aqueous NaHCO₃ (20 mL) was added and the mixture was stirred for 10 min. Then most of the methanol was evaporated and the residue was extracted by ethyl acetate (150 mL). The organic layer was washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography

over silica gel using CH₃OH-CH₂Cl₂ (1:30~15) as the eluent to afford **41** as a white solid (1.23 g; 94% yield): ¹H NMR (400 MHz, CD₃OD) δ 7.17 (s, 1H, 8-H), 6.36 (dd, J = 6.0, 8.4 Hz, 1H, 1'-H), 4.47 (m, 1H, 3'-H), 3.99 (s, 3H, 6-OCH₃), 3.96 (m, 1H, 4'-H), 3.77 (dd, J = 3.4, 12.0 Hz, 1H, one of 5'-H), 3.70 (dd, J = 3.7, 12.0 Hz, 1H, one of 5'-H), 2.55-2.64 (ddd, J = 6.0, 8.4, 13.4 Hz, one of 2'-H), 2.20-2.26 (ddd, J = 2.4, 5.9, 13.4 Hz, one of 2'-H); ¹³C NMR (100 MHz, CD₃OD) δ 164.7, 160.6, 154.3, 126.5, 101.6, 88.7, 86.0, 73.0, 63.7, 53.7, 51.3, 41.1; HRMS (FAB+) calcd for C₁₂H₁₆O₄N₄I (M+H⁺): 407.0216, found: 407.0213.

2-Amino-9-[β-D-5'-O-(tert-butyldimethylsilyl)-2'-deoxyribofuranosyl]-7-iodo-6-

methoxy-7-deazapurine (42). The procedure is the same as that of 19 and the crude product was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:2~0.5) as the eluent to afford 42 as white foam (1.38 g; 88% yield): ¹H NMR (400 MHz, CD₃OD) δ 7.23 (s, 1H, 8-H), 6.49 (dd, J = 6.1, 7.7 Hz, 1H, 1'-H), 4.46 (m, 1H, 3'-H), 3.99 (s, 3H, 6-OCH₃), 3.94 (m, 1H, 4'-H), 3.79-3.87 (m, 2H, 5'-H), 2.36-2.44 (ddd, J = 5.8, 7.7, 13.3 Hz, one of 2'-H), 2.24-2.31 (ddd, J = 3.1, 6.0, 13.3 Hz, one of 2'-H), 0.96 (s, 9H, C(CH₃)₃), 0.14 (s, 3H, one of SiCH₃), 0.13 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CD₃OD) δ 164.6, 160.7, 154.7, 125.1, 101.0, 88.2, 84.2, 72.7, 64.7, 53.7, 51.7, 41.9, 26.7, 19.4, -5.0, -5.1; HRMS (FAB+) calcd for C₁₈H₃₀O₄N₄SiI (M+H⁺): 521.1081, found: 521.1068. **9-**[β-D-3'-O-Allyl-5'-O-(*tert*-butyldimethylsilyl)-2'-deoxyribofuranosyl]-2-amino-7iodo-6-methoxy-7-deazapurine (43). The procedure is the same as that of 20 and the crude product was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:3) as the eluent to afford 43 as a white solid (1.37 g; 92% yield): ¹H NMR (400 MHz, CD₃OD) δ 7.20 (s, 1H, 8-H), 6.43 (dd, J = 6.2, 7.9 Hz, 1H, 1'-H), 5.89-5.99 (m, 1H, CH₂CH=CH₂), 5.29-5.35 (dm, J = 17.3 Hz, 1H, one of CH₂CH=CH₂), 5.16-5.21 (dm, J = 10.5 Hz, 1H, one of CH₂CH=CH₂), 4.24 (m, 1H, 3'-H), 4.01-4.11 (m, 3H, 4'-H and CH₂CH=CH₂), 3.99 (s, 3H, 6-OCH₃), 3.76-3.84 (m, 2H, 5'-H), 2.32-2.44 (m, 2H, 2'-H), 0.95 (s, 9H, C(CH₃)₃), 0.14 (s, 3H, one of SiCH₃), 0.13 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CDCl₃) δ 163.3, 158.6, 153.6, 134.1, 123.7, 116.9, 100.6, 84.4, 83.0, 79.1, 70.0, 63.6, 53.3, 51.1, 38.1, 26.1, 18.5, -5.1, -5.3; HRMS (FAB+) calcd for C₂₁H₃₄O₄N₄SiI (M+H⁺): 561.1394, found: 561.1390.

9-[β -D-3'-O-Allyl-5'-O-(*tert*-butyldimethylsilyl)-2'-deoxyribofuranosyl]-2-amino-6methoxy-7-{3-[(trifluoroacetyl)amino]-prop-1-ynyl}-7-deazapurine (46). The procedure is the same as that of 12 and the crude product was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:3~1.5) and CH₃OH-CH₂Cl₂ (1:30) respectively as the eluents to afford 46 as a yellow solid (1.34 g; 94% yield): ¹H NMR (400 MHz, CD₃OD) δ 7.34 (s, 1H, 8-H), 6.42 (dd, J = 6.2, 7.7 Hz, 1H, 1'-H), 5.88-5.99 (m, 1H, CH₂CH=CH₂), 5.28-5.35 (dm, J = 17.3 Hz, 1H, one of CH₂CH=CH₂), 5.16-5.21 (dm, J = 10.5 Hz, 1H, one of CH₂CH=CH₂), 4.29 (s, 2H, C=CCH₂), 4.24 (m, 1H, 3'-H), 4.00-4.09 (m, 3H, 4'-H and $CH_2CH=CH_2$), 3.98 (s, 3H, 6-OC H_3), 3.76-3.84 (m, 2H, 5'-H), 2.32-2.45 (m, 2H, 2'-H), 0.94 (s, 9H, C(CH_3)₃), 0.12 (s, 3H, one of SiC H_3), 0.11 (s, 3H, one of SiC H_3); ¹³C NMR (100 MHz, CD₃OD) δ 165.0, 161.2, 158.1 (q, J = 36 Hz, COCF₃), 154.2, 135.6, 125.0, 117.2 (q, J = 284 Hz, COCF₃), 117.0, 99.2, 97.3, 86.0, 84.6, 84.5, 80.3, 78.0, 71.0, 64.8, 53.8, 39.0, 30.9, 26.5, 19.3, -5.1, -5.2; HRMS (FAB+) calcd for C₂₆H₃₇O₅N₅F₃Si (M+H⁺): 584.2516, found: 584.2491.

3'-O-Allyl-7-deaza-7-{3-[(trifluoroacetyl)amino]prop-1-ynyl}-2'-deoxyguanosine

(24). To a stirred solution of 46 (1.34 g; 2.30 mmol) in anhydrous CH₃CN (86 mL) were added NaI (363 mg; 2.42 mmol) and chlorotrimethylsilane (306 μ L; 2.42 mmol). The reaction mixture was stirred at room temperature for 1 h and then at 50 °C for 12 h. The solvent was evaporated and the residue was dissolved in anhydrous THF (76 mL). 1 M TBAF in THF solution (4.80 mL; 4.80 mmol) was added and the reaction mixture was stirred at room temperature for 1 h. The solvent was evaporated and the residue was dissolved in ethyl acetate (150 mL). The solvent was evaporated and the residue was dissolved in ethyl acetate (150 mL). The solution was washed with saturated aqueous NaCl and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₃OH-ethyl acetate (1:30) as the eluent to afford **24** as a yellow solid (356 mg; 34% yield): ¹H NMR (400 MHz, CD₃OD) δ 7.21 (s, 1H, 8-H), 6.30 (dd, *J* = 6.0, 8.4 Hz, 1H, 1'-H), 5.88-5.99 (m, 1H, CH₂CH=CH₂), 5.28-5.35 (dm, *J* = 17.3 Hz, 1H, one of CH₂CH=CH₂), 5.15-5.20 (dm, *J* = 10.5 Hz, 1H, one of CH₂CH=CH₂), 4.29 (s, 2H, C=CCH₂), 4.23 (m, 1H, 3'-H), 4.00-4.10

(m, 3H, 4'-H and CH₂CH=CH₂), 3.65-3.75 (m, 2H, 5'-H), 2.41-2.49 (ddd, J = 5.8, 8.4, 13.6 Hz, 1H, one of 2'-H), 2.34-2.40 (ddd, J = 2.3, 6.0, 13.6 Hz, 1H, one of 2'-H); ¹³C NMR (100 MHz, CD₃OD) δ 160.9, 158.0 (q, J = 36 Hz, COCF₃), 154.1, 151.8, 135.6, 124.4, 117.2 (q, J = 284 Hz, COCF₃), 117.0, 101.4, 99.7, 86.4, 85.5, 84.8, 80.7, 78.0, 71.0, 63.7, 38.5, 31.2; HRMS (FAB+) calcd for C₁₉H₂₁O₅N₅F₃ (M+H⁺): 456.1495, found: 456.1493.

3'-O-Allyl-7-(3-aminoprop-1-ynyl)-7-deaza-2'-deoxyguanosine-5'-triphosphate (2).

The procedure is the same as that of 1 to yield 2 as a colorless syrup: ¹H NMR (300 MHz, D₂O) δ 7.56 (s, 1H, 6-H), 6.37 (t, J = 7.3 Hz, 1H, 1'-H), 5.89-6.02 (m, 1H, CH₂CH=CH₂), 5.31-5.39 (dm, J = 17.3 Hz, 1H, one of CH₂CH=CH₂), 5.21-5.28 (dm, J = 10.5 Hz, 1H, one of CH₂CH=CH₂), 4.49 (s, 2H, C=CCH₂), 4.32 (m, 1H, 3'-H), 4.06-4.18 (m, 3H, 4'-H and CH₂CH=CH₂), 3.92-3.99 (m, 2H, 5'-H), 2.44-2.60 (m, 2H, 2'-H); ³¹P NMR (121.4 MHz, D₂O) δ -6.1 (d, J = 20.8 Hz, 1P, γ -P), -10.8 (d, J = 18.9Hz, 1P, α -P), -21.9 (t, J = 19.8 Hz, 1P, β -P).

5'-O-(*tert***-Butyldimethylsilyl)-5-iodo-2'-deoxycytidine (48).** The procedure is the same as that of **19** and the crude product was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:20) as the eluent to afford **48** as a white solid (1.18 g; 89% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.18 (s, 1H, 6-H), 6.17 (dd, *J* = 5.8, 7.5 Hz, 1H, 1'-H), 4.34 (m, 1H, 3'-H), 4.04 (m, 1H, 4'-H), 3.93 (dd, *J* = 2.5, 11.6 Hz, 1H, one of

5'-H), 3.84 (dd, J = 2.9, 11.6 Hz, 1H, one of 5'-H), 2.41-2.48 (ddd, J = 2.5, 5.8, 13.5 Hz, 1H, one of 2'-H), 2.01-2.08 (ddd, J = 5.9, 7.6, 13.5 Hz, 1H, one of 2'-H), 0.95 (s, 9H, $C(CH_3)_3$), 0.17 (s, 3H, one of SiCH₃), 0.16 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CD₃OD) δ 165.5, 156.8, 147.8, 89.4, 88.3, 72.8, 64.6, 57.1, 43.1, 26.7, 19.4, -4.8, -4.9; HRMS (FAB+) calcd for C₁₅H₂₇O₄N₃SiI (M+H⁺): 468.0816, found: 468.0835.

3'-O-Allyl-5'-O-(tert-butyldimethylsilyl)-5-iodo-2'-deoxycytidine (49). To a stirred solution of 48 (1.18 g; 2.52 mmol) in anhydrous THF (43 mL) was added 95% NaH powder (128 mg; 5.07mmol). The suspension was stirred at room temperature for 45 min. Allyl bromide (240 µL, 2.79 mmol) was then added at 0 °C and the reaction mixture was stirred at room temperature for 14 h with exclusion of moisture. Saturated aqueous NaHCO₃ (10 mL) was added at 0 °C and stirred for 10 min. Most of THF was evaporated and the residue was dissolved in ethyl acetate (150 mL). The solution was washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using ethyl acetate as the eluent to afford starting material (378 mg; 32%) and 49 as a white solid (601 mg; 47% yield): ¹H NMR (400 MHz. CD₃OD) δ 8.15 (s, 1H, 6-H), 6.12 (dd, J = 5.6, 8.0 Hz, 1H, 1'-H), 4.17 (m, 1H, 4'-H), 4.14 (m, 1H, 3'-H), 3.98-4.10 (m, 2H, $CH_2CH=CH_2$), 3.93 (dd, J = 2.8, 11.5 Hz, 1H, one of 5'-H), 3.83 (dd, J = 2.8, 11.5 Hz, 1H, one of 5'-H), 2.53-2.60 (ddd, J = 1.7, 5.6, 13.6 Hz, 1H, one of 2'-H), 1.94-2.02 (ddd, J = 5.9, 8.0, 13.6 Hz, 1H, one of 2'-H), 0.94 (s, 9H,

C(CH₃)₃), 0.17 (s, 3H, one of SiCH₃), 0.16 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CD₃OD) δ 165.4, 156.7, 147.7, 135.5, 117.2, 88.2, 87.0, 80.4, 70.9, 64.8, 57.3, 40.1, 26.7, 19.4, -4.7, -4.9; HRMS (FAB+) calcd for C₁₈H₃₁O₄N₃SiI (M+H⁺): 508.1129, found: 508.1123.

3'-O-Allyl-5-iodo-2'-deoxycytidine (50). To a stirred solution of 49 (601 mg; 1.18 mmol) in anhydrous THF (28 mL) was added 1 M TBAF in THF solution (1.31 mL; 1.31 mmol) and the reaction mixture was stirred at room temperature for 1 h. The solvent was evaporated and the residue was dissolved in ethyl acetate (100 mL). The solution was washed with saturated aqueous NaCl and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:10) as the eluent to afford **50** as white crystals (329 mg; 71% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.47 (s, 1H, 6-H), 6.15 (dd, J = 6.2, 6.7 Hz, 1H, 1'-H), 5.87-5.98 (m, 1H, $CH_2CH=CH_2$), 5.26-5.33 (dm, J = 17.2 Hz, 1H, one of CH₂CH=CH₂), 5.14-5.19 (dm, J = 10.5 Hz, 1H, one of CH₂CH=CH₂), 4.18 (m, 1H, 3'-H), 4.08 (m, 1H, 4'-H), 3.98-4.10 (m, 2H, $CH_2CH=CH_2$), 3.82 (dd, J = 3.2, 13.0 Hz, 1H, one of 5'-H), 3.72 (dd, J = 3.3, 13.0 Hz, 1H, one of 5'-H), 2.44-2.51 (ddd, J = 3.2, 6.0, 13.6 Hz, 1H, one of 2'-H), 2.07-2.15 (m, 1H, one of 2'-H); ¹³C NMR (100 MHz, CD₃OD) & 165.4, 156.9, 148.8, 135.6, 117.0, 87.9, 86.9, 79.6, 71.2, 62.7, 57.2, 39.7; HRMS (FAB+) calcd for $C_{12}H_{17}O_4N_3I$ (M+H⁺): 394.0264, found: 394.0274.

3'-O-Allyl-5-{3-[(trifluoroacetyl)amino]prop-1-ynyl}-2'-deoxycytidine (51). The procedure is the same as that of **12** and the crude product was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (0~1:10) as the eluent to afford **51** as yellow crystals (290 mg; 83% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.31 (s, 1H, 6-H), 6.17 (dd, *J* = 6.0, 7.3 Hz, 1H, 1'-H), 5.87-5.97 (m, 1H, CH₂C*H*=CH₂), 5.26-5.33 (dm, *J* = 17.3 Hz, 1H, one of CH₂CH=CH₂), 5.15-5.19 (dm, *J* = 10.4 Hz, 1H, one of CH₂CH=CH₂), 4.31 (s, 2H, C=CCH₂), 4.17 (m, 1H, 3'-H), 4.09 (m, 1H, 4'-H), 3.98-4.10 (m, 2H, CH₂CH=CH₂), 3.80 (dd, *J* = 3.4, 12.0 Hz, 1H, one of 5'-H), 3.72 (dd, *J* = 3.6, 12.0 Hz, 1H, one of 5'-H), 2.46-2.53 (ddd, *J* = 2.9, 5.3, 13.6 Hz, 1H, one of 2'-H), 2.04-2.12 (m, 1H, one of 2'-H); ¹³C NMR (100 MHz, CD₃OD) δ 166.0, 158.4 (q, *J* = 38 Hz, COCF₃), 156.3, 145.8, 135.6, 117.1 (q, *J* = 284 Hz, COCF₃), 117.0, 91.9, 90.7, 88.0, 87.0, 79.8, 75.5, 71.2, 62.8, 39.6, 31.0; HRMS (FAB+) calcd for C₁₇H₂₀O₅N₄F₃ (M+H⁺): 417.1386, found: 417.1377.

3'-O-Allyl-5-(3-aminoprop-1-ynyl)-2'-deoxycytidine-5'-triphosphate (3). The procedure is the same as that of 1 to yield 3 as a colorless syrup: ¹H NMR (300 MHz, D₂O) δ 8.43 (s, 1H, 6-H), 6.21 (t, J = 6.7 Hz, 1H, 1'-H), 5.85-6.00 (m, 1H, CH₂CH=CH₂), 5.28-5.38 (dm, J = 17.3 Hz, 1H, one of CH₂CH=CH₂), 5.19-5.27 (dm, J = 10.4 Hz, 1H, one of CH₂CH=CH₂), 4.22-4.41 (m, 3H, 3'-H and C=CCH₂), 4.05-4.18 (m, 3H, 4'-H and CH₂CH=CH₂), 3.94-4.01 (m, 2H, 5'-H), 2.47-2.59 (m, 1H, one of 2'-H), 2.20-2.32 (m, 1H, one of 2'-H); ³¹P NMR (121.4 MHz, D₂O) δ -7.1 (d, J = 19.8 Hz, 1P, γ -P), -11.1 (d, J

= 19.1 Hz, 1P,
$$\alpha$$
-P), -21.9 (t, J = 19.5 Hz, 1P, β -P).

5'-O-(*tert***-Butyldimethylsilyl)-5-iodo-2'-deoxyuridine (53).** The procedure is the same as that of **19** and the crude product was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:20) as the eluent to afford **53** as a white solid (1.19 g; 90% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.17 (s, 1H, 6-H), 6.21 (dd, *J* = 5.9, 7.9 Hz, 1H, 1'-H), 4.36 (m, 1H, 3'-H), 4.02 (m, 1H, 4'-H), 3.93 (dd, *J* = 2.4, 11.5 Hz, 1H, one of 5'-H), 3.85 (dd, *J* = 2.9, 11.5 Hz, 1H, one of 5'-H), 2.30-2.37 (ddd, *J* = 2.3, 5.8, 13.4 Hz, 1H, one of 2'-H), 2.08-2.15 (ddd, *J* = 5.9, 7.9, 13.4 Hz, 1H, one of 2'-H), 0.96 (s, 9H, C(CH₃)₃), 0.19 (s, 3H, one of SiCH₃), 0.17 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CD₃OD) δ 162.4, 151.5, 145.8, 89.3, 87.2, 72.8, 68.7, 64.6, 42.3, 26.8, 19.5, -4.8, -4.9; HRMS (FAB+) calcd for C₁₅H₂₆O₅N₂SiI (M+H⁺): 469.0656, found: 469.0672.

3'-O-Allyl-5'-O-(*tert*-butyldimethylsilyl)-5-iodo-2'-deoxyuridine (54). The procedure is the same as that of **49** and the crude product was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:2.5) as the eluent to afford **54** as a white solid (1.03 g; 80% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.15 (s, 1H, 6-H), 6.15 (dd, *J* = 5.6, 8.3 Hz, 1H, 1'-H), 5.87-5.97 (m, 1H, CH₂CH=CH₂), 5.27-5.33 (dm, *J* = 17.3 Hz, 1H, one of CH₂CH=CH₂), 5.16-5.21 (dm, *J* = 10.4 Hz, 1H, one of CH₂CH=CH₂), 4.13-4.18 (m, 2H, 3'-H and 4'-H), 3.99-4.10 (m, 2H, CH₂CH=CH₂), 3.92 (dd, *J* = 2.7, 11.5 Hz, 1H, one of 5'-H), 3.84 (dd, *J* = 2.7, 11.5 Hz, 1H, one of 5'-H), 2.43-2.49 (ddd, *J* = 1.7, 5.6, 13.6 Hz, 1H, one of 2'-H), 2.02-2.10 (ddd, J = 5.6, 8.4, 13.6 Hz, 1H, one of 2'-H), 0.96 (s, 9H, C(CH₃)₃), 0.18 (s, 3H, one of SiCH₃), 0.17 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CD₃OD) δ 162.3, 151.4, 145.5, 135.5, 117.2, 87.0, 86.8, 80.3, 70.9, 69.0, 64.8, 39.4, 26.8, 19.4, -4.7, -4.8; HRMS (FAB+) calcd for C₁₈H₃₀O₅N₂SiI (M+H⁺): 509.0969, found: 509.0970.

3'-O-Allyl-5'-O-(tert-butyldimethylsilyl)-5-{3-[(trifluoroacetyl)amino]prop-1-ynyl}-2'-deoxyuridine (58). The procedure is the same as that of 12 and the crude product was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (0~1:40) as the eluent to afford **58** as yellow crystals (786 mg; 73% yield): ¹H NMR (400 MHz, CD_3OD) δ 8.11 (s, 1H, 6-H), 6.18 (dd, J = 5.8, 7.9 Hz, 1H, 1'-H), 5.87-5.97 (m, 1H, $CH_2CH=CH_2$), 5.27-5.33 (dm, J = 17.2 Hz, 1H, one of $CH_2CH=CH_2$), 5.16-5.21 (dm, J = 17.2 Hz, 1H, 10.4 Hz, 1H, one of CH₂CH=CH₂), 4.27-4.32 (dd, J = 17.7 Hz, 1H, one of C=CCH₂), 4.21-4.27 (dd, J = 17.7 Hz, 1H, one of C=CCH₂), 4.14-4.18 (m, 2H, 3'-H and 4'-H), 3.98-4.10 (m, 2H, CH₂CH=CH₂), 3.93 (dd, J = 2.4, 11.5 Hz, 1H, one of 5'-H), 3.84 (dd, J= 2.2, 11.5 Hz, 1H, one of 5'-H), 2.44-2.50 (ddd, J = 1.8, 5.7, 13.5 Hz, 1H, one of 2'-H), 2.04-2.12 (ddd, J = 5.6, 8.0, 13.5 Hz, 1H, one of 2'-H), 0.94 (s, 9H, C(CH₃)₃), 0.16 (s, 3H, one of SiCH₃), 0.15 (s, 3H, one of SiCH₃); ¹³C NMR (100 MHz, CD₃OD) δ 164.1, 158.0 $(q, J = 37 \text{ Hz}, \text{COCF}_3)$, 150.6, 144.3, 135.5, 117.3, 117.1 $(q, J = 284 \text{ Hz}, \text{COCF}_3)$, 99.5, 88.9, 87.2, 86.9, 80.3, 76.0, 71.0, 64.7, 39.6, 30.7, 26.6, 19.3, -5.0, -5.2; HRMS (FAB+) calcd for $C_{23}H_{33}O_6N_3F_3Si(M+H^+)$: 532.2091, found: 532.2074.

3'-O-Allyl-5-{3-[(trifluoroacetyl)amino]prop-1-ynyl}-2'-deoxyuridine (56). The procedure is the same as that of **50** and the crude product was purified by flash column chromatography over silica gel using ethyl acetate-hexane (3:1) as the eluent to afford **56** as a yellow solid (302 mg; 49% yield): ¹H NMR (400 MHz, CD₃OD) δ 8.29 (s, 1H, 6-H), 6.19 (dd, J = 6.1, 7.4 Hz, 1H, 1'-H), 5.87-5.99 (m, 1H, CH₂CH=CH₂), 5.27-5.33 (dm, J = 17.2 Hz, 1H, one of CH₂CH=CH₂), 5.15-5.20 (dm, J = 10.4 Hz, 1H, one of CH₂CH=CH₂), 4.27 (s, 2H, C=CCH₂), 4.20 (m, 1H, 3'-H), 3.99-4.09 (m, 3H, 4'-H and CH₂CH=CH₂), 3.80 (dd, J = 3.3, 12.0 Hz, 1H, one of 5'-H), 3.72 (dd, J = 3.4, 12.0 Hz, 1H, one of 5'-H), 2.39-2.46 (ddd, J = 2.6, 5.9, 13.7 Hz, 1H, one of 2'-H), 2.14-2.22 (ddd, J = 6.3, 7.5, 13.7 Hz, 1H, one of 2'-H); ¹³C NMR (100 MHz, CD₃OD) δ 164.2, 158.0 (q, J = 38 Hz, COCF₃), 150.8, 145.3, 135.6, 117.2 (q, J = 285 Hz, COCF₃), 117.1, 99.5, 88.3, 87.1, 87.0, 79.9, 75.9, 71.2, 62.9, 39.0, 30.8; HRMS (FAB+) calcd for C₁₇H₁₉O₆N₃F₃ (M+H⁺): 418.1226, found: 418.1213.

3'-O-Allyl-5-(3-aminoprop-1-ynyl)-2'-deoxyuridine-5'-triphosphate (4). The procedure is the same as that of 1. For the ¹H and ¹³P NMR data, see *Reference* [4].

N-propargyltrifluoroacetamide.³⁶



To a stirred solution of propargylamine (5.0 g; 90.8 mmol) in methanol (25 mL) was

added methyl trifluoroacetate (9.2 mL; 92.0 mmol) slowly at 0 °C and the reaction mixture was stirred at room temperature for 3 h. Most of methanol was evaporated and the residue was distilled under vacuum (b.p.: 29-32 °C/0.3 mmHg) to afford *N*-propargyl-trifluoroacetamide as a colorless oil (10.7 g; 78% yield): ¹H NMR (400 MHz, CD₃OD) δ 4.06 (d, *J* = 2.7 Hz, 2H, HC=CCH₂), 2.63 (t, *J* = 2.7 Hz, 1H, *H*C=CCH₂); ¹³C NMR (100 MHz, CD₃OD) δ 158.3 (q, *J* = 38 Hz, *C*OCF₃), 117.1 (q, *J* = 284 Hz, COCF₃), 78.8, 72.9, 29.8.

2-Triphenylmethoxymethyl-2-propen-1-ol (60). To a stirred solution of trityl chloride (4.05 g; 14.3 mmol) and 2-methylenepropane-1,3-diol **59** (1.20 mL; 14.3 mmol) in dry CH₂Cl₂ (20 mL) was added triethylamine (4.0 mL; 28.5 mmol) slowly at room temperature. The reaction mixture was stirred at room temperature for 1 h and then ethyl acetate (100 mL) and saturated aqueous NaHCO₃ (30 mL) were added. The organic layer was separated and washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using ethyl acetate-hexane (1:10~5) as the eluent to afford **60** as a white solid (2.89 g; 62% yield): ¹H NMR (400 MHz, CDCl₃) δ 7.42-7.48 (m, 6H, six of ArH), 7.27-7.33 (m, 6H, six of ArH), 7.20-7.27 (m, 3H, three of ArH), 5.26 (s, 1H, one of C=CH₂), 5.17 (s, 1H, one of C=CH₂), 4.13 (d, *J* = 6.1 Hz, 2H, CH₂OH), 3.70 (s, 2H, Ph₃COCH₂), 1.72 (t, *J* = 6.1 Hz, 1H, CH₂OH); ¹³C NMR (100 MHz, CDCl₃) δ 145.4, 143.6, 128.3, 127.6, 126.8, 111.6, 87.0, 65.3, 64.5.

1-Bromo-2-triphenylmethoxymethyl-2-propene (61). To a stirred solution of **60** (2.56 g; 7.74 mmol) in CH₂Cl₂ (75 ml) were added CBr₄ (3.63 g; 10.83 mmol) and triphenylphosphine (2.47 g; 9.31 mmol) respectively at 0 °C and the reaction mixture was stirred at room temperature for 40 min. Cooled to 0 °C, ethyl acetate (100 mL) and saturated aqueous NaHCO₃ (30 mL) were added. The organic layer was separated and washed with saturated aqueous NaCl, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₂Cl₂-hexane (1:5) as the eluent to afford **61** as a white solid (3.02 g; 92% yield): ¹H NMR (400 MHz, CDCl₃) δ 7.42-7.48 (m, 6H, six of ArH), 7.27-7.33 (m, 6H, six of ArH), 7.20-7.27 (m, 3H, three of ArH), 5.37 (s, 1H, one of C=CH₂), 5.31 (s, 1H, one of C=CH₂), 4.01 (s, 2H, CH₂Br), 3.75 (s, 2H, Ph₃COCH₂); ¹³C NMR (100 MHz, CDCl₃) δ 143.6, 142.6, 128.4, 127.6, 126.9, 115.8, 86.9, 64.2, 33.5.

3-Triphenylmethoxymethyl-3-butenonitrile (62). To a stirred solution of **61** (1.45 g; 3.69 mmol) and in dry CH₃CN (37 mL) was added trimethylsilyl cyanide (0.49 mL; 3.69 mmol). Then 1 M tetrabutylammonium fluoride (TBAF) in THF solution (3.69 mL, 3.69 mmol) was added slowly at room temperature and the reaction mixture was stirred for 20 min. Most of the solvents were evaporated and ethyl acetate (100 mL) and saturated aqueous NaHCO₃ (30 mL) were added. The organic layer was washed with saturated aqueous NaCl and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using acetate-

hexane (1:10) as the eluent to afford **62** as a white solid (1.01 g; 64% yield): ¹H NMR (400 MHz, CDCl₃) δ 7.39-7.45 (m, 6H, six of ArH), 7.21-7.34 (m, 9H, nine of ArH), 5.31 (s, 2H, C=CH₂), 3.64 (s, 2H, Ph₃COCH₂), 3.11 (s, 2H, CH₂CN); ¹³C NMR (100 MHz, CDCl₃) δ 143.3, 135.5, 128.2, 127.7, 126.9, 116.8, 114.7, 87.0, 65.7, 21.9.

3-Triphenylmethoxymethyl-3-buten-1-amine (63). To a stirred solution of LiAlH₄ (119 mg; 2.98 mmol) in dry ether (5 mL) was added AlCl₃ (400 mg; 2.94 mmol) carefully at 0 °C and the mixture was stirred for 15 min. Then a solution of 62 (829 mg; 2.44 mmol) in dry ether (9 mL) was added and the reaction mixture was stirred at 0 °C for 3h. After that 10% aqueous NaOH (10 mL) was added to quench the reaction. The organic layer was separated and washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous K_2CO_3 . After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:20~5) as the eluent to afford **63** as a colorless oil (545 mg; 65% yield): ¹H NMR (400 MHz, CDCl₃) δ 7.41-7.48 (m, 6H, six of ArH), 7.26-7.33 (m, 6H, six of ArH), 7.19-7.26 (m, 3H, three of ArH), 5.33 (s, 1H, one of C=C H_2), 4.96 (s, 1H, one of C=C H_2), 3.53 (s, 2H, Ph₃COC H_2), 2.70 (m, 2H, CH₂CH₂NH₂), 2.18 (t, J = 6.7 Hz, 2H, CH₂CH₂NH₂), 2.06 (br s, 2H, NH₂); ¹³C NMR (100 MHz, CDCl₃) δ 143.6, 143.4, 128.1, 127.4, 126.5, 111.3, 86.5, 66.1, 39.8, 37.4; HRMS (FAB+) calcd for $C_{24}H_{26}ON$ (M+H⁺): 344.2014, found: 344.2017.

4-Amino-2-methylene-1-butanol (5). To a stirred solution of 63 (540 mg; 1.57 mmol) in

CH₃OH (11 mL) was added HCl (2M solution in ether; 5.5 mL) at room temperature and the reaction mixture was stirred for 1 h. Then 7M methanolic ammonia (2.7 mL) was added at room temperature and the reaction mixture was stirred for 10 min. The white solid formed was removed by filtration and washed with CH₃OH. The combined solution was evaporated and the residue was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:4) as the eluent to afford **5** as a colorless oil (151 mg; 95% yield): ¹H NMR (400 MHz, CD₃OD) δ 5.19 (s, 1H, one of C=CH₂), 5.01 (s, 1H, one of C=CH₂), 4.06 (s, 2H, CH₂OH), 3.10 (t, *J* = 7.5 Hz, 2H, CH₂CH₂NH₂), 2.46 (t, *J* = 7.5 Hz, 2H, CH₂CH₂NH₂); ¹³C NMR (100 MHz, CD₃OD) δ 145.3, 113.7, 65.5, 39.5, 32.0; LRMS (FAB+) calcd for C₃H₁₂ON (M+H⁺): 102.09, found: 102.12.

5-(Bromomethyl)-2-nitroacetophenone (68) was prepared from 66 using the exact procedure in *Reference [35]*.

1-[5-(Bromomethyl)-2-nitrophenyl]ethanol (69). To a stirred solution of 68 (500 mg; 1.94 mmol) in CH₃OH (9 mL) was added NaBH₄ (75 mg; 1.97 mmol) at 0 °C. The reaction mixture was stirred at room temperature for 1 h. Acteone (1 mL) was added and the reaction was stirred for another 10 min. Most of the solvents were evaporated and ethyl acetate (50 mL) and saturated aqueous NaHCO₃ (20 mL) were added. The organic layer was washed with saturated aqueous NaCl and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₃CN-CH₂Cl₂ (1:20) as the eluent to afford **69** as a white solid (499 mg; 99% yield): ¹H NMR (400 MHz, CDCl₃) δ 7.90 (d, J = 8.4 Hz, 1H, one of ArH), 7.86 (d, J = 2.0 Hz, 1H, one of ArH), 7.44 (dd, J = 2.0, 8.4 Hz, 1H, one of ArH), 5.45 (q, J = 6.4 Hz, 1H, CH₃CH), 4.51 (s, 2H, CH₂Br), 2.39 (br s, 1H, OH), 1.57 (d, J = 6.4 Hz, 3H, CH₃CH).

1-[5-(Aminomethyl)-2-nitrophenyl]ethanol (65). To a stirred 7M methanolic ammonia (20 mL) was added a solution of 69 (450 mg; 1.73 mmol) in CH₃OH (2 mL) slowly. The reaction mixture was stirred at room temperature for 4 h. Most solvents were evaporated and ethyl acetate (50 mL) and saturated aqueous NaCl (20 mL) were added. The organic layer was separated and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:2.5) as the eluent to afford 69 as a white solid (244 mg; 72% yield): ¹H NMR (400 MHz, CDCl₃) δ 7.88 (d, *J* = 8.4 Hz, 1H, one of ArH), 7.78 (d, *J* = 1.3 Hz, 1H, one of ArH), 7.33 (dd, *J* = 1.7, 8.4 Hz, 1H, one of ArH), 5.42 (q, *J* = 6.4 Hz, 1H, CH₃CH), 3.95 (s, 2H, CH₂NH₂), 2.28 (br s, 3H, OH and NH₂), 1.54 (d, *J* = 6.4 Hz, 3H, CH₃CH); ¹³C NMR (100 MHz, CD₃OD) δ 149.4, 147.3, 143.0, 127.4, 127.0, 125.1, 66.0, 46.2, 25.2.

Allyl-Bodipy-FL-510 (71). To a stirred solution of 5 (3.5 mg; 34.6 μ mol) in DMF (450 μ L) was added 1M aqueous NaHCO₃ (100 μ L) at room temperature and the solution was

stirred for 5 min. Bodipy-FL-510 NHS (*N*-hydroxysuccinimide) ester **70** (5 mg; 12.8 μ mol) in DMF (450 μ L) was added and the reaction mixture was stirred at room temperature for 6 h. The crude product was purified by a preparative silica gel TLC plate using CH₃OH-CH₂Cl₂ (1:15) as the eluent to afford **71**: ¹H NMR (400 MHz, CD₃OD) δ 7.42 (s, 1H), 7.00 (d, *J* = 4.0 Hz, 1H), 6.32 (d, *J* = 4.0 Hz, 1H), 6.21 (s, 1H), 5.06 (s, 1H, one of C=CH₂), 4.87 (s, 1H, one of C=CH₂, partly superimposed by solvent signal), 4.01 (s, 2H, CH₂OH), 3.33 (t, *J* = 7.5 Hz, 2H, partly superimposed by solvent signal), 3.21 (t, *J* = 7.7 Hz, 2H), 2.59 (t, *J* = 7.7 Hz, 2H), 2.51 (s, 3H, one of ArCH₃), 2.28 (s, 3H), 2.26 (t, *J* = 7.1 Hz, 2H); HRMS (FAB+) calcd for C₁₉H₂₄O₂N₃F₂B (M⁺): 375.1933, found: 375.1957.

Allyl-Bodipy-FL-510 NHS ester (72). To a stirred solution of 71 in dry DMF (350 μ L) were added *N*,*N*^{*}-disuccinimidyl carbonate (DSC) (8.0 mg; 31.2 μ mol) and triethylamine (5.0 μ L; 35.4 μ mol). The reaction mixture was stirred at room temperature for 10 h and then most of DMF was removed under vacuum. The crude product was purified quickly by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:50) as the eluent to afford 72, which was used directly for the next step.

3'-O-allyl-dCTP-allyl-Bodipy-FL-510 (73). To a stirred solution of **72** in DMF (300 μ L) was added 3'-O-allyl-dCTP-NH₂ **3** in 1M NaHCO₃-Na₂CO₃ buffer (300 μ L; pH 8.7). The reaction mixture was stirred at room temperature for 7 h and the crude product was

purified by a preparative silica gel TLC plate using CH₃OH-CH₂Cl₂ (1:1) as the eluent. The crude product was then purified by reverse-phase HPLC using a 50×4.6 mm C18 column to afford pure product (retention time = 35 min). Mobile phase: A, 4.4 mM Et₃N / 98.3 mM 1,1,1,3,3,3-hexafluoro-2-propanol in water (pH = 8.1); B, methanol. Elution was performed from 100% A isocratic over 10 min followed by a linear gradient of 0-50% B for 20 min and then 50% B isocratic over 20 min. The product was characterized by testing in a single base extension reaction.

Allyl-R6G (75). The procedure is the same as that of 71 and the crude product was purified using a preparative silica gel TLC plate with CH₃OH-CH₂Cl₂ (1:2) as the eluent to afford 75: ¹H NMR (400 MHz, CD₃OD) δ 8.12 (d, J = 8.1 Hz, 1H), 8.05 (dd, J = 1.8, 8.1 Hz, 1H), 7.66 (d, J = 1.6 Hz, 1H), 7.02 (s, 2H), 6.88 (s, 2H), 5.08 (s, 1H, one of C=CH₂), 4.92 (s, 1H, one of C=CH₂), 4.06 (s, 2H, CH₂OH), 3.48-3.56 (m, 6H), 2.40 (t, J = 7.2 Hz, 2H), 2.13 (s, 6H), 1.36 (t, J = 7.2 Hz, 6H); HRMS (FAB+) calcd for C₃₂H₃₆O₅N₃ (M+H⁺): 542.2655, found: 542.2648.

Allyl-ROX (79). The procedure is the same as that of 71 and the crude product was purified using a preparative silica gel TLC plate with CH₃OH-CH₂Cl₂ (1:2) as the eluent to afford 79: ¹H NMR (400 MHz, CD₃OD) δ 8.03 (d, J = 8.1 Hz, 1H), 7.98 (dd, J = 1.6, 8.1 Hz, 1H), 7.60 (d, J = 1.4 Hz, 1H), 6.75 (s, 2H), 5.08 (s, 1H, one of C=CH₂), 4.91 (s, 1H, one of C=CH₂), 4.05 (s, 2H, CH₂OH), 3.45-3.57 (m, 10H), 3.03-3.10 (m, 4H),

2.64-2.73 (m, 4H), 2.38 (t, J = 7.1 Hz, 2H), 2.04-2.15 (m, 4H), 1.89-1.99 (m, 4H); HRMS (FAB+) calcd for C₃₈H₄₀O₅N₃ (M+H⁺): 618.2968, found: 618.2961.

Allyl-Bodipy-650 (83). The procedure is the same as that of 71 and the crude product was purified using a preparative silica gel TLC plate with CH₃OH-CH₂Cl₂ (1:10) as the eluent to afford 83: ¹H NMR (400 MHz, CD₃OD) δ 8.12 (dd, J = 0.9, 3.8 Hz, 1H), 7.63 (m, 3H), 7.54 (d, J = 6.4 Hz, 2H), 7.35 (s, 1H), 7.18-7.22 (m, 2H), 7.12 (m, 2H), 7.06 (d, J = 8.8 Hz, 2H), 6.85 (d, J = 4.2 Hz, 1H), 5.06 (s, 1H, one of C=CH₂), 4.86 (s, 1H, one of C=CH₂),4.56 (s, 2H), 4.00 (s, 2H, CH₂OH), 3.28 (m, 4H), 2.23 (t, J = 7.1 Hz, 2H), 2.14 (t, J = 7.5 Hz, 2H), 1.49-1.62 (m, 4H), 1.25-1.34 (m, 2H); HRMS (FAB+) calcd for C₃₄H₃₇O₄N₄F₂SB (M⁺): 646.2603, found: 646.2620.

Allyl-R6G NHS ester (76). The procedure is the same as that of 72 and the crude product was purified by flash column chromatography over silica gel using $CH_3OH-CH_2Cl_2$ (1:4) as the eluent to afford 76, which was used directly in the next step.

AllyI-ROX NHS ester (81). The procedure is the same as that of 72 and the crude product was purified by flash column chromatography over silica gel using $CH_3OH-CH_2Cl_2$ (1:5) as the eluent to afford 81, which was used directly in the next step.

Allyl-Bodipy-650 NHS ester (84). The procedure is the same as that of 72 and the crude

product was purified by flash column chromatography over silica gel using $CH_3OH-CH_2Cl_2$ (1:30) as the eluent to afford **84**, which was used directly in the next step.

3'-O-allyl-dUTP-allyl-R6G (77), 3'-O-allyl-dATP-allyl-ROX (81), and 3'-O-allyldGTP-allyl-Bodipy-650 (85) were prepared using the same procedure as that of 73.

3'-O-allyl-dGTP-PC-Bodipy-FL-510 (86), 3'-O-allyl-dUTP-PC-R6G (87), 3'-O-allyldATP-PC-ROX (88) and 3'-O-allyl-dCTP-PC-Bodipy-650 (89) were prepared using similar procedures as above and in *Reference [3]*.

References

- [1] a) Ju, J.; Li, Z.; Edwards, J. R.; Itagaki, Y. US Patent 2002102586. b) Ju, J.; Li, Z.; Edwards, J. R.; Itagaki, Y. WO Patent 2002029003.
- [2] Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J. Science 1994, 264(5167), 1891-1903.
- [3] Seo, T. S.; Bai, X.; Kim, D. H.; Meng, Q.; Shi, S.; Ruparel, H.; Li, Z.; Turro, N. J.; Ju, J. Proc. Natl. Acad. Sci. USA 2005, 102(17), 5926-5931.
- [4] Ruparel, H.; Bi, L.; Li, Z.; Bai, X.; Kim, D. H.; Turro, N. J.; Ju, J. Proc. Natl. Acad. Sci. USA 2005, 102(17), 5932-5937.
- [5] Metzker, M. L. Genome Research 2005, 15(12), 1767-1776.
- [6] Mitra, R. D.; Shendure, J.; Olejnik, J.; Olejnik, E. K.; Church, G. M. Anal. Biochem. 2003, 320(1), 55-65.
- [7] a) Pleasants, J. C.; Guo, W.; Rabenstein, D. L. J. Am. Chem. Soc. 1989, 111(17), 6553-6558.
 b) Huyghues-Despointes, B. M. P.; Nelson, J. W. Biochemistry 1992, 31(5), 1476-1483.
- [8] Bi, L.; Kim, D. H.; Ju, J. J. Am. Chem. Soc. 2006, 128(8), 2542-2543.
- [9] Milton, J.; Wu, X.; Smith, M.; Brennan, J.; Barnes, C.; Liu, X.; Ruediger, S. WO Patent 2004 018497.
- [10] Pudlo, J. S.; Nassiri, M. R.; Kern, E. R.; Wotring, L. L.; Drach, J. C.; Townsend, L. B. J.
 Med. Chem. 1990, 33(7), 1984-1992.
- [11] Seela, F.; Zulauf, M. Synthesis 1996, (6), 726-730.
- [12] Seela, F.; Muth, H.-P.; Bindig, U. Synthesis 1988, (9), 670-674.
- [13] Kazimierczuk, Z.; Cottam, H. B.; Revankar, G. R.; Robins, R. K. J. Am. Chem. Soc. 1984, 106(21), 6379-6382.
- [14] Buhr, C. A.; Wagner, R. W.; Grant, D.; Froehler B. C. Nucleic Acids Res. 1996, 24(15), 2974-2980.
- [15] Robins, M. J.; Robins, R. K. J. Am. Chem. Soc. 1965, 87(21), 4934-4940.
- [16] Lin, W.; Zhang, X.; Seela, F. Hel. Chim. Acta 2004, 87(9), 2235-2244.

- [17] For introduction of silyl protective groups to alcohols and the cleavage, see: Greene, T. W.;
 Wuts, P. G. M. *Protective Groups in Organic Synthesis*, 3rd edition; Wiley-Interscience: New York, 1999 (and references therein).
- [18] For nucleophilic aromatic displacement of chloro substituent by amino, see: a) Ugarkar, B.
 G.; Castellino, A. J.; DaRe, J. S.; Ramirez-Weinhouse, M.; Kopcho, J. J.; Rosengren, S.;
 Erion, M. D. J. Med. Chem. 2003, 46(22), 4750-4760. b) Zhang, L.; Zhang, Y.; Li, X.;
 Zhang, L. Bioorg. Med. Chem. 2002, 10(4), 907-912.
- [19] For Sonogashira cross-coupling reaction of iodonucleosides with the terminal alkynes, see: a) Robins, M. J.; Vinayak, R. S.; Wood, S. G. Tetrahedron Lett. 1990, 31(26), 3731-3734. b) Seela, F.; Zulauf, M. Hel. Chim. Acta 1999, 82(11), 1878-1898. c) McKeen, C. M.; Brown, L. J.; Nicol, J. T. G.; Mellor, J. M.; Brown, T. Org. Biomol. Chem. 2003, 1(13), 2267-2275.
- [20] For general method to synthesize modified triphosphate, see: a) Lee, S.E.; Sidorov, A.; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M. Nucleic Acids Res. 2001, 29(7), 1565-1573. b) Giller, G; Tasara, T.; Angerer, B.; Mühlegger, K.; Amacker, M.; Winter, H. Nucleic Acids Res. 2003, 31(10), 2630-2635.
- [21] This part of work has been published at: Meng, Q.; Kim, D. H.; Bai, X.; Bi, L.; Turro, N. J.;
 Ju, J. J. Org. Chem. 2006, 71(8), 3248-3252.
- [22] Seela, F.; Driller, H. Nucleosides Nucleotides 1989, 8(1), 1-21.
- [23] a) Ramzaeva, N.; Seela, F. Hel. Chim. Acta 1995, 78(7), 1083-1090. b) Barnett, C. J.;
 Kobierski, M. E. J. Heterocyclic Chem. 1994, 31(5), 1181-1183.
- [24] Lüpke, U.; Seela, F. Chem. Ber. 1979, 112(10), 3432-3440.
- [25] Cocuzza, A. J. Tetrahedron Lett. 1988, 33(29), 4061-4064.
- [26] a) Ramzaeva, N.; Mittelbach, C.; Seela, F. Hel. Chim. Acta 1999, 82(1), 12-18. b) Seela, F.;
 Shaikh, K. Hel. Chim. Acta 2004, 87(6), 1325-1332. c) Seela, F.; Peng, X. J. Org. Chem.
 2006, 71(1), 81-90.
- [27] For conversion of 2-amino-6-methoxy-7-deazapurine derivatives into the corresponding

7-deazaguanine moiety by aqueous sodium hydroxide solution, see: a) Prakash, T. P.; Prhavc,
M.; Eldrup, A. B.; Cook, P. D.; Carroll, S. S.; Olsen, D. B.; Stahlhut, M. W.; Tomassini, J.
E.; MacCoss, M.; Galloway, S. M.; Hilliard, C.; Bhat, B. J. Med. Chem. 2005, 48(4),
1199-1210. b) Seela, F.; Driller, H.; Steker, H. Nucleic Acids Res. 1991, 19(2), 302-307. c)
Seela, F.; Rosemeyer, H.; Biesewig, A.; Juergens, T. Nucleosides Nucleotides 1988, 7(5-6),
581-584.

- [28] For conversion of 2-amino-6-methoxy-7-deazapurine derivatives into the corresponding 7-deazaguanine moieties by iodotrimethylsilane, see: a) Kawate, T.; Allerson, C. R.; Wolfe, J. L. Org. Lett. 2005, 7(18), 3865-3868. b) Ramasamy, K.; Imamura, N.; Robins, R. K.; Revankar, G. R. J. Heterocyclic Chem. 1988, 25(6), 1893-1898. c) Ramasamy, K.; Imamura, N.; Robins, R. K.; Revankar, G. R. Tetrahedron Lett. 1987, 28(43), 5107-5110. d) Jung, M. E.; Lyster, M. A. J. Org. Chem. 1977, 42(23), 3761-3764.
- [29] For conversion of 2-amino-6-methoxy-7-deazapurine derivatives into the corresponding 7-deazaguanine moieties by other methods, see: a) McGee, D. P. C.; Martin, J. C.; Verheyden, J. P. H. J. Heterocyclic Chem. 1985, 22(4), 1137-1140. b) Seela, F.; Kehne, A.; Winkeler, H. D. Liebs. Ann. der Chem. 1983, (1), 137-146.
- [30] Kamal, A.; Laxman, E.; Rao, N. V. Tetrahedron Lett. 1999, 40(2), 371-372.
- [31] Kang, J.-H.; Siddiqui, M. A.; Sigano, D. M.; Krajewski, K.; Lewin, N. E.; Pu, Y.; Blumberg,
 P. M.; Lee, J.; Marquez, V. E. Org. Lett. 2004, 6(14), 2413-2416.
- [32] Honma, M.; Sawada, T.; Fujisawa, Y.; Utsugi, M.; Watanabe, H.; Umino, A.; Matsumura, T.;
 Hagihara, T.; Takano, M.; Nakada, M. J. Am. Chem. Soc. 2003, 125(10), 2860-2861.
- [33] Soli, E. D.; Manoso, A. S.; Patterson, M. C.; DeShong, P. J. Org. Chem. 1999, 64(9), 3171-3177.
- [34] Jacobson, M. A.; Williard, P. G. J. Org. Chem. 2002, 67(11), 3915-3918.
- [35] Bai, X.; Kim, S.; Li, Z.; Turro, N. J.; Ju, J. Nucleic Acids Res. 2004, 32(2), 535-541.
- [36] a) Khan, S. I.; Grinstaff, M. W. J. Am. Chem. Chem. 1999, 121(19), 4704-5. b) Hobbs, F. W.
 Jr.; Cocuzza; A. J. Eur. Pat. Appl. 19880107.
Chapter 3. Application of Modified Fluorescent Nucleotides as Reversible Terminators for DNA Sequencing by Synthesis

3.1 Survey of Methods to Cleave 3' Allyl Capping Group and Allyl Linker Moiety

To apply our modified fluorescent nucleotides in DNA sequencing by synthesis successfully, an efficient method to remove both the 3' allyl group and allyl linker moiety in each cycle must be established.

An ideal method to cleave allyl groups in both the allyl ethers and allyl carbamates for the DNA sequencing system requires that: 1) the reaction must be performed in an aqueous medium due to the hydrophilic property of the DNA strand; 2) the reaction conditions must be mild enough that the DNA strand is neither damaged nor incurs any negative effects, for example the reaction cannot be performed at high temperature or in strong acid; 3) the conversion yield must be high enough to increase the potential sequencing read length of SBS as much as possible; and 4) the reaction time should be short so that the total sequencing is fast enough.

Traditional methods to remove an allyl group from allyl ethers usually involve isomerization of the allyl group to the corresponding prop-1-enyl group followed by acid cleavage. The isomerization can be achieved by a strong base such as potassium *tert*-butoxide or transition metal catalysts like rhodium or iridium.¹ More methodologies have been developed in recently years to cleave allyl groups from allyl ethers, such as the use of *N*-bromosuccinimide,² samarium (II) iodide,³ chlorotrimethylsilane and sodium iodide,⁴ ruthenium (II) complex,⁵ and zirconocene (Cp_2Zr).⁶ Most of the methods above are performed in an organic solvent or in a mixture of organic solvent and water, or under harsh conditions such as strong acid, which makes them unsuitable for the DNA sequencing system because of the susceptibility of the DNA strand.

In the past decades, chemists have developed efficient catalysts to cleave allyl groups from allyl ethers, allyl carbonates or allyl carbamates; these are composed of palladium (0) or (II) combined with suitable ligands or other reagents.¹ Some of them are performed in relatively mild conditions and may be a potential choice for our DNA sequencing system. The following are some examples:

Milton *et al* employed an aqueous palladium (II)-phosphine system to cleave the allyl group from allyl ethers in their patent, specifically hydrophilic sodium tetrachloropalladate (Na₂PdCl₄) and trisodium triphenylphosphine-3,3',3"-trisulfonate (TPPTS) ligand were used (Scheme 3-1).⁷ This reaction was performed in water at 50 °C. Monitored by HPLC, the conversion efficiency of this reaction was observed to be higher than 98% in 30 minutes.



Scheme 3-1. Milton's method to cleave a 3' allyl group using an aqueous palladium (II)-phosphine system.⁷

Genêt *et al* developed an efficient water-soluble catalyst generated from palladium (II) acetate and TPPTS. This catalyst could cleave the allyl group from allyl esters, allyl carbonates or allyl carbamates smoothly in a homogenous solution (acetonitrile / water) or in a biphasic mixture (butyronitrile / water) (Scheme 3-2).⁸ The reaction conditions were very mild, and most reactions were complete in less than 15 minutes at room temperature.

Scheme 3-2. Genêt's method to cleave an allyl group from allyl esters, allyl carbonates or allyl carbamates using an aqueous palladium-phosphine system.⁸

Thayumanavan described a strategy to cleave the allyl group from allyl ethers and allyl carbamates under basic conditions in the presence of a palladium (0) catalyst (Scheme

3-3).⁹ These reactions were usually performed in methanol at room temperature for several hours with high isolation yields.

$$\begin{array}{c} R \\ O \\ O \\ R' \\ N \\ H \end{array} \xrightarrow{Pd(PPh_3)_4 / K_2CO_3} \\ CH_3OH, 2-12 h \\ room temperature \\ 84-97\% \\ Substrate : Pd(PPh_3)_4 : K_2CO_3 = 1: 0.01 : 3~6 \end{array}$$

Scheme 3-3. Thayumanavan's strategy to cleave an allyl group from allyl esters and allyl carbamates using a palladium (0) catalyst.⁹

Smith used an excess of palladium (II) chloride in a medium of acetic acid and sodium acetate to cleave the allyl group in the following reaction during his total synthesis of (+)-phyllanthoside (Scheme 3-4).¹⁰ This reaction was performed at room temperature for 18 hours with 89% yield.



Scheme 3-4. Smith's cleavage of an allyl group by palladium (II) chloride in acetic acid and sodium acetate.¹⁰

Considering reaction conditions, solvents, conversion efficiency and reaction time, Milton's method employing sodium tetrachloropalladate and TPPTS appears to be most promising for our DNA sequencing system. Therefore a series of experiments based on Milton's method were performed. Here 3'-O-allylthymidine was used as substrate, due to its good solubility in warm water and similarity to the 3' allyl-modified nucleotides incorporated into the DNA strand. The results are shown in Table 3-1.



Table 3-1. Deallylation of 3'-O-allylthymidine by Na₂PdCl₄ and TPPTS in water.

Entry	[Substrate]	Pd	[Pd]	[TPPTS]	Temp	Converted
	mM		mM	mM	°C	yield, %*
1	3.2	Na ₂ PdCl ₄	1.6	9.7	60	80
2	3.2	Na_2PdCl_4	1.6	12.8	60	>90
3	1.6	Na ₂ PdCl ₄	0.8	6.4	60	85
4	1.6	Na ₂ PdCl ₄	0.8	6.4	70	>90
5	6.7	Na_2PdCl_4	3.3		60	50
6	6.0	PdCl ₂	3.0		60	85
7	3.0	PdCl ₂	3.0		60	>90
8	1.0	$PdCl_2$	1.0		60	90
9	3.1	$Pd(OAc)_2$	3.2		60	<15

* The yield is based on TLC analysis by comparing the spot containing the reaction mixture and commercial thymidine under UV light. >90% means the spot of substrate could barely be observed.

It can be seen from Table 3-1 that a ratio of TPPTS/ Pd of 8: 1 is necessary to maintain a high conversion yield of the substrate (entries 1 and 2). At higher temperature (70 °C), even a dilute concentration of substrate could still afford a high conversion yield (entry 4).

This is very desirable for the DNA sequencing system as the DNA amount being analyzed is usually very small. It is also noteworthy that some palladium reagents, such as Na_2PdCl_4 and $PdCl_2$, gave satisfactory deallylation results without the TPPTS ligand (entries 5 to 8), while $Pd(OAc)_2$ itself gave a low conversion yield (entry 9).

3.2 Application of Photocleavable Modified Fluorescent Nucleotides as Reversible Terminators for DNA Sequencing by Synthesis in the solution¹¹

For photocleavable fluorescent 3'-O-allyl-dNTP-PC-Dye nucleotides (**86**, **87**, **88** and **89**, Figure 3-1) to act as reversible terminators for DNA SBS, it is important to establish that they can be used to determine a repeated DNA sequence in a polymerase reaction. This requires that only one complementary nucleotide can be incorporated into the primer strand in each sequencing cycle.

For this purpose, DNA polymerase extension reactions in the solution phase using **86-89** as substrates were performed. The DNA products at each step were analyzed using matrix-assisted laser desorption / ionization time-of-flight mass spectrometry (MALDI-TOF MS) to characterize their detailed molecular structures. As can be seen in Scheme 3-5A, there is a repeating sequence of 3'-CC-5' in the template adjacent to the annealing site of the primer strand. Thus DNA polymerase extension reactions on this repeat C template in solution phase using a mixture of 3'-O-allyl-dNTP-PC-Dye nucleotides **86-89**



Figure 3-1. Structures of fluorescent 3'-O-allyl-dNTP-PC-Dye nucleotides 86-89.

as substrates were performed in the presence of 9° N_m DNA polymerase and only one molecule of **86** was incorporated per primer. After that, the fluorophore was cleaved by near UV laser irradiation efficiently (\approx 355 nm). The 3' allyl capping group was then removed in aqueous palladium (II)-phosphine solution to regenerate the 3'-hydroxyl group, so that the next cycle of nucleotide incorporation was able to proceed. The above steps constitute one complete sequencing cycle on the template.

The MALDI-TOF MS spectra in Scheme 3-5A show the sequential results at each step of

DNA SBS using 86-89 as reversible terminators. As can be seen in Figure (1), the distinct peak at m/z 7.052 corresponds to the single base extension product with 86 at 100% incorporation efficiency, confirming that the reversible terminator 86 can be incorporated base-specifically by DNA polymerase into the growing DNA strand. The small peak at m/z 6,556 resulted from the partial photocleavage caused by the nitrogen laser pulse (337) nm) used for ionization in MALDI-TOF MS. Figure (2) shows the photocleavage result after laser irradiation at 355 nm for 10 sec on the DNA extension product. Now the peak at m/z 7,052 disappeared completely, while the peak at m/z 6,556 corresponding to the photocleavage product appears as the sole peak. Figure (3) shows the deallylation result in aqueous Na₂PdCl₄-P(PhSO₃Na)₃ (1:8) solution at 70 °C for 30 sec. The single peak at m/z 6,516 corresponds to the deallylated product and the disappearance of the peak at m/z6,556 revealed that the deallylation reaction was completed with high efficiency. This regenerated the 3'-hydroxyl group for the subsequent sequencing cycle. Thus 3'-O-allyl-dGTP-PC-Bodipy-FL-510 86 has been validated as a very good reversible terminator for DNA SBS.

DNA polymerase extension reactions on other templates in the solution phase using a mixture of **86-89** as substrates were also performed. As shown in Scheme 3-5 B-D, the MALDI-TOF MS spectra show sequential results at each step of DNA SBS using **87-89** as reversible terminators, which correspond to the expected product peaks of extension, photocleavage and deallylation respectively. Therefore the photocleavable fluorescent 3'-







Scheme 3-5. Extension with 3'-O-allyl dNTP-PC-Dye on a DNA template.

Scheme 3-5 continued

C) Extension with 3'-O-allyl-dNTP-PC-Dye on repeat T template:







O-allyl-dNTP-PC-Dye nucleotides **86-89** were all proven to be good reversible terminators for DNA SBS.

3.3 Application of Photocleavable Modified Fluorescent Nucleotides as Reversible Terminators for DNA Sequencing on a Glass Chip

The photocleavable fluorescent 3'-O-allyl-dNTP-PC-Dye nucleotides **86-89** were also employed for DNA sequencing by synthesis on a glass chip to expand this technique into a more practical application. To realize this application, a necessary requirement is to construct a chip with immobilized self-primed DNA templates, so that the primer strand is stably linked to the surface. The template can be immobilized on the surface either by physical adsorption or by chemical linkage via a covalent bond. The latter has demonstrated superior advantages with regard to stability, flexibility and feasibility. In this method a functional group on the DNA template will react with another specific group on the surface to form a covalent bond. This chemical bond will immobilize the DNA template on the surface. Both of the functional groups have to be carefully selected so that the chemical bond can be easily formed with high efficiency and have sufficient stability to endure the subsequent sequencing conditions.

Prof. Ju has applied the 1,3-dipolar cycloaddition *click chemistry*¹² of an alkyne and an azide to label a single-stranded DNA with a fluorophore.¹³ This copper (I)-catalyzed

cycloaddition reaction was performed in aqueous solution with almost quantitative yields. Later, this reaction was also used to immobilize a DNA template on a glass chip for DNA SBS due to its high efficiency and mild reaction conditions.¹⁴ As shown in Scheme 3-6, a terminal azido group was first attached to the surface of a glass chip via a chemical coupling reaction. A 5' alkyne-modified DNA strand was also prepared by attaching a DNA strand to an alkyne-modified linker. This modified DNA strand was finally immobilized on the glass chip via the cycloaddition reaction between the alkyne and the azido group under the catalysis of copper (I) iodide.



Scheme 3-6. Illustrative description of immobilizing a DNA template on a glass chip via a copper (I)-catalyzed 1,3-dipolar cycloaddition reaction.¹⁴

After being immobilized with DNA templates, the glass chip was subjected to DNA polymerase extension reactions using 3'-O-allyl-dNTP-PC-Dye nucleotides **86-89** to investigate its stability and functionality. As shown in Figure 3-2, DNA polymerase extension reactions on the glass chip using a mixture of the four nucleotides **86-89** as

substrates was performed in the presence of 9° N_m DNA polymerase. One molecule of 3'-O-allyl-dGTP-PC-Bodipy-FL-510 86 that was complementary to the corresponding base on the template was first incorporated into the self-primed template, which was confirmed by observing the unique fluorescent signal using a ScanArray Express microarray scanner. After that, near UV laser irradiation was applied to the chip to cleave the fluorophore efficiently, which was confirmed by the blank fluorescent signal from the scanner. Later, the chip was treated with aqueous palladium (II)-phosphine solution to remove the 3' allyl capping group so that the hydroxyl group was regenerated on the 3' end of the template to initiate the next nucleotide incorporation. Thus the sequencing cycle could go on and on. In this sequencing system a complete sequencing cycle was composed of nucleotide incorporation, fluorophore detection, photocleavage of fluorescent dye and deallylation of the capping group. As shown in Figure 3-2, the sequencing cycles have been repeated seven times, correctly identifying seven successive bases on the DNA template. It has been proven that photocleavable 3'-O-allyl-dNTP-PC-Dye nucleotides 86-89 could be used on the glass chip to identify the sequence of a DNA strand. The relatively strong laser irradiation power used in the photocleavage did not result in any evident of DNA strand damage after these cycles. As photocleavage was performed in each sequencing cycle, any residual fluorescent dyes left from previous cycles could be cleaved completely thereafter to ensure cleaner background for the subsequent sequencing cycles. Therefore the signal / noise ratio of this sequencing method could be maintained even after repeated sequencing cycles and the read length



could be increased to an acceptable level.

Figure 3-2. Application of photocleavable modified fluorescent 3'-O-allyl-dNTP-PC-Dye nucleotides **86-89** for DNA sequencing by synthesis on a glass chip.

With current sequencing efficiency and further optimization in the near future, this sequencing technique is expected to sequence more than 25 bases on the glass chip using our modified fluorescent nucleotides, and will be potentially developed into a highly efficient DNA analysis system for DNA sequencing.

3.4 Application of Chemically Cleavable Modified Fluorescent Nucleotides as Reversible Terminators for DNA Sequencing by Synthesis in the solution Chemically cleavable fluorescent nucleotide 3'-O-allyl-dGTP-allyl-Bodipy-FL-510 **90** was first synthesized and tested on a repeat C template to validate its properties as a reversible terminator for DNA SBS.

For this purpose a DNA polymerase extension reaction on a repeat C template in solution phase using 90 as substrate was performed in the presence of 9° N_m DNA polymerase, as shown in Scheme 3-7, and as a result only one molecule of 90 was incorporated per template. After that, the fluorescent dye moiety was removed in aqueous palladium (II)-phosphine solution along with the 3' allyl capping group in a single step. Thus, the original two steps of fluorescent dye cleavage and 3' capping group removal required when using 3'-O-allyl-dNTP-PC-Dye nucleotides as terminators were combined in one step. The efficiency of the entire sequencing cycle was thereby improved and the speed of SBS was greatly increased. At this point the free 3'-hydroxyl group had been regenerated for the next cycle of nucleotide incorporation. This DNA strand was then used as a primer in the subsequent DNA polymerase extension reaction in the presence of 9° N_m DNA polymerase to incorporate another molecule of 90. The fluorescent dye and 3' allyl capping group were again removed together in aqueous palladium (II)-phosphine solution to regenerate the 3'-hydroxyl group. Thus the above steps constitute two complete sequencing cycles on the template resulting in a two-base-extended primer ready for the third cycle of nucleotide incorporation.

177

The MALDI-TOF MS spectra in Scheme 3-7 show the sequential results at each step of DNA SBS using 90 as a reversible terminator. As can be seen in Figure (1), the distinct peak at m/z 6,953 corresponds to the product of a single base extension with 100% incorporation efficiency, confirming that only one molecule of 90 was incorporated specifically by DNA polymerase into the primer strand. Figure (2) shows the deallylation result in aqueous Na₂PdCl₄-P(PhSO₃Na)₃ (1:8) solution at 70 °C for 30 sec. The single peak at m/z 6,498 confirms that both the fluorescent dye and 3' allyl capping group were cleaved together and the disappearance of the peak at m/z 6,953 proved that the deallylation reaction was completed efficiently. In Figure (3), the peak at m/z 7,334 indicates that another molecule of 90 was incorporated specifically into the primer strand. The single peak at m/z 6,879 in Figure (4) confirms that both the fluorescent dye and 3' allyl capping group were cleaved completely, and the 3'-hydroxyl group was regenerated for subsequent sequencing cycles. These two complete sequencing cycles validated the chemically cleavable fluorescent 3'-O-allyl-dGTP-allyl-Bodipy-FL-510 nucleotide 90 as a very good reversible terminator for DNA SBS.

The above results demonstrating **90** to be a good reversible terminator for DNA SBS encouraged us to synthesize a full set of chemically cleavable modified fluorescent 3'-*O*-allyl-dNTP-allyl-Dye nucleotides (**73**, **77**, **81** and **85**, Figure 3-3) and they too were applied in DNA polymerase extension reactions on a DNA template to validate their properties as reversible terminators for DNA SBS.

allyl-Bodipy-FL-510 Pd (II) / ligand Primer 5'- GTTGATGTACACATTGTCAA 90 -3' 5 -allyl -3' G (1) Template 3'- --CCTG -5' 9° N_m DNA polymerase 3'-CCTG -5 Deallylation Extension H₂N NH₂ H₂N allyl-Bodipy-FL-510 NH₂ •**G**-OH Pd (II) / ligand 90 5'-5'--GG-allyl -3' 5'--ĠĠ-он -з' -3 3'--CCTG -5' 9° N_m DNA polymerase -CCTG -5' 3'--CCTG -5' 3'-Deallylation Extension (2) (3) (4) ÇH₃ 5 Primer (1) 6953 Calcd: 6952 (1) Extension Ó H_2N product 0 ò -NH₂ 5 Primer (2) 6512 Calcd: Deallylated (2) 6511 Hal product a Intensity NH₂ Primer (3) 7334 Extension product 0= Calcd: 7332 H₂N (3) 6893 (4) ń ì 5 Deailylated Primer product NH2 Ć H₂N 0=1 Calcd: 5200 6200 7200 200 100 6891 mh (4) H₂I 3'-O-allyl-dGTP-allyl-Bodipy-FL-510 90 ó òσ ò-ố ò-Ó

Extension with 3'-O-allyI-dGTP-allyI-Bodipy-FL-510 90 on repeat C template:

Scheme 3-7. Extension with 3'-O-allyl dGTP-allyl-Bodipy-FL-510 90 on a repeat C template.



Figure 3-3. Structures of fluorescent 3'-O-allyl-dNTP-allyl-Dye nucleotides 73, 77, 81 and 85.

As shown in Scheme 3-8A, DNA polymerase extension reaction on a repeat G template in solution phase using a mixture of 73, 77, 81 and 85 as substrates was performed in the presence of 9° N_m DNA polymerase and only one molecule of 73 was incorporated specifically. After that, both the fluorescent dye moiety and 3' allyl capping group were removed in aqueous palladium (II)-phosphine solution in one step to regenerate the free 3'-hydroxyl group. Thus the next cycle of nucleotide incorporation was ready to proceed.

180

The MALDI-TOF MS spectra in Scheme 3-8A show the sequential results at each step of DNA SBS using **73** as a reversible terminator. As can be seen in Figure (1), the distinct peak at m/z 5,935 corresponds to the product of single base extension with 100% incorporation efficiency, confirming that only one molecule of **73** was incorporated specifically by DNA polymerase into the primer strand. Figure (2) shows the deallylation result in aqueous Na₂PdCl₄-P(PhSO₃Na)₃ (1:8) solution at 70 °C for 30 sec. The single peak at m/z 5,489 confirms that both the fluorescent dye and 3' allyl capping group were cleaved together and the disappearance of the peak at m/z 5,935 proved that the deallylation reaction was completed efficiently. This complete sequencing cycle validated 3'-*O*-allyl-dCTP-allyl-Bodipy-FL-510 **73** as a very good reversible terminator for DNA SBS.

DNA polymerase extension reactions on repeat A, C and T templates in solution phase using a mixture of **73**, **77**, **81** and **85** as substrates were also performed. As can be seen in Scheme 3-8B-D, the MALDI-TOF MS spectra show sequential results at each step of DNA SBS using **73**, **77**, **81** and **85** as substrates that correspond to the product peaks of extension and deallylation respectively. Both the fluorescent dye and the 3' capping group were cleaved together in aqueous palladium (II)-phosphine solution conveniently. Therefore the modified fluorescent 3'-O-allyl-dNTP-allyl-Dye nucleotides **73**, **77**, **81** and **85** have been proven to be good reversible terminators for DNA SBS. The efficiency of an entire sequencing cycle has been improved and the speed and sequence read length of

SBS obviously could be increased due to the simplified sequencing steps.



A) Extension with 3'-O-allyl-dNTP-allyl-Dye on a repeat G template:

B) Extension with 3'-O-allyl-dNTP-allyl-Dye on a repeat A template:



Scheme 3-8. Extension with 3'-O-allyl dNTP-PC-Dye on a DNA template.

Scheme 3-8 continued

C) Extension with 3'-O-allyl-dNTP-allyl-Dye on a repeat T template:



D) Extension with 3'-O-allyl-dNTP-allyl-Dye on a repeat C template:



3.5 Proposed Photocleavage and Deallylation Mechanisms of Modified Fluorescent Nucleotides in DNA Sequencing by Synthesis

The mechanism of photolysis of 2-nitrobenzyl linker moiety in the modified fluorescent nucleotides **86-89** is proposed to proceed via a *Norrish-type II* reaction, as shown in Scheme 3-9.¹⁵ The nitro group in the 2-nitrobenzyl linker is first excited by the UV light to form a free radical, which in turn pursues a γ -hydrogen abstraction to generate an intermediate. This intermediate is then converted to a five-membered ring, which finally produces a carboxylic acid and a 2-nitrosoacetophenone as final products.



Scheme 3-9. Photolysis mechanism of 2-nitrobenzyl linker moiety.¹⁵

The deallylation process in aqueous palladium (II)-phosphine solution proceeds in a more complicated manner. Up to now there has been no clear explanation of its mechanism. J. P. Genêt described the cleavage of the allyl group from allyl esters, allyl carbonates and allyl carbamates in a homogenous solution (acetonitrile / water) or in a biphasic mixture (butyronitrile / water) and he also postulated a mechanism for these reactions.⁸ Based on this mechanism of allyl group cleavage, we propose the deallylation mechanism of removing allylic fluorophores from nucleotide molecules in aqueous palladium (II)-phosphine solution shown in Scheme 3-10A. The essential process of this mechanism



Scheme 3-10. Proposed mechanism of cleaving allyl groups from allyl ethers and allyl carbamates in aqueous palladium-phosphine solution.⁸

involves the formation of a π -allyl complex between palladium and allylic cation via oxidative addition. The other part of the nucleotide molecule is then hydrolyzed to form an amine and a carbon dioxide molecule. Later, an allyl scavenger (here it is hydroxide anion) attacks the π -allyl complex to form an allylic alcohol and the palladium catalyst is then regenerated for the next catalytic cycle. In this way the allylic carbamate group in the fluorescent nucleotide molecules is hydrolyzed under the catalysis of the palladium catalyst.

The removal of the 3' allyl capping group was assumed to proceed via a similar pathway, as shown in Scheme 3-10B. A π -allyl complex is first formed via oxidative addition, which then reacts with hydroxide anion to form allyl alcohol and the palladium catalyst is then regenerated for the next catalytic cycle. In this way the allylic ether in the nucleotide molecules is hydrolyzed under the catalysis of the palladium catalyst.

3.6 Experimental Section

General Information. ¹H NMR spectra were recorded on a Bruker DPX-400 (400 MHz) spectrometer and are reported in ppm from CD₃OD internal standard (3.31 ppm). Data are reported as follows: (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublets, ddd = doublet of doublets of doublets; coupling constant(s) in Hz; integration; assignment). Proton decoupled ¹³C NMR spectra were recorded on a Bruker

DPX-400 (100 MHz) spectrometer and are reported in ppm from CDCl₃ internal standard (77.0 ppm). High Resolution Mass Spectra (HRMS) were obtained on a JEOL JMS HX 110A mass spectrometer. All other chemicals were purchased from Sigma-Aldrich. Mass measurement of DNA was made on a Voyager DE matrix-assisted laser desorption / ionization time-of-flight (MALDI-TOF) mass spectrometer (Applied Biosystems). Photolysis was performed at 340 nm (10 nm band pass) by using a 450 W high-pressure xenon lamp (Thermo Oriel, Stratford, CT) in conjunction with a 340 nm UV interference filter (CVI Laser, Albuquerque, NM) with a light intensity of 6 mW/cm² or at 355 nm using a Spectra Physics GCR-150-30Nd-YAG laser (ca. 50 mJ/ pulse, pulse length ca. 7 ns, 30 Hz) with a light intensity of 1.5 mW/cm². The scanned fluorescence images were obtained by using a ScanArray Express scanner (Perkin Elmer Life Sciences), which was equipped with four lasers with excitation wavelengths of 488, 543, 594 and 633 nm and emission filters centered at 522, 570, 614 and 670 nm.

3'-O-Allylthymidine (91). To a stirred solution of 5'-O-(4, 4'-dimethoxytrityl)thymidine (1.00 g; 1.84 mmol) in THF (29 mL) were added NaOH powder (180 mg; 4.5 mmol), allyl bromide (179 μ L; 2.07 mmol) and tris[2-(2-methoxyethoxy)ethyl]amine (TDA-1) (121 μ L; 0.359 mmol). The reaction mixture was stirred at room temperature for 10 h. Most of THF was evaporated and the residue was dissolved in ethyl acetate (100 mL). The solution was washed with saturated aqueous NaHCO₃ and NaCl respectively, and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified

187

by flash column chromatography over silica gel using ethyl acetate-hexane (1:1) as the eluent to afford a white solid (947 mg; 88% yield). The product was then dissolved in CH₂Cl₂ (16 mL) and trifluoroacetic acid (1.60 mL; 20.8 mmol) was added. The reaction mixture was stirred at room temperature for 20 min. Most of the CH₂Cl₂ was evaporated and the residue was purified by flash column chromatography over silica gel using ethyl acetate-hexane (4:1) as the eluent to afford 91 as a white solid (407 mg; 89% yield): ¹H NMR (400 MHz, CD₃OD) δ 7.79 (m, 1H, 6-H), 6.23 (dd, J = 6.0, 8.0 Hz, 1H, 1'-H), 5.88-5.98 (m, 1H, $CH_2CH=CH_2$), 5.27-5.33 (dm, J = 17.2 Hz, 1H, one of $CH_2CH=CH_2$), 5.14-5.19 (dm, J = 10.4 Hz, 1H, one of CH₂CH=CH₂), 4.19 (m, 1H, 3'-H), 3.99-4.10 (m, 3H, 4'-H and CH₂CH=CH₂), 3.78 (dd, J = 3.5, 12.0 Hz, 1H, one of 5'-H), 3.72 (dd, J =3.5, 12.0 Hz, 1H, one of 5'-H), 2.32-2.39 (ddd, J = 2.4, 5.9, 13.6 Hz, 1H, one of 2'-H), 2.12-2.21 (ddd, J = 6.2, 8.0, 13.6 Hz, 1H, one of 2'-H), 1.88 (d, J = 1.2 Hz, 3H, CH₃); ¹³C NMR (100 MHz, CDCl₃) δ 163.4, 150.1, 136.9, 133.9, 117.4, 111.0, 87.3, 85.1, 78.5, 70.5, 62.8, 37.2, 12.7; HRMS (FAB+) calcd for $C_{13}H_{19}O_5N_2$ (M+H⁺): 283.1294, found: 283.1292.

General deallylation procedure of 3'-O-allylthymidine by Na₂PdCl₄ and TPPTS in water. Entry 3 (Table 3-1) is given as an example: Na₂PdCl₄ (4.7 mg; 16.0 μ mol) and TPPTS (72.8 mg; 128 μ mol) were dissolved in deionized water (20 mL) and stirred at room temperature for 5 min. 3'-O-Allylthymidine (9.0 mg; 31.9 μ mol) was added and the reaction mixture was stirred at 60 °C for 30 min. The converted percentage of

3'-O-allylthymidine was measured using TLC analysis by comparing the reaction mixture with the starting material and commercial thymidine. Ethyl acetate (50 mL) was added to the reaction mixture, and the solution was washed with saturated aqueous NaCl and dried over anhydrous Na₂SO₄. After evaporation of the solvent, the residue was purified by flash column chromatography over silica gel using CH₃OH-CH₂Cl₂ (1:6) as the eluent to afford thymidine as a white solid.

General procedure of primer extension by using 3' modified fluorescent nucleotides and photocleavage of the extension product. The procedure for 86 is given as an example: the polymerase extension reaction mixture consisted of 60 pmol of primer (5'-GTTGATGTACACATTGTCAA-3'), 80 pmol of 100-mer template (5'-TACCCGGA GGCCAAGTACGGCGGGTACGTCCTTGACAATGTGTACATCAACATCACCTACC ACCATGTCAGTCTCGGTTGGATCCTCTATTGTGTCCGGG-3'), 120 pmol of 3'-Oallyl-dGTP-PC-Bodipy-FL-510 86, 1X Thermopol reaction buffer (20 mM Tris-HCl/ 10 mM (NH₄)₂SO₄/10 mM KCl/2 mM MgSO₄/0.1 % Triton X-100, pH 8.8, New England Biolabs), and 6 units of 9° N_m Polymerase (exo-)A485L/Y409V (generously supplied by New England Biolabs) in a total volume of 20 μ L. The reaction consisted of 20 cycles at 94 °C for 20 sec, 46 °C for 40 sec, and 60 °C for 90 sec. After the reaction, a small portion of the DNA extension product was desalted using a ZipTip and analyzed by MALDI-TOF MS, and showed a dominant peak at m/z 7,052 corresponding to the DNA extension product (Scheme 3-5A, figure (1)). The rest of the product mixture was

freeze-dried, resuspended in 200 μ L of deionized water, and irradiated at 355 nm for 10 sec to cleave the fluorophore from the DNA to yield the photocleavage product and then analyzed by MALDI-TOF MS (Scheme 3-5A, figure (2)).

General deallylation procedure of DNA extension products. The DNA product (20 pmol) was added to a mixture of degassed 1X Thermopol reaction buffer (20 mM Tris-HCl/10 mM (NH₄)₂SO₄/10 mM KCl/2 mM MgSO₄/0.1 % Triton X-100, pH 8.8, 1 μ L), Na₂PdCl₄ in degassed H₂O (7 μ L, 23 nmol) and P(PhSO₃Na)₃ in degassed H₂O (10 μ L, 176 nmol) to perform deallylation. The reaction mixture was then placed in a heating block and incubated at 70 °C for 30 seconds to yield quantitatively deallylated DNA product which was analyzed by MALDI-TOF MS.

References

- [1] For cleavage of allyl group from an allyl ether or allyl carbamate, see: Greene, T. W.; Wuts, P.
 G. M. Protective Groups in Organic Synthesis, 3rd edition; Wiley-Interscience: New York, 1999 (and references therein).
- [2] Diaz, R. R.; Melgarejo, C. R.; López-Espinosa, M. T. P.; Cubero, I. I. J Org. Chem. 1994, 59(25), 7928-7929.
- [3] Dahlén, A.; Sundgren, A.; Lahmann, M.; Oscarson, S.; Hilmersson, G. Org. Lett. 2003, 5(22), 4085-4088.
- [4] Kamal, A.; Laxman, E.; Rao, N. V. Tetrahedron Lett. 1999, 40(2), 371-372.
- [5] Tanaka, S.; Saburi, H.; Ishibashi, Y.; Kitamura, M. Org. Lett. 2004, 6(11), 1873-1875.
- [6] Ito, H.; Taguchi, T.; Hanzawa, Y. J Org. Chem. 1993, 58(3), 774-775.
- [7] Milton, J.; Wu, X.; Smith, M.; Brennan, J.; Barnes, C.; Liu, X.; Ruediger, S. WO Patent 2004 018497.
- [8] For Genêt's work on cleavage of allyl group, see: a) Genêt, J. P.; Blart, E.; Savignac, M.; Lemeune, S.; Paris, J.-M. *Tetrahedron Lett.* 1993, 34(26), 4189-4192. b) Lemaire-Audoire, S.; Savignac, M.; Blart, E.; Pourcelot, G.; Genêt, J. P. *Tetrahedron Lett.* 1994, 35(47), 8783-8786. c) Genêt, J. P.; Blart, E.; Savignac, M.; Lemeune, S.; Lemaire-Audoire, S.; Paris, J.-M.; Bernard, J.-M. *Tetrahedron* 1994, 50(2), 497-503. d) Lemaire-Audoire, S.; Savignac, M.; Blart, E.; Bernard, J.-M.; Genêt, J. P. *Tetrahedron Lett.* 1997, 38(17), 2955-2958. e) Lemaire-Audoire, S.; Savignac, M.; Pourcelot, G.; Genêt, J. P.; Bernard, J.-M. *J. Mol. Catal A : Chemical* 1997, 116(1-2), 247-258.
- [9] Vutukuri, D. R.; Bharathi, P.; Yu, Z.; Rajasekaran, K.; Tran, M.-H.; Thayumanavan, S. J Org. Chem. 2003, 68(3), 1146-1149.
- [10] Smith, A. B.; Rivero, R. A.; Hale, K. J.; Vaccaro, H. A. J. Am. Chem. Soc. 1991, 113(6), 2092-2112.
- [11] This part of the work has been published at: Meng, Q.; Kim, D. H.; Bai, X.; Bi, L.; Turro, N. J.; Ju, J. J. Org. Chem. 2006, 71(8), 3248-3252.

- [12] For click chemistry, see: a) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Angew. Chem. Int. Ed.
 2001, 40(11), 2004-2021. b) Li, Z.; Seo, T. S.; Ju, J. Tetrahedron Lett. 2004, 45(15), 3143-3146.
- [13] Seo, T. S.; Li, Z.; Ruparel, H.; Ju, J. J. Org. Chem. 2003, 68(2), 609-612.
- [14] a) Ju, J.; Li, Z.; Edwards, J. R.; Itagaki, Y. US Patent 2002102586. b) Ju, J.; Li, Z.;
 Edwards, J. R.; Itagaki, Y. WO Patent 2002029003.
- [15] Bochet, C. G. J. Chem. Soc., Perkin Trans. 1 2002, (2), 125-142.

Photocleavable and chemically cleavable 3' modified fluorescent nucleotides have been synthesized in support of this thesis. They have been applied for DNA sequencing by synthesis both in solution phase and on a glass chip and were proven to be good reversible terminators for DNA SBS. Among them, the chemically cleavable fluorescent nucleotides demonstrated great advantages for DNA SBS as both the fluorescent dye moiety and the 3' capping group could be cleaved together in aqueous palladium (II)-phosphine solution conveniently after fluorophore detection. Thus with the smaller number of sequencing steps, the efficiency of an entire sequencing cycle has been improved and the sequencing speed and read length of SBS could potentially be increased. This valuable approach has offered an alternative DNA analysis method to current DNA sequencing techniques. All the modified fluorescent nucleotides described in this thesis were shown to be good substrates for at least some DNA polymerases and they could be accurately incorporated into a growing DNA strand in the DNA polymerase reaction.

The photocleavable 3' modified fluorescent nucleotides have been successfully applied on the glass chip-based DNA sequencing system to identify the sequence the identity of a DNA strand. By running this method on a solid surface, certain disadvantages and limitations of conventional electrophoresis-based sequencing could be overcome. The identity of each nucleotide is recognized by observing the unique fluorescent signal with a scanner right after it is incorporated into a DNA strand. Four color fluorescent dyes (Bodipy-FL- 510, R6G, ROX and Bodipy-650) were selected to label the four nucleotide molecules respectively to produce unique fluorescent signals for nucleotide detection.

Application developments using chemically cleavable 3' modified fluorescent nucleotides on the glass chip-based DNA sequencing system are in progress now. With these fluorescent nucleotides employed on the solid surface, the fluorescent reporter and the 3' capping group can be cleaved together in a convenient chemical method after fluorophore detection. Thus an entire sequencing cycle is simplified. This work is expected to introduce a new massively parallel DNA sequencing technique to the DNA analysis.

Up to now there are still several factors that could be improved in our current massively parallel DNA sequencing technique. One of them is the 3' capping group, as discussed in Chapter 2. A highly effective capping group plays a critical role in the DNA sequencing technique. The prerequisite is that the capped modified nucleotide must be recognized by a polymerase so that it can be incorporated into a DNA strand with high efficiency. The capping group should be cleaved quickly under mild conditions to avoid damage to the DNA strand and prevent any negative effects on subsequent steps. Right now the allyl group is the most suitable capping group among those we have explored. It has been verified in this thesis that a nucleotide molecule with a 3' allyl group and a fluorescent dye attached to the base via a photocleavable or chemically cleavable linker, can be efficiently incorporated into a DNA strand in a DNA polymerase reaction. The removal of the allyl group could be completed effectively in an aqueous palladium (II)-phosphine solution in 30 sec. However, it is still worth pursuing the feasibility of using other reversible 3' capping groups, such as a photocleavable 2-nitrobenzyl group, as photolysis of a capping group is usually more desirable because it is fast, mild, easy to control, and can be conducted without the addition of other reagents that need subsequent purification. Along with a photocleavable linker attached to the dyes, this would be equally or even more efficient than the nucleotides with chemically cleavable (e.g. allyl) groups at both positions.

Other factors that can be improved in the massively parallel DNA sequencing process include the technique for immobilizing the DNA templates to the glass chip, the polymerase extension reaction and deallylation processes themselves, etc. Optimization of all these experimental protocols will ultimately improve the sequencing efficiency and increase the sequencing speed and read length.

With the sequencing technique we have already developed and expected further optimization in the near future, our DNA sequencing by synthesis system using the modified fluorescent nucleotides is expected to be developed into a highly efficient DNA analysis system for DNA sequencing that can overcome the limitations of current sequencing techniques.

Appendix: ¹H and ¹³C NMR Spectra

Р	ar	t	Ι





































































































Т



¥

ĊH₃

(Z)**-10**





239

2. S











































~

Part II























-6.461 -6.624	-11.024	-21,645 -21,805
V	V	Ŵ


























Part II





.







