COMPUTATIONAL AND SYNTHETIC STUDIES ON ANTIMETABOLITES FOR ANTICANCER-, ANTIVIRAL-, AND ANTIBIOTIC DRUG DISCOVERY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By
Rohit Tiwari, M.S.
Graduate Program in Pharmacy

The Ohio State University
2010

Dissertation Committee:
Dr. Werner Tjarks, Advisor
Dr. Pui-Kai Li
Dr. Karl Werbovetz
Dr. Robert Baiocchi
ABSTRACT

Thymidine kinase (TK) and dihydrofolate reductase (DHFR) are two enzymes that play major roles in the salvage- and de novo pathways of DNA synthesis in many eukaryotes, prokaryotes, and viruses. Both enzymes have been exploited extensively as molecular targets for the therapy and imaging of cancer and infectious diseases. This thesis deals with three different research projects that are related to these two enzymes.

The first project focuses on the development of novel computational techniques for the structure-based design of carborane-containing therapeutics. These new methods were developed at the example of DHFR-targeting carboranyl antifolates using AutoDock, Glide, FlexX, and Surflex as software platforms. Applying these new techniques, AutoDock and Glide were found to be superior to Surflex and FlexX in docking of closo-carboranyl antifolates into the active site of hDHFR. In the case nido-carboranyl antifolates, AutoDock, Glide, and Surflex produced more accurate docking results than FlexX.

The second project describes the design and synthesis of a novel type of tumor imaging agents. These are 3-carboranyl thymidine analogues (3CTAs) that are labeled with iodine-125 at the carborane cluster. 3CTAs are substrates of human TK1 (hTK1) and are selectively entrapped in tumor cells through phosphorylation by this enzyme. The advantage of radiolabeling at the carborane cluster in these compounds is that boron-iodine bonds are apparently far less susceptible to cleavage than carbon-iodine bonds in conventional imaging agents, which may reduce the undesired accumulation of radioactive iodine in the thyroid and stomach. The key step in the synthesis of a $^{125}$I-
labelled 3CTA designated as N5-\textsuperscript{125}I was a palladium catalyzed isotope exchange of cold iodine-127 with hot iodine-125 using commercially available Na\textsuperscript{125}I as the radiohalogen source. Two different methods were developed for the synthesis of \textsuperscript{127}I-labelled N5 (N5-\textsuperscript{127}I), the starting material in this reaction.

The third project deals with the design and synthesis of novel agents for a “combination therapy” of malignancies that are associated with the Epstein-Barr Virus (EBV). These include Burkitt’s lymphoma, Hodgkin’s disease, nasopharyngeal carcinoma, T/NK-cell lymphoma, and post-transplant lymphoproliferative disease. The agents were designed to selectively target EBV thymidine kinase (EBVTK) following induction of the lytic cycle of the virus within the proliferating cells. The EBV lytic cycle can be induced e.g. by DNA methyltransferase inhibitors, histone deacetylase (HDAC) inhibitors, and various chemotherapeutic agents. A small focused library of 24 compounds was synthesized for future evaluation in enzyme assay with hTK1 and EBVTK.
Dedicated

to

Trupti & my parents
ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor, Dr. Werner Tjarks, for his guidance, motivation, and support. He not only advised me in my research but also honed my presentation and teaching skills. I appreciate his thorough reading and editing of this thesis and other publications.

I thank my committee members, Drs. Pui-Kai Li, Karl Werbovetz, and Robert Baiocchi, for their valuable time, advice, and patience for addressing my graduate curricular needs. I would like to thank the faculty of the Division of Medicinal Chemistry & Pharmacognosy, especially Dr. Michael Darby and Dr. Chenglong Li, for advising me on the radiochemical and computational aspects of my thesis. I wish to thank Dr. Robert Curley for his willingness and enthusiasm to lend scientific insights on a variety of research problems.

The progression of my graduate education was made smoother by invaluable contributions of Kathy Brooks, John Fowble, Edward Beranek, Bruce Posey, Casey Hoerig, and James Logan. They were always there for me whenever I needed them the most.

I wish to express my thanks to our collaborators, Drs. Staffan Eriksson and Elena Sjouvarsson at the Swedish University of Agricultural Sciences in Uppsala, Sweden, and Drs. Robert Baiocchi and Fengting Yan at the Comprehensive Cancer Center of The Ohio State University. Special thanks go to Dr. Antonio Toppino for collaboration related to the research work described in Chapter 3, and Kiran Mahasenan and Ryan Pavlovicz for collaboration related to the research work described in the Chapter 2 of this thesis. The x-
ray crystallography work described in this thesis would not have been possible without a lending hand of Dr. Judith Gallucci of the Department of Chemistry at The Ohio State University.

I am grateful to my former and current lab members for their help in numerous matters. Especially, I would like to mention Drs. Youngjoo Byun, Achintya Bandyopadhyaya, Jayaseharan Johnsamuel, and Sureshbabu Narayanasamy for sharing their thoughts and insights on a variety of scientific matters, Dr. Hitesh Agarwal for the active collaboration related to the molecular biology studies and carrying out mass spectrometry studies on some of the compounds described in this thesis, and Sherifa Hasabelnaby for her careful and detailed proofreading of this thesis. Special thanks also go to Naum Sayfullin, Sara Jordan, Trang Nguyen, Minchen Shi, Melea Ward, Saffiatu Sillah, Kevin Ricks, and Sina Mahaboubi. I would like to acknowledge the wonderful company of fellow graduate students, especially Ye Deng, Mark Bahar, Jian Yang, Deepak Bhasin, and Kiran Mahasenan.

I am indebted to my parents, who inculcated the values and beliefs that shaped and groomed my personality, my parents-in-law, and many family members in India for their support and encouragement. I am blessed to have been showered with the love and affection of my wife, Trupti and our daughter, Ambar. Their selfless love and affection not only made me a stronger and mature human being but also provided me with a very sturdy and rock solid buffer, which maintained and refreshed my positive outlook and enthusiasm in difficult times.
VITA

June 24, 1980 ...........................................Born – Shujalpur (M.P), India

2001......................................................................B. Pharmacy, University of Pune, India

2001-2002 ...................................................M.S., NIPER, Chandigarh, India

2003-2004....................................................Graduate Teaching Associate,

The University of Oklahoma

2004-2010........................................................Graduate Teaching and Research Associate,

The Ohio State University

PUBLICATIONS


**FIELD OF STUDY**

Major Field: Pharmacy

Specialization: Medicinal Chemistry & Pharmacognosy
TABLE OF CONTENTS

1) Page

Abstract ............................................................................................................................... ii

Dedication .......................................................................................................................... iv

Acknowledgments............................................................................................................... v

Vita.................................................................................................................................... vii

List of tables..................................................................................................................... xiv

List of figures.................................................................................................................... xv

List of schemes ................................................................................................................. xx

Abbreviations.................................................................................................................. xxii

Chapters:

1

1.1 Introduction ..................................................................................................................... 2
1.2 Nucleotide biosynthetic pathways ................................................................................... 3
1.3 dNKs (2’-deoxyribonucleoside kinases/nucleoside kinases) ......................................... 4
1.4 Substrate specificity of mammalian dNKs ................................................................ 6
1.5 Substrate specificity of viral dNKs .............................................................................. 10
1.6 The importance of dNKs in the therapy of cancer and infectious diseases ............... 12
   1.6.1 dNKs in boron neutron capture therapy (BNCT) ............................................. 12
   1.6.2 Other therapeutic anticancer applications of dNKs ....................................... 15
   1.6.3 The role of dNKs in the development of antiviral therapeutics ....................... 16
1.7 The importance of dNKs in the prognosis and diagnosis of cancer and infectious diseases ................................................................. 17
1.8 The “lasso loop” in TK1-like kinases: Impact of protein flexibility on drug design ........... 18
1.9 Dihydrofolate reductase (DHFR) .............................................................................. 26
2. A comparative docking evaluation of carborane clusters using Autodock, Flexx, Glide and Surflex

2.1 Introduction ................................................................................................................. 31
2.2 Docking ......................................................................................................................... 34
2.2.1 Docking algorithms ................................................................................................. 34
2.2.2 Crystal structures ..................................................................................................... 36
2.2.3 Ligand construction and optimization ....................................................................... 36
2.2.4 Ligand preparation .................................................................................................... 37
2.2.5 Docking comparison .................................................................................................... 38
2.2.6 Scoring ......................................................................................................................... 38

2.3 Results and Discussion ................................................................................................. 39
2.3.1 Binding interactions of 1 C-prot and 2 N-prot .......................................................... 39
2.3.2 Docking analysis for docking with Autodock ............................................................. 41
2.3.3 Docking analysis for docking with Glide ................................................................. 46
2.3.4 Docking analysis for docking with FlexX ................................................................. 48
2.3.5 Docking analysis for docking with Surflex ............................................................... 51
2.3.6 Binding energy calculations of the docked poses (scoring) ........................................ 55

2.4 Conclusions and future outlook ..................................................................................... 56

3.1 Monoradioiodination of carboranyl thymidine analogues for the imaging and treatment of cancer: A case study of B-I bond formation with a controlled level of iodination and regioselectivity

3.1.1 Introduction ................................................................................................................. 61
3.1.2 Synthetic strategies for the iodination of N5 ................................................................ 66
3.1.3 Chemistry .................................................................................................................... 70
3.1.3.1 Direct iodination of preformed N5 (3.1) [strategy 1] .............................................. 70
3.1.3.2 Iodination of unsubstituted α-carborane, followed by the attachment of the iodination of α-carborane to the thymidine scaffold [strategy 2] .............................................. 77
3.1.4 NMR spectroscopic and x-ray crystallographic analysis of various halogenated species described in Section 3.1.3 ................................................................................. 86
3.1.4.1 1H NMR analysis ................................................................................................ 86
3.1.4.2 11B-NMR analysis .............................................................................................. 88
3.1.4.3 X-ray crystallographic analysis .......................................................................... 92

3.1.5 Halogen exchange reaction ......................................................................................... 93
3.1.6 Isotope exchange reaction of N5-I (3.2 a / 3.2 b) with Na\textsuperscript{125}I ........................................ 99
3.1.7 Summary and future direction .................................................................................... 102

3.2 A General Study of the deglycosylation and anomeration of N3-substituted thymidine and uridine derivatives caused by electrophilic iodine (I\textsuperscript{+})

3.2.1 Introduction ............................................................................................................... 103
3.2.2 Importance / Significance of α-nucleosides, α-nucleotides and α-oligonucleotides in medicinal chemistry .......................................................... 103
3.2.3 Importance / Significance of N3-substituted thymine nucleobases and nucleosides in medicinal chemistry .................................................. 104
3.2.4 Result and Discussion.................................................................................. 106
  3.2.4.1 Investigation of factors impacting deglycosylation and anomerization of \( \beta \)-pyrimidine nucleosides by iodine species .................................................. 106
  3.2.4.2 Hypothetical mechanism of deglycosylation and anomerization of 3.24 .................................................. 111
  3.2.4.3 The effects of various modifications at the ribose portion of 3.24 on the deglycosylation and anomerization by ICl .................................................. 112
3.2.5 Applications for the developed deglycosylation and anomerization strategies .................................................. 118
3.2.6 NMR analysis of compound 3.24 and 3.26 .................................................. 119

4. The synthesis of novel N3-substituted D- and L-dioxolanyl pyrimidine nucleoside analogues to utilize thymidine kinase as a molecular target for the treatment of Epstein-Barr virus (EBV) associated cancers
  4.1 Introduction .................................................................................................. 122
  4.2 EBV life cycle ............................................................................................. 122
  4.3 Association of EBV with B-cell malignancies ............................................. 127
  4.4 The lytic phase and associated genes and gene products .............................. 129
  4.5 Novel therapies for EBV associated malignancies ........................................ 133
  4.6 Problems associated with current therapies directed towards the lytic induction EBV positive tumors .................................................. 135
  4.7 Homology modeling of EBVTK .................................................................... 138
    4.7.1 The substrate binding site and the catalytic site of EHV4TK ......................... 139
    4.7.2 EBVTK homology model ...................................................................... 141
  4.8 Design of substrate and/or inhibitor of EBVTK ............................................. 144
  4.9 Chemistry .................................................................................................... 151
    4.9.1 Synthetic strategies .............................................................................. 151
    4.9.2 Synthesis of N3-substituted nucleobases ................................................ 152
      4.9.2.1 Synthesis of N3-substituted thymine analogues using the deglycosylation strategy .................................................. 152
      4.9.2.2 Synthesis of N3-substituted uracil and thymine analogues using the N1-Boc protection strategy .................................................. 153
    4.9.3 Synthesis of (2R, 4S/R)-4-acetoxy-2-[(tert-butyldiphenylsilyl)oxy]methyl] dioxolane (“D-dioxolanyl acetate”) for the condensation with nucleobases .................................................. 155
    4.9.4 Glycosylation/condensation of the D-dioxolanyl acetate with nucleobases and the deprotection of the TBDPS .................................................. 160
    4.9.5 Synthesis of L-(2S, 4S/R)-4-acetoxy-2-[(benzoxyloxy)-methyl]dioxolane (“L-dioxolanyl acetate”) for the condensation with nucleobases .................................................. 161
A7  $^{13}$C DEPT135 NMR spectra of 3.26 .................................................................295
A8  HSQC NMR spectra of 3.26 ..............................................................................296
A9  HMBC spectra of N1-Boc-thymine .................................................................297
A10 The HMBC spectra of N3-butyl thymine .......................................................298
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Optimization characteristics and Autodock docking patterns of carboranyl antifolates 2.1 and 2.2</td>
</tr>
<tr>
<td>2.2</td>
<td>The effect of boron parameter implementation on Autodock docking patterns of carboranyl antifolates 2.1 and 2.2</td>
</tr>
<tr>
<td>2.3</td>
<td>Glide docking of closo- and nido-carboranyl antifolates listing the impact of partial charges on docking outcomes</td>
</tr>
<tr>
<td>2.4</td>
<td>The effects of docking mode base fragment for placement (BFP), and placement algorithm (PA) on FlexX docking of 1 C-crys and 2 N-crys</td>
</tr>
<tr>
<td>2.5</td>
<td>The effects of docking mode base fragment for placement (BFP) and placement algorithm (PA) on FlexX docking of 1 C-con and 2 N-con</td>
</tr>
<tr>
<td>2.6</td>
<td>Surflex docking of 1 C-crys and 2 N-crys</td>
</tr>
<tr>
<td>2.7</td>
<td>Surflex docking of 1C-con and 2 N-con</td>
</tr>
<tr>
<td>3.1</td>
<td>Radionuclides useful in diagnosis and imaging</td>
</tr>
<tr>
<td>3.2</td>
<td>Effect of various reaction conditions on the level of iodination of 3.1</td>
</tr>
<tr>
<td>3.3</td>
<td>Effect of various reaction conditions on the deglycosylation and anomerization of 3.24</td>
</tr>
<tr>
<td>3.4</td>
<td>Effect of various reaction conditions on the deglycosylation and anomerization of 3.27 and 3.28</td>
</tr>
<tr>
<td>3.5</td>
<td>The effects of various reaction conditions on compound 3.29 and 3.30</td>
</tr>
<tr>
<td>4.1</td>
<td>EBV latency program and latency associated genes and malignancies</td>
</tr>
<tr>
<td>4.2</td>
<td>Latency associated viral gene products and their proposed function/characteristics</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Structure of 2´-deoxynucleoside triphosphates (dNTPs)</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Biosynthetic pathways for nucleotides</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Nucleoside analogues for the anticancer and antiviral therapy</td>
<td>8-9</td>
</tr>
<tr>
<td>1.4</td>
<td>Unrooted phylogenetic relationship among dNKs</td>
<td>12</td>
</tr>
<tr>
<td>1.5</td>
<td>Representative structure of 3CTAs</td>
<td>14</td>
</tr>
<tr>
<td>1.6</td>
<td>Structure of N5-^{125}I, a carborane cage-radiohalogenated 3CTA.</td>
<td>17</td>
</tr>
<tr>
<td>1.7</td>
<td>Part of a single monomer of the X-ray crystal structure of hTK1 (PDB # 1W4R) showing primarily the “lasso domain” and dTTP in the substrate/ATP binding site</td>
<td>20</td>
</tr>
<tr>
<td>1.8</td>
<td>Part of a single monomer of the X-ray crystal structure of BcTK (PDB # 2JA1) showing primarily the “lasso domain” in the substrate/ATP binding site</td>
<td>21</td>
</tr>
<tr>
<td>1.9</td>
<td>Part of a single monomer of a homology model of hTK1 based on CaTK as a template primarily showing the “lasso domain” and no substrate in the substrate/ATP binding site</td>
<td>22</td>
</tr>
<tr>
<td>1.10</td>
<td>(A) Homology model of hTK1 (open form)</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>(B) Crystal structure of BcTK [PDB # 2JA1] (semi open form)</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>(C) Crystal structure of hTK1 [PDB # 1W4R] (closed form)</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>(D) Overlay of 1.6 A, 1.6 B, 1.6 C</td>
<td>24</td>
</tr>
<tr>
<td>1.11</td>
<td>Docked pose of N5-2OH (a 3CTA) in the crystal structure of “semi-open” BcTK</td>
<td>25</td>
</tr>
</tbody>
</table>
1.12 The metabolic roles of DHFR and TS in the pyrimidine nucleoside, purine nucleoside, and amino acid biosynthesis ...............................................................26

1.13 Traditional inhibitors of DHFR employed for the treatment of cancer and microbial infections ......................................................................................27

2.1 Structures and nomenclature of closo- and nido-carboranes .......................................................... 32

2.2 Dimensions and isosurfaces of adamantane (A) and o-carborane (B) .............................................33

2.3 Trimethoprim and the carboranyl antifolates 2.1 and 2.2 ................................................................. 35

2.4 Structure of 2 N-crys (A) and 2 N-con (B) ...................................................................................... 37

2.5 (A) Binding of 1C-prot with various active site amino acid residues of hDHFR ................................. 40
(B) Overlap of the crystal structures poses of 1C-prot (red) and 2 N-prot (green) ............................ 40
(C) Binding pattern of Trimethoprim in hDHFR ............................................................................... 40

2.6 (A) Overlap of the docked pose (Autodock) of 1 C-con with the corresponding pose of 1C-prot in the original crystal structure ........................................ 43
(B) Overlap of the docked pose (Autodock, major cluster with 68 solutions) of 2 N-con with the corresponding pose of 2 N-prot in the original crystal structure ................................................................................. 43

2.7 Overlap of the docked pose (Autodock) of 2 N-con with the corresponding pose of 2 N-prot in the original crystal structure ........................................................................................................ 44

2.8 (A) Pyrimidyl Portion (PP) ............................................................................................................. 49
(B) closo-o-Carboranyl Portion (CP) placed inside the active site of hDHFR using base fragment placement (BFP) option implemented in FlexX ............................................................................... 49

2.9 Average total binding energies of docked 1 C-crys and 2 N-crys with hDHFR .............................................. 56

2.10 Comparative docking evaluation by Autodock, Glide, FlexX and Surflex for
(A) 1 C-crys ..................................................................................................................................... 58
(B) 2 N-crys ..................................................................................................................................... 58

2.11 (A) 1 C-con ..................................................................................................................................... 59
(B) 2 N-con ..................................................................................................................................... 59
3.1 The 2D structures of 3.1 and 3.2 along with
(A) The molecular surface of N5 (3.1) ................................................................. 64
(B) The molecular surface of N5-I (3.2) ............................................................... 64
(C) The molecular surface of an unsubstituted o-carborane .................................. 64
(D) The molecular surface of monoiodinated o-carborane ................................... 64

3.2 Structure of o-carborane with atoms numbers ................................................... 66

3.3 The change in the $^1$H NMR chemical shifts of the carboranyl C-H of 3.1 as a result of various levels of iodination ...................... 87

3.4 Change in the chemical shifts in the proton coupled $^{11}$B NMR spectra of o-closo-carborane and N5 (3.1) as a result of iodination .............................................. 89

3.5 Change in the $^{11}$B NMR (decoupled) and $^1$H NMR chemical shifts of o-carborane as a result of substitutions with either iodine or bromine .......... 91

3.6 (A) Carboranyl C-H signals in the $^1$H NMR spectra of purified 3.7 a, purified 3.7 b, and the mixture 3.7 a and 3.7 b ..................................................... 92
(B) X-ray crystal structure of 3.7 a ........................................................................ 92

3.7 Change in the $^1$H NMR $^{11}$B NMR (decoupled) chemical shifts of 1-o-carboranyl-pentane as a result of substitutions with either iodine or bromine .............................................................. 96

3.8 $^1$H NMR spectroscopy as a tool to monitor the halogen exchange reactions .... 97

3.9 The $^{11}$B and $^1$H NMR (decoupled) spectra of 3.2 a / 3.2 b, and 3.22 a / 3.22 b as a result of substitutions with either iodine or bromine ..... 98

3.10 Radio-TLC analysis of the isotope exchange reaction of 3.2 a / 3.2 b with Na$^{125}$I in presence of Hermann’s catalyst ...................................................... 100

3.11 HPLC analysis for isotope exchange reaction of N5-$^{127}$I (3.2 a / 3.2 b) with Na$^{125}$I to yield N5-$^{125}$I (3.23 a / 3.23 b) ................................................................. 101

3.12 Strategies for the synthesis of various 3-substituted thymine/uracil based nucleosides ........................................................................................................ 105

3.13 Hypothetical mechanism, shown at the example of the initial attack of I$^+$ on C-2 carbonyl oxygen, for the deglycosylation and anomerization of pyrimidine nucleosides by ICl ................................................................. 111

xvii
3.14 Hypothetical mechanism for the resistance of the glycoside bond in ribonucleosides towards cleavage by ICl

3.15 Difference in chemical shifts and splitting pattern between 3.24 and 3.26

4.1 EBV life cycle

4.2 The sequence of events in the life cycle of EBV development inside B-cells that may be linked with the development of malignancies in B-cells

4.3 Important lytic cycle associated genes

4.4 The EBV early promoter Zp

4.5 Nucleosides and non-nucleosides agents used for studies with EBV associated malignancies

4.6 EHV4TK (EQ1, PDB # 1P672)

4.7 The amino acid binding interactions of dThd and ADP in EQ1

4.8 Amino acid sequence alignment of EBVTK on EHV4TK

4.9 Side view of the aligned “open lid” and the “closed lid” forms of the EBVTK homology models

4.10 Proposed novel and existing compounds to be evaluated in EBVTK enzyme assays

4.11 Comparison of the docking patterns of the dThd (A), and L-N3-methyl-dioxolanyl thymine (B, compound 4.72) in the active site of the “closed” homology model of EBVTK

4.12 Synthetic Strategy for the synthesis of D- and L-dioxolanyl nucleoside analogues

4.13 β-Radiogram for the phosphoryl transfer assays of 4.34 and L-FMAU with hTK1 and BaTK

4.14 Sequencing of the plasmid vector using Bgl II and BamH1 for the confirmation of hTK1 gene insertion

4.15 Western blot of hTK1 expression showing that the induction is at its peak at a concentration of 0.2 mM IPTG after 6 h of induction
Western blot of EBVTK expression showing that the induction is at its peak at 0.004 % of L-ARA after 8 h of induction .................................................................173

SDS-PAGE of EBVTK after FPLC purification showing the enrichment of EBVTK (69 KDa) in the bands T10 and T9 .................................................................174
# LIST OF SCHEMES

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 General reaction scheme for the synthesis of 3-iodo-(o)-carborane using (nido-o)-carborane (7,8-dicarba-(nido)-undecaborane)</td>
<td>67</td>
</tr>
<tr>
<td>3.2 Direct iodination of (o)-carborane using (I_2) and Friedel Craft’s catalyst (AlCl_3)</td>
<td>68</td>
</tr>
<tr>
<td>3.3 The halogen/isotope exchange reaction using “cold” B-halocarboranes</td>
<td>69</td>
</tr>
<tr>
<td>3.4 Direct iodination of 3.1 using (ICl) or (I_2) in presence or absence of 10 equivalents of (AlCl_3)</td>
<td>72</td>
</tr>
<tr>
<td>3.5 Synthetic strategy for 3.2 (a) / 3.2 (b) (N5-I)</td>
<td>78</td>
</tr>
<tr>
<td>3.6 Synthesis of 3.2 (a) and 3.2 (b) using 1, 5-diiodopentane</td>
<td>80</td>
</tr>
<tr>
<td>3.7 Synthesis and the deprotection of 3.10 (a) / 3.10 (b)</td>
<td>82</td>
</tr>
<tr>
<td>3.8 Synthesis and deprotection of 3.14 (a) / 3.14 (b)</td>
<td>83</td>
</tr>
<tr>
<td>3.9 Optimized synthesis of 3.2 (a) and 3.2 (b)</td>
<td>85</td>
</tr>
<tr>
<td>3.10 Synthesis of 3.20 (a) / 3.20 (b)</td>
<td>94</td>
</tr>
<tr>
<td>3.11 Bromine-iodine exchange reaction in presence of Hermann’s catalyst</td>
<td>95</td>
</tr>
<tr>
<td>3.12 Radiohalogen exchange reaction of 3.2 (a) / 3.2 (b) with (Na^{125})I to obtain 3.23 (a) / 3.23 (b)</td>
<td>99</td>
</tr>
<tr>
<td>3.13 The synthesis of substituted N3-butyl thymidine and reaction with (I_2)</td>
<td>106</td>
</tr>
<tr>
<td>3.14 The synthesis of substituted N3-butyl thymidine</td>
<td>113</td>
</tr>
<tr>
<td>3.15 Synthesis of N3-substituted uridines and their reaction with (ICl)</td>
<td>116</td>
</tr>
<tr>
<td>3.16 Application of (ICl) mediated deglycosylation strategy for the synthesis of N3-substituted thymine nucleobases</td>
<td>118</td>
</tr>
<tr>
<td>4.1 A) Synthesis of N3-methyl thymine using the deglycosylation strategy.</td>
<td>xx</td>
</tr>
</tbody>
</table>
B) Synthesis of nucleobases using the N1-Boc-protection strategy ..................154

4.2  Synthesis of the D-dioxolanyl acetate 4.23 for condensation with various nucleobases ..........................................................................................................................................................................................157

4.3  Condensation of silylated nucleobases with D-dioxolanyl acetate 4.23 ..........159

4.4  Removal of the TBDPS group from 4.24 A/B - 4.28 A/B and 4.30 A/B ..........161

4.5  Synthesis of the L-dioxolanyl acetate 4.54 for condensation with various nucleobases ..........................................................................................................................................................................................162

4.6  Step-by-step formation of 4.49 from 4.47 ........................................................163

4.7  Proposed alternative synthetic route for the L-dioxolanyl acetate 4.63 with a tert-butyldiphenylsilyl protective group ..........................................................................................................................................................................................165

4.8  Condensation of silylated nucleobases with the L-dioxolanyl acetate 4.54 .......166

4.9  Removal of the benzoyl group from 4.64 A / B - 4.68 A / B and 4.70 A / B ..................................................................................................................................................................................................................168

4.10 Proposed synthetic strategies for 4.41 ..................................................................175
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>3D</td>
<td>3-Dimensional</td>
</tr>
<tr>
<td>3CTAs</td>
<td>3-Carboranyl thymidine analogues</td>
</tr>
<tr>
<td>5-FU</td>
<td>5-Fluorouracil</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine diphosphate</td>
</tr>
<tr>
<td>ADT</td>
<td>Autodock tools</td>
</tr>
<tr>
<td>APT</td>
<td>Atomic Polar Tensors</td>
</tr>
<tr>
<td>Arg</td>
<td>Arginine</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>Avg</td>
<td>Average</td>
</tr>
<tr>
<td>AZT</td>
<td>3’-Azido-3’-deoxythymidine</td>
</tr>
<tr>
<td>^{10}B</td>
<td>Boron-10 isotope</td>
</tr>
<tr>
<td>^{11}B</td>
<td>Boron-11 isotope</td>
</tr>
<tr>
<td>B. anthracis</td>
<td><em>Bacillus anthracis</em></td>
</tr>
<tr>
<td>BaTK</td>
<td><em>Bacillus anthracis</em> thymidine kinase</td>
</tr>
<tr>
<td>BBB</td>
<td>Blood-brain barrier</td>
</tr>
<tr>
<td>BFP</td>
<td>Base fragment placement</td>
</tr>
<tr>
<td>BNCT</td>
<td>Boron neutron capture therapy</td>
</tr>
<tr>
<td>Boc</td>
<td>Butyloxycarbamate</td>
</tr>
<tr>
<td>Bu</td>
<td>Butyl</td>
</tr>
<tr>
<td>Cat</td>
<td>Catalytic amount</td>
</tr>
<tr>
<td>CaTK</td>
<td><em>Clostridium acetobutylicum</em> thymidine kinase</td>
</tr>
<tr>
<td>CED</td>
<td>Convection-enhanced delivery</td>
</tr>
<tr>
<td>CP</td>
<td><em>closo-o-carboranyl- or nido-o-carboranyl</em> portion</td>
</tr>
<tr>
<td>COSY</td>
<td>Correlation spectroscopy</td>
</tr>
<tr>
<td>DA</td>
<td>Dalton</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>DCP-AES</td>
<td>Direct current plasma-atomic emission spectroscopy</td>
</tr>
<tr>
<td>DEPT</td>
<td>Distortionless enhancement of polarization transfer</td>
</tr>
<tr>
<td>dCK</td>
<td>2’-Deoxycytidine kinase</td>
</tr>
<tr>
<td>dGK</td>
<td>2’-Deoxyguanosine kinase</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-(Dimethylamino) pyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>N, N-Dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>2’-Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNK</td>
<td>Deoxyribonucleoside kinase</td>
</tr>
<tr>
<td>dThd</td>
<td>Thymidine</td>
</tr>
<tr>
<td>dTMP</td>
<td>Thymidine monophosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>Thymidine triphosphate</td>
</tr>
<tr>
<td>dUrd</td>
<td>2’-Deoxyuridine</td>
</tr>
<tr>
<td>EBVTK</td>
<td>Epstein Barr virus thymidine kinase</td>
</tr>
<tr>
<td>EHV4TK</td>
<td>Equine herpes virus 4 thymidine kinase</td>
</tr>
<tr>
<td>eq</td>
<td>Equivalents</td>
</tr>
<tr>
<td>ESI</td>
<td>Electro spray ionization</td>
</tr>
<tr>
<td>ESPfit</td>
<td>Electrostatic Potential Fit method</td>
</tr>
<tr>
<td>Et</td>
<td>Ethyl</td>
</tr>
<tr>
<td>g</td>
<td>Gram</td>
</tr>
<tr>
<td>GBM</td>
<td>Glioblastoma multiforme</td>
</tr>
<tr>
<td>Glu</td>
<td>Glutamic acid</td>
</tr>
<tr>
<td>Gln</td>
<td>Glutamine</td>
</tr>
<tr>
<td>h</td>
<td>Hour</td>
</tr>
<tr>
<td>hDHFR</td>
<td>Human Dihydrofolate Reductase</td>
</tr>
<tr>
<td>HHV8TK</td>
<td>Human herpes virus 8 thymidine kinase</td>
</tr>
<tr>
<td>HMBC</td>
<td>Heteronuclear multiple-bond correlation</td>
</tr>
<tr>
<td>HMQC</td>
<td>Heteronuclear multiple-quantum coherence</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>HRMS</td>
<td>High resolution mass spectroscopy</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HSV1TK</td>
<td>Herpes simplex virus 1 thymidine kinase</td>
</tr>
<tr>
<td>HSV2TK</td>
<td>Herpes simplex virus 2 thymidine kinase</td>
</tr>
<tr>
<td>hTK1</td>
<td>Human thymidine kinase 1</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>IC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>Inhibitory concentration 50%</td>
</tr>
<tr>
<td>IR</td>
<td>Infra-red</td>
</tr>
<tr>
<td>LDL</td>
<td>Low-density lipoprotein</td>
</tr>
<tr>
<td>LGA</td>
<td>Lamarckian Genetic Algorithm</td>
</tr>
<tr>
<td>LUMO</td>
<td>Lowest occupied molecular orbital</td>
</tr>
<tr>
<td>m</td>
<td>Meta</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>Me</td>
<td>Methyl</td>
</tr>
<tr>
<td>MHz</td>
<td>Megahertz</td>
</tr>
<tr>
<td>min</td>
<td>Minute</td>
</tr>
<tr>
<td>mol</td>
<td>Mole</td>
</tr>
<tr>
<td>MP</td>
<td>Monophosphate</td>
</tr>
<tr>
<td>N</td>
<td>Normality</td>
</tr>
<tr>
<td>NADP</td>
<td>Dihydronicotinamide Adenine Dinucleotide</td>
</tr>
<tr>
<td>NADPH</td>
<td>Dihydronicotinamide Adenine Dinucleotide Phosphate</td>
</tr>
<tr>
<td>n-BuLi</td>
<td>n-Butyllithium</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>o</td>
<td>Ortho</td>
</tr>
<tr>
<td>p</td>
<td>Para</td>
</tr>
<tr>
<td>PA 1</td>
<td>Placement algorithm 1</td>
</tr>
<tr>
<td>PA 2</td>
<td>Placement algorithm 2</td>
</tr>
<tr>
<td>PA 3</td>
<td>Placement algorithm 3</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>PP</td>
<td>Pyrimidyl portion</td>
</tr>
<tr>
<td>ppm</td>
<td>Parts per million</td>
</tr>
<tr>
<td>PR</td>
<td>Phosphorylation rate</td>
</tr>
</tbody>
</table>

xxiv
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTA</td>
<td>Phosphoryl transfer assay</td>
</tr>
<tr>
<td>rdf</td>
<td>receptor description file</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root mean square deviation</td>
</tr>
<tr>
<td>RP-18</td>
<td>C18 reversed phase</td>
</tr>
<tr>
<td>rt</td>
<td>Room temperature</td>
</tr>
<tr>
<td>SAR</td>
<td>Structure activity relationship</td>
</tr>
<tr>
<td>SP</td>
<td>Standard precision</td>
</tr>
<tr>
<td>TBAF</td>
<td>Tetrabutylammonium fluoride</td>
</tr>
<tr>
<td>TBDMS</td>
<td>tert-Butyl dimethyl silyl</td>
</tr>
<tr>
<td>TBDPS</td>
<td>tert-Butyl diphenyl silyl</td>
</tr>
<tr>
<td>THF</td>
<td>Tetrahydrofuran</td>
</tr>
<tr>
<td>TK2</td>
<td>Thymidine kinase 2</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin layer chromatography</td>
</tr>
<tr>
<td>TS</td>
<td>Thymidylate Synthetase</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
<tr>
<td>UuTK</td>
<td><em>Ureaplasma urealyticum</em> thymidine kinase</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>VZVTK</td>
<td>Varicella zoster virus thymidine kinase</td>
</tr>
</tbody>
</table>
CHAPTER 1

Nucleoside kinases and dihydrofolate reductase in anticancer- and antiviral therapy and imaging
1.1. Introduction

The hereditary molecule in all cellular life forms is DNA (deoxyribonucleic acid). The basic structural components of DNA are nucleosides, which are linked through phosphate bridges. DNA is synthesized from the 5′-triphosphates of four different deoxynucleosides (deoxynucleoside triphosphates or dNTPs), which are 2′-deoxyadenosine triphosphate (dATP), 2′-deoxyguanosine triphosphate (dGTP), 2′-deoxycytidine triphosphate (dCTP), and thymidine triphosphate (TTP). All dNTPs are composed of a deoxyribose sugar, a triphosphate, and either a purine- or pyrimidine nucleobase (Figure 1.1). Purine nucleobases are adenine (A) and guanine (G) whereas pyrimidine nucleobases are cytosine (C) and thymine (T).

![Figure 1.1: Structure of 2′-deoxynucleoside triphosphates (dNTPs).](image_url)

dATP: 2′-deoxyadenosine triphosphate, dGTP: 2′-deoxyguanosine triphosphate, dCTP: 2′-deoxycytidine triphosphate, dTTP: 2′-deoxythymidine triphosphate.
1.2. Nucleotide biosynthetic pathways

dNTPs are biosynthesized either \textit{de novo} or via salvage pathway, as depicted in Figure 1.2. The \textit{de novo} synthesis of dNTPs utilizes basic metabolic precursors such as glutamine, glycine, aspartate, 5-phosphoribosyl-1-pyrophosphate, CO$_2$, and folic acid. Inosine monophosphate (IMP) and uridine monophosphate (UMP) are key intermediates in the purine- and the pyrimidine-based dNTP \textit{de novo} synthesis, respectively.

![Figure 1.2: Biosynthetic pathways for nucleotides.](image)

dN: 2′-deoxynucleosides, dNMP: 2′-deoxynucleoside monophosphates, dNDP: 2′-deoxynucleoside diphosphates, dNTP: deoxynucleoside triphosphates, NMP: nucleoside monophosphates, NDP: nucleoside diphosphates, dNKs: 2′-deoxynucleoside kinase, NMPK: nucleoside monophosphate kinase, NDPK: nucleoside diphosphate kinase, RR: ribonucleotide reductase, 5′-NT: 5′-nucleotidase.
Three key enzymes in the *de novo* biosynthetic pathway of pyrimidine nucleotides are ribonucleotide reductase (RR), dihydrofolate reductase (DHFR), and thymidylate synthetase (TS). RR catalyzes the conversion of NDPs into dNDPs via reduction of the 2′-hydroxyl group of ribose. Thus, it plays an important metabolic role in the *de novo* pathway. DHFR and TS are essential for the synthesis of dTMP. The individual roles of DHFR and TS in catalyzing the biosynthesis of dTMP are discussed in more detail in Section 1.9.

2′-Deoxynucleosides (dNs) generated from DNA catabolism are utilized by the salvage pathway. The dNs are converted to the corresponding 2′-deoxynucleoside-5′-monophosphates (dNMPs) by deoxyribonucleoside kinases (dNKs). These are enzymes that catalyze the transfer of the γ-phosphate mostly of ATP to the 5′-hydroxyl groups of dNs. The dNMPs are subsequently phosphorylated to dNDPs (2′-deoxynucleoside-5′-diphosphates) by nucleoside monophosphate kinases (NMPKs) and finally to dNTPs by nucleoside diphosphate kinases (NDPKs). The initial phosphorylation of dNs by dNKs is irreversible whereas the subsequent phosphorylation steps of are reversible.1-3 5′-Nucleotidases (5′-NTs) are intracellular nucleoside catabolizing enzymes responsible for the dephosphorylation of dNMPs.1,3

1.3. dNKs (2′-deoxyribonucleoside kinases)

2′-Deoxynucleoside kinases are key enzymes of the salvage pathway because they carry out the rate limiting 5′-phosphorylation of endogenous nucleosides as well as that of many anticancer and antiviral nucleoside analogues.4 Mammalian cells possess four distinct dNKs, i.e. thymidine kinase 1 (TK1), deoxycytidine kinase (dCK), thymidine
kinase 2 (TK2), and deoxyguanosine kinase (dGK). TK1 and dCK are cytosolic enzymes whereas TK2 and dGK are located in mitochondria. TK1 carries out the initial 5′-phosphorylation of thymidine (dTd) and 2′-deoxyuridine (dUrd) to the corresponding monophosphates (dTMP and dUMP). Deoxycytidine kinase (dCK) on the other hand has broad substrate specificity. It catalyzes the phosphorylation of endogenous 2′-deoxycytidine (dCyd), 2′-deoxyguanosine (dGuo), and 2′-deoxyadenosine (dAdo) to the corresponding monophosphates (dCMP, dGMP, and dAMP) using either ATP or UTP as phosphate donors. Mitochondrial TK2 utilizes dThd, dCyd, and dUrd as substrates whereas dGuo, dAdo, and deoxyinosine (dIno) are the endogenous substrates of dGK.

The activities of the cytosolic kinases, in particular that of hTK1, is higher in tissues with proliferating cells whereas the mitochondrial kinases are equally active in all tissues. Based on the amino acid sequences and substrate specificities, known dNKs are classified into three subgroups:

1) TK1-like kinases, which include human TK1 (hTK1), poxviruses TK, and TKs from various bacteria, e.g. Bacillus anthracis TK (BaTK), Bacillus cereus TK (BcTK), Ureaplasma urealyticum TK (UuTK), Clostridium acetobutylicum TK (CaTK), and Thermotoga maritima TK (TmTK) [Section 1.4]

2) Non-TK1-like enzymes, such as TK2, dCK, dGK, and dNK from fruit fly (Drosophila melanogaster) (Section 1.4)

3) Herpes virus TKs. (Section 1.5)
1.4. Substrate specificity of mammalian dNKs.

TK1-like kinases have the most restricted substrate specificity of all salvage pathway enzymes. They only accommodate minor modifications at the C-5 position of dThd, such as 5-halo or 5-ethyl substituents and at the 3′-position, such as hydrogen, fluorine, or azido groups. In 1996, Lunato et al. demonstrated that hTK1 can also tolerate significantly larger substituents at the N3-position of the thymidine analogues. The phosphorylation of nucleoside analogues by hTK1 is stereospecific in as far as almost exclusively natural and unnatural D-nucleosides are accepted as substrates. The only known exception is L-FMAU [1-(2-fluoro-5-methyl-β-L-arabinofuranosyl)uracil] (Figure 1.3).

As already mentioned above, dCK has a fairly broad substrate specificity and this may be the reason why it plays a pivotal role in the activation of several anticancer and antiviral nucleoside analogues such as araA, Emtricitabine (FTC), Lamivudine (3TC), Cytarabine (araC), Cladribine (CdA), Fludarabine, Gemcitabine, and Zalcitabine (Figure 1.3). dCK phosphorylates nucleoside analogues with both 2′- and 3′-modifications (araA, araC, Zalcitabine, 3TC, Gemcitabine) as well as acyclic sugars (e.g. Cytallene). Also, dCK does not distinguish between D- and L-nucleoside analogues. The ability of dCK to phosphorylate L-nucleosides led to the development of several L-nucleoside analogues, primarily for antiviral therapy (e.g. Emtricitabine, Lamivudine, and L-FMAU) but also for the experimental treatment of cancers (L-dioxolanyl cytidine [L-OddC]) (Figure 1.3).

dGK tolerates a fairly wide range of modifications to the purine nucleoside scaffold except for 3′-modifications to the deoxyribose portion and acyclic sugars. Similar to
dCK, dGK accepts both D- and L-nucleosides.\textsuperscript{4,13,15} dGK has been shown to monophosphorylate several anticancer and antiviral nucleoside analogues, such as araG, araA, CdA, and Clofarabine.\textsuperscript{4,13,15}

In contrast to hTK1, TK2 has a relaxed enantioselectivity in that it accepts both D- and L-nucleosides. It also tolerates a fairly wide range of modification in particular at the carbohydrate portion of the endogenous substrates.\textsuperscript{4,5,10,12} Also, TK2 is evolutionarily closer to dCK and dGK than to TK1 (See Figure 1.4). Its amino acid sequence shows approximately 40\% identity to that of dCK and dGK.\textsuperscript{4,5,10,12} In addition, the amino acid residues responsible for substrate recognition in all three enzymes are conserved, which explains the nearly identical substrate specificities.\textsuperscript{4,5,10,12}
Figure 1.3: Nucleoside analogues for anticancer and antiviral therapy. The numbering of the ring system of purine- and pyrimidine nucleosides is shown at the examples of araA and AZT, respectively (continued on the next page).
Figure 1.3: (contd.): Nucleoside analogues for the anticancer and antiviral therapy.
1.5. Substrate specificity of viral dNKs

The human herpes viruses are a group of double stranded DNA viruses that are divided into three subfamilies, i.e. α-, β-, and γ-herpes viruses (see Section 1.3.1 for details). Several herpes viruses encode thymidine kinases, including Herpes Simplex Viruses 1 and 2 (HSV1TK and HSV2TK), Varicella Zoster Virus (VZVTK), Equine Herpes Virus 4 (EHV4TK), Epstein-Barr Virus (EBVTK), and Kaposi’s Sarcoma-Associated Herpes Virus (KSHV)/Human Herpes Virus 8 (HHV8TK). Four major differences exist between TK1-like kinases and viral TKs. 1) Except for HSV2TK, all viral TKs have an additional thymidylate kinase activity that converts dTMP into dTDP, 2) the specificity requirements of the viral TKs are much less stringent than those of cellular dNKs, both with respect to the phosphate donor (NTP) and the substrates (both purines and pyrimidine nucleosides are phosphorylated), 3) viral TKs cannot distinguish between D- and L-nucleosides, and 4) several viral TKs, such as HSV1TK, HSV2TK, and VZVTK, accept nucleoside derivatives with acyclic sugar moieties as substrates. Since viral TKs are evolutionary closer to dCK, TK2, and dGK (See Figure 1.4), it was hypothesized that the viral enzymes might have evolved from a captured human deoxycytidine kinase gene. The broad substrate specificity of viral TKs has been the basis for several antiviral chemotherapeutic approaches.

There are distinct differences among the herpes virus TKs. Gustafson et al. demonstrated that EBVTK is able to phosphorylate dThd but not dCyd, which is a substrate of HSV1TK and of HSV2TK. EBVTK was also found to have weaker thymidylate kinase activity than HSV1TK. In addition, it was found that Ganciclovir and Acyclovir (Figure 1.3) are not substrates of EBVTK whereas they are excellent
substrates of HSV1TK, HSV2TK, and VZVTK. In contrast to HSV1TK and HSV2TK, dThd analogues such as D4T and AZT (Figure 1.3) were found to be effective in inhibiting dThd phosphorylation by the EBVTK whereas as several cytidine and guanosine analogues such as Gemcitabine and Cidofovir (Figure 1.3) were ineffective. The phylogenetic tree displayed in Figure 1.4 indicates that EBVTK and HHV8TK evolved distinctly from the other viral TKs. This may explain their specific substrate characteristics, which almost seem to be closer to those of the TK1-like kinases.
hTK1-like dNKs are in red, non-TK1-like dNKs are in black, and viral TKs are in blue and purple. BaTK, *Bacillus Anthracis* thymidine kinase; BcTK, *Bacillus cereus* thymidine kinase; CaTK, *Clostridium Acetobutylicum* thymidine kinase; dCK, deoxycytidine kinase; dGK, deoxyguanosine kinase; EBVTk, Epstein Barr Virus thymidine kinase; EHV4-TK, Equine Herpes Virus-4 thymidine kinase; HHV8-TK, Human Herpes Virus 8 thymidine kinase; HSV1TK, Herpes Simplex Virus 1 thymidine kinase; HSV2TK, Herpes Simplex Virus 2 thymidine kinase; hTK1, human thymidine kinase 1; TK2, thymidine kinase 2; TmTK, *Thermatoga maritime* thymidine kinase; VZVTK, Varicella Zoster Virus thymidine kinase.

1.6. The importance of dNKs in the therapy of cancer and infectious diseases

1.6.1. dNKs in Boron Neutron Capture Therapy (BNCT)

Human TK1 has also been used as a molecular target in BNCT, a binary cancer treatment modality that is based on both a chemical and a radiation component. It relies
on a nuclear reaction between a stable $^{10}$B isotope and thermal neutrons. When a $^{10}$B atom is irradiated by thermal neutrons, an unstable $^{11}$B isotope is produced, which undergoes fission to yield a $^4$He- ($\alpha$-particle) and a $^7$Li nucleus. These high linear energy transfer (LET) particles can cause DNA double strand breaks. Because of their short range (< 10 µm) in biological tissue, however, the lethal damage is largely restricted to the targeted (tumor) cells.$^{16,36,37}$ A minimum of 20 µg of $^{10}$B per gram of tumor tissue is needed for successful BNCT. The boron should be selectively accumulated/delivered inside the tumor cells in order to avoid toxicity to the surrounding cells.$^{16,36,37}$

Boron-containing dThd analogues may be highly promising BNCT agents because they may accumulate selectively inside cancer cells following monophosphorylation by hTK1, which is overexpressed in cancer cells (see Section 1.7 for further discussion). Such selective entrapment of nucleoside analogues in tumor cells has been referred as “kinase mediated trapping (KMT).”$^{16,38}$

In 1996, Luanto et al. reported the synthesis of N3 substituted carboranyl thymidine analogues, which were later designated as 3CTAs (3-carboranyl thymidine analogues, see Figure 1.5 for representative 3CTAs).$^{19,20}$ These compounds are substituted with carborane clusters at the N3 position, which are prototype boron moieties for BNCT because of their high boron content and stability coupled with synthetic versatility (see Chapter 2, Figure 2.1 for the representative structures of carboranes).
Figure 1.5: Representative structure of 3CTAs

The biological evaluations of 3CTAs using phosphoryl transfer assays and, subsequently, more detailed kinetic studies with hTK1 identified N5 and N5-2OH as two lead 3CTAs, with N5-2OH being slightly superior to N5.20,36,39,40 The \( \frac{r_{\text{cat}}}{K_M} \) values for both N5 and N5-2OH were found to be 26.7% and 35.8%, respectively, which are higher than that of the established TK1 substrate FLT (7.6%) and, by a small margin, lower than that of AZT (43.7%) (see Figure 1.3 for structures). Additional enzymatic studies with N5 and N5-2OH showed that these compounds were not the substrates of TK2 and dCK as well as nucleoside catabolizing 5’-nucleotidases (5’-NTs) and nucleoside phosphorylases.16,36,39,40

In vivo BNCT studies with mice bearing TK1 (+) L929 tumors or TK1 (-) L929 tumors that were treated with N5-2OH indicated that the former tumor was far more
responsive to neutron irradiation. This was presumably due to the fact that the TK1 (+) tumor accumulated approximately three times more boron than the TK1 (-) tumor.\textsuperscript{41} In comparative BNCT studies with \textbf{N5-2OH} and $p$-dihydroxyboryl phenyl alanine (BPA), a clinically used BNCT agent, in rats bearing intracerebral RG2 gliomas, the former agent produced a 2.4 times longer survival than the latter.\textsuperscript{41} It was concluded that KMT might play a major role in the selective entrapment of \textbf{N5-2OH} in the TK1 (+) L929 tumor and the RG2 glioma.\textsuperscript{41}

1.6.2. Other therapeutic anticancer applications of dNKs

The roles of various dCK and dGK targeting nucleoside analogues in anticancer therapy were already briefly discussed in \textbf{Section 1.4}. dNKs have also been explored extensively in suicide gene therapy. In this treatment modality for cancer, a viral dNK-encoded gene is inserted into the genome of targeted cancer cells. The encoded viral dNK must be able to selectively activate (monophosphorylate) a nucleoside prodrug, which is not activated by host cell dNKs.\textsuperscript{42-44} The most commonly used dNK/drug combination in suicide gene therapy is HSV1TK/ganciclovir. Ganciclovir monophosphate, selectively produced in transfected cancer cells by HSV1TK activity, undergoes sequential di- and tri-phosphorylation by host cell NMPK and NDPK, respectively. The triphosphate form of ganciclovir is an effective inhibitor of human DNA polymerase. In addition, its incorporation into DNA chains causes termination of DNA replication, and consequently, tumor cell death.
1.6.3. The role of dNKs in the development of antiviral therapeutics

The roles of various dCK substrates in antiviral therapy were already briefly discussed in Section 1.4. Apart from the natural substrate dThd and dUrd, hTK1 also catalyzes the phosphorylation of AZT and D4T (Figure 1.3), two important clinically used anti-HIV agents. The subsequent di- and triphosphorylation of these dThd analogues is carried out by host cell kinases. The specific toxic effects towards HIV are predominantly exerted by the active triphosphate forms of these prodrugs, which block DNA synthesis by inhibition HIV reverse transcriptase and/or incorporation at 3´-terminal ends of viral DNA chains.4,5,8,9,11,14,16,25,45

Nucleoside analogues approved for the treatment of α- and β- herpes virus infections comprise Acyclovir, Valacyclovir, Penciclovir, Ganciclovir, Valganciclovir, Famciclovir, Brivudine, Idoxuridine, Trifluridine, and Cidofovir (Figure 1.3).46-48 Most of the these nucleoside prodrugs are selectively monophosphorylated by α- or β-herpes virus TKs or protein kinases followed by conversion to the active triphosphate forms by host cell kinases. As will be discussed in detail in Chapter 4, some of these prodrugs have been explored for the treatment of EBV infections and also EBV-associated malignancies. However, none of them have been approved for these purposes. One reason for the lack of efficacy of most of these nucleoside prodrugs against EBV-associated diseases may be that they are not effectively monophosphorylated by EBV kinases. As discussed in Section 1.5, EBVTK has evolved separately from α- and β-herpes virus TKs, and thus it has dissimilar substrate characteristics. The development of novel nucleosides prodrugs that are specific for EBVTK is therefore part of this thesis and will be presented in Chapter 4.
1.7. The importance of dNKs in the prognosis and diagnosis of cancer and infectious diseases

During the G1/S transition of normal cells, hTK1 activity increases 10- to 20-fold and remains constant in the cell until M phase when it is rapidly degraded. Cancer cells lose cell-cycle control of hTK1, which leads to even further increased levels of hTK1. Serum hTK1 levels correlate strongly with the proliferative stage and aggressiveness of cancers, although the exact mechanism for the transport from the inside of cancer cells into the serum is not known. This underscores the basis for using hTK1 as a prognostic indicator for many solid tumors, including breast cancer, prostate cancer, bladder carcinoma, and small-cell carcinoma of the lung.

![Figure 1.6: Structure of N5-125I, a carborane cage-radiohalogenated 3CTA.](image)

Deoxynucleoside kinases have also been explored as molecular targets for the imaging of cancer as well as viral and bacterial infections. In particular, nucleoside prodrugs such as FLT, L-FMAU, and D-FIAU (Figure 1.3) have been explored in this context. The phosphorylation of these agents by human, viral or bacterial thymidine kinases leads to their monophosphorylation, and their subsequent KMT in tumor- or
infected cells. A novel type of hTK1-targeting cancer diagnostics are carborane cage radiohalogenated 3-CTAs, such as N5-125I (Figure 1.6). The development of these types of cancer diagnostics is part of this thesis and will be presented in Chapter 3. They constitute a novel type of radiopharmaceuticals, which should be excellent tools to study the general suitability of low molecular weight (MW) compounds as carriers for radio-iodinated boron clusters for the diagnosis and therapy of cancer.

1.8. The “lasso loop” in TK1-like kinases: Impact of protein flexibility on drug design.

X-ray crystallographic analysis revealed that hTK1 is a tetramer, and that each subunit of hTK1 consists of one large α/β domain and another smaller zinc binding domain, which also has been described as the “lasso domain”.\textsuperscript{17,27,28,54} The larger α/β domain contains the substrate (dThd) and an ATP binding site along with a Mg\textsuperscript{2+} ion (Figure 1.7).\textsuperscript{17,27,28,54} The smaller lasso domain covers the dThd binding site and is characterized by the presence of a highly conserved Arg-Tyr couple. The Arg-Tyr couple along with dThd, which in the case of the crystal structure shown in Figure 1.7 is part of dTTP, holds the lasso loop in place by hydrogen bonding with the main chain amino acids valine (2 x) and glutamic acid. The orientation of the lasso loop shown in Figure 1.7 gives the impression of a “closed lid” and indicates that the substrate binding site is full and no longer available. This form of hTK1 could be intuitively referred to as a “closed” (holo) form.\textsuperscript{16,36}

On the other hand, in the absence of substrate, dThd, the lasso domain should have the form of an “open lid”, and this hTK1 conformation could be considered an “open”
(apo) form. This hypothesis has found support in the recent crystallographic investigation of BcTK. In the crystal structure of BcTK, an exogenous substrate, 2-methyl-2, 4-pentanediol (MPD), has displaced dThd (See Figure 1.8). This led to a drastic change in the conformation of the lasso loop because of the breakdown of the hydrogen bonding network of the Arg-Tyr couple and dThd with the lasso loop, resulting in an apparently partial opening of the dThd binding site. While exogenous MPD still interacts with some of the amino acids of the lasso domain, such as Tyr from Arg-Tyr couple and isoleucine, it certainly forms less hydrogen bonds than dThd. Thus, this hTK1 conformation could be considered as “semi-open”.

It is conceivable that in TK1-like enzymes without any substrate, the highly flexible lasso loop may even take a different conformation than in BcTK. This hypothesis can be rationalized by inspecting a homology model of hTK1 that is based on the crystal structure of CaTK (Figure 1.9)
Figure 1.7: Part of a single monomer of the X-ray crystal structure of hTK1 (PDB # 1W4R)\textsuperscript{54} showing primarily the “lasso domain” (green) and dTTP in the substrate/ATP binding site.
Figure 1.8: Part of a single monomer of the X-ray crystal structure of BcTK (PDB # 2JA1) showing primarily the “lasso domain” (brown) and MPD near the substrate/ATP binding site. MPD shows weak hydrogen bonding interactions with Tyr of Arg-Tyr couple and the backbone carbonyl oxygen of isoleucine.
Figure 1.9: Part of a single monomer of a homology model of hTK1 based on CaTK as the template (PDB # 1XX6, amino acid homology = 39 %) primarily showing the “lasso domain” (blue) and no substrate in the substrate/ATP binding site.
The choice of this template was based on the fact it does not contain any substrate in substrate/ATP binding site and is characterized by the presence of only a partial lasso loop with a apparently different conformation than that of hTK1. This homology model of hTK1 shows that the lasso loop is “wide open” making the substrate binding site accessible for the entry and the binding of dThd (See Figure 1.9). The conformational change of the lasso loop in this case is such that the main chain amino acids, valine, valine, and glutamic acid move upwards, which results in the disruption of all hydrogen bond interaction with Arg-Tyr couple. Only a weaker hydrogen bond interaction between Arg from the Arg-Tyr couple with the backbone carbonyl oxygen of glycine of the lasso domain remains (see Figure 1.9).

Based of the analysis of protein structures shown in Figures 1.7 – 1.9, we hypothesize that the lasso loop undergoes conformational transformation upon dThd binding from an “open form” to a “semi-open/partially closed form” to a completely “closed form”. This sequence of events is clearly depicted in Figure 1.10 A-D. This conformational flexibility was also discussed in detail by Lavie et al. at the example of the apo and holo forms of TmTK.56,57
Figure 1.10: (A) Homology model of hTK1 (open form)

(B) Crystal structure of BcTK [PDB # 2JA1] (semi open form)

(C) Crystal structure of hTK1 [PDB # 1W4R] (closed form)

(D) Overlay of (A), (B) and (C)
Figure 1.11: Docked pose of N5-2OH in the crystal structure of “semi-open” BcTK.\(^\text{16}\)

The ChemDraw (2D) structure of N5-2OH is shown next to the enzyme.

This flexibility of the lasso domain may explain specific hTK1 substrate/inhibitor characteristics of 3CTAs and other N3-substituted dThd analogues that have bulky groups bound to the N3-position.\(^\text{16}\) As discussed in Section 1.6.1, 3CTAs such as N5-2OH are excellent substrates of hTK1. However, they can only be docked into
“open/semi-open” forms of TK1-like enzymes, such as BcTK.\textsuperscript{16} As shown in Figure 1.11, the bulky carboranyl substituent (see Chapter 2, Figure 2.2 B for the dimension of larger substituents such as carboranes and adamantanes) at the N3-position of N5-2OH is oriented towards the gap provided by the “semi-open” lasso domain, which enables the binding of this agent to the substrate binding site of BcTK. Docking of N5-2OH to “closed forms” of TK1-like enzymes is not possible, because of the absence of such an opening that could accommodate the bulky carborane cluster.

1.9. Dihydrofolate reductase (DHFR)

![Dihydrofolate reductase (DHFR) diagram]

**Figure 1.12:** The metabolic roles of DHFR and TS in the pyrimidine nucleoside-, purine nucleoside-, and amino acid biosynthesis.
In cooperation with thymidylate synthetase (TS), DHFR plays an integral role in maintaining the pool of reduced folates, such as tertahydrofolate (THF), dihydrofolate (DHF), and N5, N10-mehtylenetetrahydrofalte, which are important for the biosynthesis of purine nucleoside-, pyrimidine nucleoside-, and several amino acids (Figure 1.12).^{58-60}

![Methotrexate](image1.png)

![Trimetrexate](image2.png)

![Aminopterin](image3.png)

![Pemetrexed](image4.png)

![Trimethoprim](image5.png)

![Piritrexim](image6.png)

![Pyrimethamine](image7.png)

**Figure 1.13:** Traditional inhibitors of DHFR used for the treatment of cancer and microbial infections.
Prokaryotes synthesize their own folates and lack the transporters for their uptake from the host system, whereas mammals do not synthesize folates and depend on their diet for supply. In either case, DHFR plays a key metabolic role, and thus, has been an attractive molecular target for the development of chemotherapeutics for microbial infections and cancer.

Selectivity of agents such as trimethoprim (Figure 1.13) for microbial DHFR was the basis for the treatment of parasitic infection. The 4-amino folic acid analogues, aminopterin and methotrexate (MTX) (Figure. 1.13), were the first antifolates used for the chemotherapy of childhood acute lymphoblastic leukemia (ALL). Rationally designed novel antifolates were developed, which also have TS inhibitory activity. These include raltitrexed and pemetrexed (Figure 1.13), which have been used successfully for the treatment of advanced colorectal and non-small cell lung cancer. Furthermore, many antifolates have become important components of different chemotherapeutic regimens currently used for the treatment of other human malignancies including osteosarcoma, breast cancer, and primary central nervous system lymphoma.

The x-ray crystal and NMR-based 3D-structures of various DHFRs along with various inhibitors have been studied extensively. The binding of DHFR inhibitors mainly involves two primary types of interactions, H-bonding and lipophilic/hydrophobic. Crystal structures of closo- and nido-carboranyl antifolates with human dihydrofolate reductase (hDHFR) also have been reported recently, which show the same type of interactions (see Chapter 2). This was the first report on carborane-containing compounds (see Chapter 2, Figure 2.3) that were crystallized with a protein. These crystal structures provided for the first time the opportunity to verify docking
strategies that have been developed previously for carboranyl compounds. In Chapter 2, we will evaluate the docking of these carboranyl antifolates into the active site of the hDHFR crystal structures using the docking programs Autodock, Glide, FlexX, and Surflex. The described computational studies should be of great interest for scientists involved in the design and development of boron-containing therapeutics and diagnostics who had very limited access to computational tools in the past.
CHAPTER 2

A comparative docking evaluation of antifolates that contain carborane clusters using Autodock, Flexx, Glide, and Surflex

Contributors other than myself:

1) Kiran Mahasenan (Graduate Student, Division of Medicinal Chemistry and Pharmacognosy, The Ohio State University)

2) Ryan Pavlovicz (Graduate Student, Biophysics Program, The Ohio State University)

Part of this chapter was published in the Journal of Chemical Information and Modeling. Permission was obtained from the American Chemical Society for the reproduction of text, tables and figures.
2.1. Introduction

Chemical, physicochemical, and structural versatility combined with high stability under physiological conditions are distinctive features of carboranes and other boron clusters. Carboranes are three dimensional σ-aromatic structures that exist in three isomeric forms. The isomers are named according to the position of the two CH vertices with respect to each other. In ortho-carborane, the two carbon atoms are directly connected to each other in 1,2-geometric relationship, whereas in the meta- and in the para-carborane, the CH vertices have 1,7- and 1,12- geometric relationships, respectively (See Figure 2.1).

ortho-Carborane and meta-carborane undergo deboronation, and hence degradation, to the corresponding “nido” forms by treatment with the fluoride anion, as in tetrabutylammonium fluoride (TBAF), or strong Brønsted bases, such as the methoxide ion, triethylamine, hydrazine, piperidine, ammonia (Figure 2.1). para-Carborane, on the other hand, demonstrated an extraordinary stability in presence of bases or the fluoride ion and deboronation was only possible by treatment with potassium hydroxide under extreme reaction conditions in presence of nucleophilicity enhancing agents, such as crown ether. The treatment of ortho-carborane and meta-carborane with strong base removes the most electropositive boron vertex from the carborane cage leaving behind an open-cage or “nido” cluster with a bridging hydrogen occupying the empty vertex. This acidic bridging hydrogen is also known as the “extra” hydrogen.
Figure 2.1: Structures and nomenclature of closo- and nido-carboranes

Carboranes have been used for decades in the design and synthesis of therapeutics for Boron Neutron Capture Therapy (BNCT) of cancer (see Chapter 1, Section 1.6.1) and, more recently, also in other areas of drug design. Hydrophobic closo-carboranes, comparable in dimensions to adamantane, (see Figure 2.2) were used as bioisosteric replacements for (hetero)aromatic and (hetero)aliphatic ring systems and other bulky entities in the design and synthesis of carboranyl derivatives of various amino acids and peptides, estrogen receptor modulators, androgen receptor antagonists, retinoids, benzolactamic protein kinase C inhibitors, thalidomide, flufenamic acid, diflunisal, thrombin inhibitors, and trimethoprim. Many of these boronated derivatives displayed biological activities comparable or even superior to those of their non-boronated counterparts. In addition, metallocarboranes were found to be effective inhibitors of HIV-1 protease and therapeutics containing single boron atoms, such as the proteasome inhibitor bortezomib, have attracted considerable attention in recent years.
Drug design involving boron-containing agents has two major disadvantages compared with conventional drug design: (1) There is a lack of compound libraries containing boron agents for virtual screening, and (2) many software packages available for structure/ligand-based drug design do not have inbuilt parameters for boron. Several strategies to circumvent the latter problem have been reported in recent years. These include the substitution of boron with carbon and the calculation of suitable boron parameter for specific applications. Similar methods have been used to obtain physicochemical parameters of boron compounds. Other reports dealing with docking studies of boron compounds do not provide specific information on computational strategies addressing this problem.

Figure 2.2: Dimensions and isosurfaces of adamantane (A) and o-carborane (B). Structures were built and minimized (MM$^+$) with HyperChem. Longest distances between two vertices are 3.55 Å (A) and 3.23 Å (B)
Crystal structures of proteins complexed with carborane-containing agents, including those of closo- and nido-carboranyl antifolates with human dihydrofolate reductase (hDHFR), have been reported recently.\textsuperscript{66,85} These provide for the first time the opportunity to verify docking strategies that have been developed previously for carboranyl compounds.\textsuperscript{71} In this chapter, we will evaluate the docking of these carboranyl antifolates into the active site of the hDHFR crystal structures using the docking programs Autodock, Glide, FlexX, and Surflex. The obtained docking poses are compared with the poses of the corresponding carboranyl antifolates in the original hDHFR crystal structures and differences between the docking programs are discussed. The described computational studies will be of great interest for scientists involved in the design and development of boron-containing therapeutics and diagnostics who had very limited access to computational tools in the past.

2.2. Docking

The experimental details related to the hardware/software platforms employed for this study, along with the detailed description related to the ligand preparation (see Section 2.2.4 for short description), protein preparation, and the docking protocols for individual softwares can be found in the experimental section (Chapter 5, Sections 5.1.1- 5.1.4)

2.2.1. Docking algorithms

1) AutoDock 4: It is based on a Lamarckian genetic algorithm (LGA) method. Basically, this program determines total interaction energies between random pairs of ligands and various selected portions of protein to determine docking poses.\textsuperscript{94,95}
2) FlexX: It is a fragment based docking algorithm, which builds putative poses of the ligands using an incremental construction approach.\textsuperscript{96-98} The modeling of protein-ligand interactions and their binding energy predictions is based on the \textit{de novo} design tool LUDI.\textsuperscript{97}

3) Surflex: It generates putative poses for molecular fragments using a surface based molecular similarity method. This program employs the Hammerhead docking system for scoring.\textsuperscript{99,100}

4) Glide: Glide docking uses a series of hierarchical filters to find the best possible ligand binding locations in a pre-built receptor grid space. The filters include a systematic search approach, which samples the positional, conformational, and orientational space of the ligand before evaluating the energy interactions of the ligand with the protein.\textsuperscript{101,102}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{trimethoprim.png}
\caption{Trimethoprim and the carboranyl antifolates 2.1 and 2.2. The positioning of the “extra hydrogen” between B9-B10 corresponds to that in 2 N-crys}
\end{figure}
2.2.2. Crystal Structures

PDB ID # 2C2S: Human dihydrofolate reductase (hDHFR) complexed with [5-(1, 2-closo-dicarbadodecarboran-1-yl) methyl]-2, 4-diamino-6-methylpyrimidine (Compound 2.1 in Figure 2.3). This form of compound 2.1 will be referred to as 1 C-prot throughout this chapter. PDB ID # 2C2T: Human DHFR complexed with a racemic mixture of [5-(7, 8-nido-dicarbaundecarboran-7-yl) methyl]-2, 4-diamino-6-methylpyrimidine (Compound 2.2 in Figure 2.3). This form of compound 2.2 will be referred to as 2 N-prot throughout this chapter.

Small molecule crystal structure data for 2.1 and 2.2 not complexed with hDHFR were generously provided by Dr. Steven Ealick, Cornell University, Ithaca, NY, and Dr. David Borhani, Harvard Medical School, Boston, MA.66 These forms of 2.1 and 2.2 will be referred to as 1 C-crys and 2 N-crys, respectively, throughout this chapter.

2.2.3. Ligand Construction and Optimization

1 C-prot and both enantiomers of 2 N-prot were extracted from their hDHFR crystal structures. The coordinates of the “extra hydrogen” of 2 N-crys were inserted into the structure of 2 N-prot. Other hydrogen atoms were added to both 1 C-prot and 2 N-prot using the “add valence” option of GaussView. Mulliken-103 and APT-104 charges were calculated with Gaussian at AM1-105,106 and HF/6-31+G* levels.107,108 Espfit charges were calculated for 1 C-prot and 2 N-prot at HF/6-31+G* level (Table 2.1). Compounds 2.1 and 2.2 were also constructed de novo without using crystallographic coordinates with HyperChem. These forms of 2.1 and 2.2 will be referred to as 1 C-con and 2 N-con throughout this chapter. In the case of 2 N-con, the nido-carboranyl moieties were
aligned with the closo-carborane moiety of 1 C-prot, the coordinates for the “missing” boron atom were transferred, and then changed to hydrogen. Optimization and Mulliken charge calculations for 1 C-con and 2 N-con were carried at HF/6-31+G* level (Table 2.1). Choosing de novo construction and optimization of 2 N-con, the “extra hydrogen” of the nido-cluster positioned itself between B10 and B11, while in the case of and 2 N-crys, the “extra hydrogen” was located between B9 and B10 (see Figure 2.3 and 2.4). In solution, the “extra hydrogen” appears to be in a fluxional equilibrium between B9/B10 and B10/B11.\(^{109}\) Mulliken charges for 1 C-crys and 2 N-crys were calculated at HF/6-31+G* level (Table 2.1).

![Figure 2.4: Structure of 2 N-crys (A) and 2 N-con (B).](image)

2.2.4. Ligand Preparation

Because of the unavailability of the force field parameters for the hexavalent borons present in carborane, we adopted a strategy developed earlier by Johnsamuel et al,\(^ {70,71}\) which employs the replacement (or the substitution) of hexavalent boron atom type by a
carbon atom type prior to docking with all the softwares used in this chapter. (See Chapter 5, Section 5.1.2 for details regarding ligand preparation for the individual softwares)

2.2.5. Docking comparison

As stated above, the four docking programs are based on different algorithms and it may be difficult to directly compare the results obtained with each of the programs, as has been discussed previously by Cole et al.\textsuperscript{110} Apart from Autodock, the used docking programs did not generate large numbers of duplicate poses within a RMSD of 2 Å. Therefore, a comparison of the avg. RMSDs of all docked poses would be meaningless for FlexX, Surflex, and Glide. However, in the case of the latter three programs, approximately the top 25% of all docked poses (based on the binding energies) were below 4Å. In addition, the top ranked poses reproduced accurately the binding mode of the ligands in the crystal structure. Therefore, we used the top 25% of the poses and the top ranked pose to compare the docking results from FlexX, Surflex, and Glide.

2.2.6. Scoring

Autodock, FlexX, and Glide predict the free energy of binding in terms of kcal/mol.\textsuperscript{94, 97,100} Surflex predicts the binding affinity of the ligand-protein complex in the form of $-\log (K_d)$.\textsuperscript{111} In order to compare scoring features appropriately, the Surflex scoring values were converted into free energy of binding values using the following equation: kcal/mol $= 0.59 \log_e (10^{-pK_d})$.\textsuperscript{111}
2.3. Results and Discussion

2.3.1. Binding site interaction of 1 C-prot and 2 N-prot

The binding interactions of 1 C-prot and both enantiomers of 2 N-prot with amino acid residues in the original hDHFR crystal structures have been discussed previously by Reynolds et al.\textsuperscript{66} 1 C-prot and 2 N-prot showed almost identical binding poses in the hDHFR crystal structures, as shown in Figure 2.5 A-B.\textsuperscript{66} The additional BH vertex of the closo-cage of 1 C-prot is positioned slightly above the center of the open faces of the nido-cages of 2 N-prot. This binding pattern is similar to that of trimethoprim in the active site of hDHFR.\textsuperscript{66}

The carborane cages of 1 C-prot and 2 N-prot bind to the same hydrophobic pocket of hDHFR as the trimethoxyphenyl group of trimethoprim (See Figure 2.5 C).\textsuperscript{66} The 2,4-diamino-5-methylpyrimidine moieties of 1 C-prot and 2 N-prot form hydrogen bonds with various amino acid residues in the active site (Glu 30, Ile 7, Val 115),\textsuperscript{66} as shown in Figure 2.5 A using the example of 1 C-prot.
Figure 2.5: (A) Binding of 1C-prot with various active site amino acid residues of hDHFR. The 2N-prot enantiomers have similar binding interactions in the hDHFR active site (combine with Figure 2.6). 66

(B) Overlap of the crystal structures poses of 1C-prot (red) and 2 N-prot (green). 66

(C) Binding pattern of Trimethoprim in hDHFR. 66
The binding between amino acid residues of the active site with both the neutral \textit{closo}-cage of \textbf{1 C-prot} and the presumably negatively charged \textit{nido}-cages of \textbf{2 N-prot} are typical for hydrophobic interactions. Similar hydrophobic interactions were observed for a metallocarborane, consisting of two negatively charged \textit{nido}-carboranes, bound to HIV protease. Proton–hydride type bonds\cite{112,113} may not play major roles in the interactions of the carboranyl moieties of \textbf{1 C-prot} and \textbf{2 N-prot} with active site amino acid residues, as is evident from the orientations of the Thr56 hydroxyl group in the original hDHFR crystal structures, which points away from the cages (see \textbf{Figure 2.5} A)\cite{66}

\textbf{2.3.2. Docking analysis for docking with Autodock}

An analysis of the docking studies with \textbf{1 C-crys, 2 N-crys, 1 C-prot, 2 N-prot, 1 C-con,} and \textbf{2 N-con} using Autodock is shown in \textbf{Table 2.1} and \textbf{Figure 2.6 A-B}. The number docked poses of the \textit{closo}-carboranyl structures with a RMSD below 2 Å was in all cases high (83 - 100) with minimal differences in avg. RMSDs (0.53 Å– 0.86 Å). With the exception of \textbf{2 N-con}, the \textit{nido}-carboranyl structures showed similar docking patterns with 93-100 poses having an RMSD below 2 Å and avg. RMSDs between 0.5 and 0.74 Å. In the case of \textbf{2 N-con}, 68\% of the docked poses had RMSD’s lower than 2 Å while 32\% had an avg. RMSD of 5.2 Å. A detailed analysis of this minor cluster revealed that the pose of the \textit{nido-o}-carboranyl portion did not show any significant deviation from its counterpart in the original crystal structure while the pose of the 2, 4-diaminopyrimidyl portion was altered forming a strong hydrogen bond with Asp 21 rather than with Glu 30, as in the case of crystallized ligand (\textbf{Figure 2.7}).
<table>
<thead>
<tr>
<th></th>
<th>AM1</th>
<th></th>
<th></th>
<th>HF/6-31+G*</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mulliken</td>
<td>APT</td>
<td>Mulliken</td>
<td>Mulliken</td>
<td>Mulliken</td>
<td>Mulliken</td>
<td>ESPfit</td>
</tr>
<tr>
<td>1 C-prot</td>
<td>0.079</td>
<td>-0.014</td>
<td>0.073</td>
<td>0.118</td>
<td>0.110</td>
<td>-0.0491</td>
<td>0.116</td>
</tr>
<tr>
<td>1 C-crys</td>
<td>0.162</td>
<td>0.117</td>
<td>0.169</td>
<td>0.345</td>
<td>0.326</td>
<td>0.181</td>
<td>0.321</td>
</tr>
<tr>
<td># of poses</td>
<td>98</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>83</td>
<td>99</td>
<td>98</td>
</tr>
<tr>
<td>Avg RMSD</td>
<td>0.581</td>
<td>0.575</td>
<td>0.86</td>
<td>0.530</td>
<td>0.666</td>
<td>0.5515</td>
<td>0.570</td>
</tr>
<tr>
<td>2 N-prot</td>
<td>0.028</td>
<td>-0.099</td>
<td>0.038</td>
<td>0.058</td>
<td>0.056</td>
<td>-0.106</td>
<td>0.035</td>
</tr>
<tr>
<td>2 N-crys</td>
<td>0.101</td>
<td>0.033</td>
<td>0.106</td>
<td>0.247</td>
<td>0.234</td>
<td>0.096</td>
<td>0.248</td>
</tr>
<tr>
<td># of poses</td>
<td>99</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>68</td>
<td>99</td>
<td>93</td>
</tr>
<tr>
<td>Avg binding energy</td>
<td>-7.974</td>
<td>-4.337</td>
<td>-6.405</td>
<td>-2.164</td>
<td>-1.948</td>
<td>-2.195</td>
<td>-1.402</td>
</tr>
<tr>
<td>Avg RMSD</td>
<td>0.629</td>
<td>0.631</td>
<td>0.543</td>
<td>0.741</td>
<td>0.499</td>
<td>0.603</td>
<td>0.574</td>
</tr>
</tbody>
</table>

**Table 2.1:** Optimization characteristics and Autodock docking patterns of carboranyl antifolates 2.1 and 2.2

*Average partial charge of all cage hydrogen atoms connected to boron; partial charge of cage hydrogen atom bound to carbon; # of docked poses with RMSD below 2Å (out of 100 docking solutions); avg. binding energy for all the poses below 2Å ; avg. RMSDs (Å) of all the poses below 2Å.
The enantiomer of *nido*-carboranyl antifolates used for these calculations correspond to the enantiomer of 2 N-prot designated as 2B in the original crystal structure.66

In the case of 2 N-con, the *de novo* construction and optimization resulted in a different location of the “extra hydrogen” compared with 2 N-crys and 2 N-prot (Figure 2.3), which may have contributed to the occurrence of the minor cluster. In general, both enantiomers of 2 showed similar docking patterns (data not shown).

**Figure 2.6:** (A) Overlap of the docked pose (Autodock) of 1 C-con (magenta) with the corresponding pose of 1 C-prot (cyan) in the original crystal structure (avg. RMSD: 0.666 Å) (B) Overlap of the docked pose (Autodock, major cluster with 68 solutions) of 2 N-con (magenta) with the corresponding pose of 2 N-prot (cyan) in the original crystal structure (avg. RMSD: 0.499 Å)

The boron-bound hydrogen atoms (B1-B6, B9-B11) (Figure 2.3) in carboranes are predicted to be slightly more electronegative than carbon-bound hydrogen112,113 and may
have on average negative partial charges. According to ESPfit calculations, the average partial charges for these hydrogens are about -0.24 for two-fold negatively charged nido-carborane \([C_2B_9H_{11}]^2^-\), -0.15 for one-fold negatively charged nido-carborane \([C_2B_9H_{12}]^-\), and -0.07 for neutral nido-carborane \([C_2B_9H_{13}]^-\). The charges generated in our own ESPfit calculations for 1 C-prot and 2 N-prot (Table 2.1) are consistent with those reported in the literature.\(^{112,113}\) However, Mulliken charges obtained for 1 C-prot and 2 N-prot were generally more positive.

![Figure 2.7: Overlap of the docked pose (Autodock) of 2 N-con (magenta) with the corresponding pose of 2 N-prot (cyan) in the original crystal structure. The avg. RMSD of the minor cluster (32 poses) is 5.22 Å.](image)

\[\text{Figure 2.7: Overlap of the docked pose (Autodock) of 2 N-con (magenta) with the corresponding pose of 2 N-prot (cyan) in the original crystal structure. The avg. RMSD of the minor cluster (32 poses) is 5.22 Å.}\]
<table>
<thead>
<tr>
<th>AM1</th>
<th>HF/6-31+G*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mulliken</td>
<td>Mulliken</td>
</tr>
<tr>
<td>1 C-prot</td>
<td>1 C-prot</td>
</tr>
<tr>
<td>1 C-crys</td>
<td>1 C-con</td>
</tr>
<tr>
<td>ESPfit</td>
<td>Mulliken</td>
</tr>
<tr>
<td># of poses⁵</td>
<td>100</td>
</tr>
<tr>
<td>Avg. binding energy⁴ (kcal/mol)</td>
<td>-7.29</td>
</tr>
<tr>
<td>Avg. RMSD⁶ (Å)</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>2 N-prot</td>
</tr>
<tr>
<td></td>
<td>2 N-prot</td>
</tr>
<tr>
<td></td>
<td>2 N-crys</td>
</tr>
<tr>
<td></td>
<td>2 N-prot</td>
</tr>
<tr>
<td></td>
<td>2 N-prot</td>
</tr>
<tr>
<td></td>
<td>2 N-crys</td>
</tr>
<tr>
<td># of poses⁵</td>
<td>99</td>
</tr>
<tr>
<td>Avg. binding energy⁴ (kcal/mol)</td>
<td>-6.17</td>
</tr>
<tr>
<td>Avg. RMSD⁶ (Å)</td>
<td>0.799</td>
</tr>
</tbody>
</table>

Table 2.2: The effect of boron parameter implementation on autodock docking patterns of carboranyl antifolates 2.1 and 2.2.

a Average partial charge of all cage hydrogen atoms connected to boron; b partial charge of cage hydrogen atom bound to carbon; c # of docked poses with RMSD below 2Å (out of 100 docking solutions); Mean Binding Energy, d avg. RMSDs (Å) of all the poses below 2Å. e The enantiomer of nido-carboranyl antifolates used for these calculations correspond to the enantiomer of 2 N-prot designated as 2B in the original crystal structure.

Overall, the docking results obtained with Autodock indicate that minor differences in geometries, stemming from different data sources and/or optimization methods/calculations (e.g. “extra hydrogen” position), and partial atom charges did not have a major impact on docking accuracy although they seemed to effect the avg. binding
energies to some extent (Table 2.1). In general the binding energies were somewhat lower for the nido-carboranyl antifolates compared with their closo-carboranyl counterparts.

The force field parameters developed for sp3 borons by Otkidach et al.\textsuperscript{90,114} were partially implemented in AutoDock for docking of the carboranyl antifolates. The described docking model appears to remain unaffected by this substitution (Table 2.2), which further validates our strategy to enable docking of biomolecules containing boron clusters by replacing boron with carbon. It should be noted, however, that apart from AutoDock the implementation of these parameters in the other softwares packages discussed in this chapter is not straightforward.

2.3.3. Docking analysis for docking with Glide

The Glide docking program\textsuperscript{101,102} uses ConfGen, a systematic sampling method to generate several conformations of the ligand prior to the actual docking calculation. Initially, the switch of boron atom type to C.3 atom type (see Chapter 5, Section 5.1.2 for details about ligand preparation) using the ConfGen failed to generate conformations of the ligands because the atomic bonds in the cage structures of the ligands were considered as torsions. Therefore, we adopted a rigid docking protocol, which did not directly assess the conformational sampling capability in Glide.

The impact of charge differences on the carborane clusters as a result of different data sources and/or optimization methods/calculations was also explored with Glide (see Table 2.3). As in the case of Autodock docking, the binding energies were lower for the nido-carboranyl antifolates compared to the nido-carboranyl counterparts.
<table>
<thead>
<tr>
<th>AM1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mulliken</td>
<td>APT</td>
<td>Mulliken</td>
<td>Espfit</td>
<td>Mulliken</td>
<td>Mulliken</td>
<td>Mulliken</td>
<td>Mulliken</td>
</tr>
<tr>
<td>RMSD of the top ranked pose (Å)(^a)</td>
<td>0.10</td>
<td>0.12</td>
<td>0.19</td>
<td>0.15</td>
<td>0.13</td>
<td>0.17</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Avg. RMSD of the top 25% of poses (Å)(^b)</td>
<td>0.16</td>
<td>0.14</td>
<td>0.14</td>
<td>0.14</td>
<td>0.12</td>
<td>0.23</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>No. of docking solutions</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>RMSD of the top ranked pose (Å)(^a)</td>
<td>0.20</td>
<td>0.20</td>
<td>0.19</td>
<td>0.13</td>
<td>0.22</td>
<td>0.29</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Binding energy of the top ranked pose (kcal/mol)</td>
<td>-7.53</td>
<td>-7.51</td>
<td>-8.41</td>
<td>-8.13</td>
<td>-8.11</td>
<td>-7.56</td>
<td>-7.90</td>
<td></td>
</tr>
<tr>
<td>Avg. RMSD of the top 25% of poses (Å)(^b)</td>
<td>0.20</td>
<td>0.20</td>
<td>0.69</td>
<td>0.22</td>
<td>0.20</td>
<td>0.29</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>Avg. binding energy of the top 25% poses (kcal/mol)</td>
<td>-7.22</td>
<td>-7.24</td>
<td>-6.59</td>
<td>-7.84</td>
<td>-7.97</td>
<td>-7.56</td>
<td>-7.80</td>
<td></td>
</tr>
<tr>
<td>No. of docking solutions</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.3.** Glide docking of closo- and nido-carboranyl antifolates listing the impact of partial charges on docking outcomes.

\(^a\)Pose with the lowest binding energy, \(^b\)top 25% poses ordered according to binding energies. See **Table 2.1** for partial charges of cage hydrogens bound to carbon and boron.
Overall, Glide reproduced the crystal binding modes closely as can be seen by the average RMSDs of the top 25% poses. The average binding energy of the top 25% poses was also very close to the binding energy of the top ranked pose and the differences in avg. binding energies, as a result of different geometries and partial atom charges were not as pronounced as in the case of Autodock while RMSDs and pose clustering were comparable.

2.3.4. Docking analysis for docking with FlexX.

Since FlexX and Surflex operate with formal charges, only docking of 1 C-/2 N-crys and 1 C-/2 N-con was evaluated with these programs. The later carboranyl antifolates were chosen to address the influence of geometric differences on the docking performances.

Automated docking of 1 C-crys and 2 N-crys by FlexX produced top ranked poses with RMSDs of 3.19 Å and 3.18 Å, respectively (Table 2.4). The top 25% of the docked poses for both 1 C-crys and 2 N-crys had avg. RMSD’s of 5.97 Å and 5.89 Å respectively (Table 2.4). The automated mode selected the PPs (pyrimidyl portion, see Chapter 5, Section 5.1.2. for details related to the preparation of PP) of 1 C-crys and 2 N-crys as the base fragments and PA3 (placement algorithm 3, see Chapter 5, Section 5.1.4. for experimental details) before incremental construction of the ligands.

Figure 2.8 A-B shows how either PP or CP (carboranyl portion, see Chapter 5, Section 5.1.2. for details related to the preparation of PP) is placed inside the hDHFR active site pocket, before the incremental construction of the whole ligand. User-specific base fragment placement (BFP) of the PP of 1 C-crys followed by PA1 (placement algorithm 1), and PA2 (placement algorithm 2), improved the RMSDs of the top-ranking
poses by factors of 1.6 and 2, respectively, and the avg. RMSDs of top 25% poses by factors of 3.2 and 1.4, respectively.

**Figure 2.8:** (A) Pyrimidyl Portion (PP) and, (B) *closo-o*-Carboranyl Portion (CP) placed inside the active site of hDHFR using base fragment placement (BFP) option implemented in FlexX.
<table>
<thead>
<tr>
<th></th>
<th>Automated docking</th>
<th>User-Specific Docking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PP1&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PA1&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>1 C-crys</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSD of the top ranked pose (Å)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>3.19</td>
<td>2.01</td>
</tr>
<tr>
<td>Binding energy of the top ranked pose (kcal/mol)</td>
<td>-10.73</td>
<td>-25.64</td>
</tr>
<tr>
<td>Avg. RMSD of the top 25% poses (Å)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>5.97</td>
<td>1.84</td>
</tr>
<tr>
<td>Avg. binding energy of the top 25% poses (kcal/mol)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-7.18</td>
<td>-21.10</td>
</tr>
<tr>
<td>Pose with the lowest RMSD (Å) and its energy rank</td>
<td>3.19 (1)</td>
<td>0.96 (7)</td>
</tr>
<tr>
<td>Total no. of docking solutions</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td><strong>2 N-crys</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSD of the top ranked pose (Å)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>3.18</td>
<td>1.27</td>
</tr>
<tr>
<td>Avg. RMSD of the top 25% poses (Å)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>5.89</td>
<td>3.50</td>
</tr>
<tr>
<td>Avg. binding energy of the top 25% poses (kcal/mol)&lt;sup&gt;g&lt;/sup&gt;</td>
<td>-9.82</td>
<td>-15.44</td>
</tr>
<tr>
<td>Pose with the lowest RMSD (Å) and its energy rank</td>
<td>3.18 (1)</td>
<td>1.24 (5)</td>
</tr>
<tr>
<td>Total no. of docking solutions</td>
<td>28</td>
<td>26</td>
</tr>
</tbody>
</table>

**Table 2.4:** The effects of docking mode base fragment for placement (BFP), and placement algorithm (PA) on FlexX docking of **1 C-crys** and **2 N-crys**

<sup>a</sup>Docking using the pyrimidyl portion (PP) for base fragment placement (BFP); <sup>b</sup>docking using the closo-o-carboranyl portion (CP) for BFP. <sup>c</sup>Placement algorithm 1 (PA1); <sup>d</sup>Placement algorithm 2 (PA2); <sup>e</sup>Placement algorithm 3 (PA3); <sup>f</sup>Top ranking pose according to binding energies, <sup>g</sup>Top 25% poses according to binding energies. PA1 did not result in docking solutions for 1 C-crysCP and 2 N-crysCP.
The corresponding improvement factors in the case 2 N-crys were 2.5, 1.6, 1.7, and 1.0. As expected, user-specific BFP of PP, followed by PA3, produced the same results as automated docking. 1 C-con and 2 N-con produced similar FlexX docking results indicating that in the absence of crystal structure information, binding interactions of virtual carboranyl compounds with proteins can be predicted accurately (see Table 2.5).

For BFP, both in automated and user specific mode, FlexX favors PP. User-specific BFP of CPs, in particular in the cases of 2 N-crys and 1 C-/2 N-con, was inferior to BFP of PP indicating that hydrogen bonds and/or ionic interactions may be favored in this initial phase of the FlexX docking.96-98 This may also explain the fact that user-specific BFP of CPs, if at all, produced poor docking solutions with high RMSDs. PA1 was found to be superior to PA2 and PA3 following initial BFP of PP. PA1 only utilizes interactions that are hydrophobic and unspecific in nature.96 This indicates that the second phase of incremental construction in FlexX docking, involving the PAs, is mainly guided by the hydrophobicity of the CPs.

2.3.5. Docking analysis for docking with Surflex.

The results of Surflex docking of 1 C-crys and 2 N-crys are summarized in Table 2.6. Using default docking parameters without fragment placement, the top 25% poses of 1 C-crys have an avg. RMSD of 0.78 Å and the top ranking pose had an RMSD of 0.89 Å.
Table 2.5: The effects of docking mode base fragment for placement (BFP), and placement algorithm (PA) on FlexX docking of 1 C-con and 2 N-con

<table>
<thead>
<tr>
<th></th>
<th>Automated docking</th>
<th>User-Specific Docking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PP1(^a) PA1(^c)</td>
<td>CP(^b) PA2(^d) PA3(^e)</td>
</tr>
<tr>
<td><strong>1 C-con</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSD of the top ranked pose (Å)(^f)</td>
<td>2.54</td>
<td>2.12</td>
</tr>
<tr>
<td>Binding energy of the top ranked pose (kcal/mol)</td>
<td>-14.61</td>
<td>-25.40</td>
</tr>
<tr>
<td>Avg. RMSD of the top 25% poses (Å)(^g)</td>
<td>4.78</td>
<td>1.62</td>
</tr>
<tr>
<td>Pose with the lowest RMSD (Å) and its energy rank</td>
<td>2.31 (24)</td>
<td>0.51 (7)</td>
</tr>
<tr>
<td>Total no. of docking solutions</td>
<td>200</td>
<td>198</td>
</tr>
</tbody>
</table>

|                  |                   |                       |
| **2 N-con**      |                   |                       |
| RMSD of the top ranked pose (Å)\(^f\) | 3.18 | 1.81 | 1.79 | 3.18 | 3.30 | 5.93 |
| Binding energy of the top ranked pose (kcal/mol) | -14.27 | -21.00 | -23.98 | -14.27 | -16.87 | -11.86 |
| Avg. RMSD of the top 25% poses (Å)\(^g\) | 4.18 | 2.22 | 3.96 | 4.18 | 4.89 | 5.85 |
| Avg. binding energy of the top 25% poses (kcal/mol)\(^g\) | -11.68 | -17.00 | -12.44 | -11.68 | -8.70 | -8.45 |
| Pose with the lowest RMSD (Å) and its energy rank | 3.18 (1) | 0.76 (11) | 0.89 (2) | 3.18 (1) | 1.96 (32) | 2.49 (5) |
| Total no. of docking solutions | 25 | 48 | 102 | 25 | 51 | 48 |

\(^a\)Docking using the pyrimidyl portion (PP) for base fragment placement (BFP); \(^b\)docking using the closo-o-carboranyl portion (CP) for BFP. \(^c\)Placement algorithm 1 (PA1); \(^d\)Placement algorithm 2 (PA2); \(^e\)Placement algorithm 3 (PA3); \(^f\)Top ranking pose according to binding energies; \(^g\)Top 25% poses according to binding energies. PA1 did not result in docking solutions for 1 C-conCP and 2 N-conCP.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>1 C-crys</th>
<th>2 N-crys</th>
<th>2 N-crys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
<td>Default settings</td>
<td>multistart 5&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Default settings</td>
</tr>
<tr>
<td>RMSD of the top ranked pose (Å)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.883</td>
<td>1.998</td>
<td>0.741</td>
</tr>
<tr>
<td>Binding energy of the top ranked pose (kcal/mol)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-9.74</td>
<td>-9.93</td>
<td>-7.12</td>
</tr>
<tr>
<td>Avg. RMSD of the top 25% poses (Å)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.78</td>
<td>1.41</td>
<td>3.25</td>
</tr>
<tr>
<td>Avg. total binding energy of the 25% poses (kcal/mol)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-9.58</td>
<td>-9.78</td>
<td>-7.06</td>
</tr>
<tr>
<td>Pose with lowest RMSD (Å) and its energy rank indicated in bracket</td>
<td>0.448 (4)</td>
<td>0.448 (15)</td>
<td>0.646 (6)</td>
</tr>
<tr>
<td>No. of docking solutions</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**Table 2.6: Surflex docking of 1 C-crys and 2 N-crys**

<sup>a</sup>Pose with the lowest binding energy; <sup>b</sup>top 25% poses with lowest binding energy; <sup>c</sup>docking with 5 different starting positions; PP: Pyrimidyl Portion; CP: Carboranyl Portion.

Docking of 1 C-crys using the “multistart5” option (5 different starting positions) without fragment placement decreased the docking quality of both top 25% poses and top ranked pose by a factor ~ 2 (Table 2.6). Docking with fragment placement was not explored because the default settings produced satisfactory docking results. In the case 1
**C-con**, there were no significant differences in docking quality between default settings and the “-multistart5” option (see Table 2.7).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>1 C-con</th>
<th>2 N-con</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters default settings</td>
<td>Default settings</td>
<td>Default settings</td>
</tr>
<tr>
<td>RMSD of the top ranked pose (Å)(^a)</td>
<td>1.07</td>
<td>3.94</td>
</tr>
<tr>
<td>Binding energy of the top ranked pose (kcal/mol)</td>
<td>-10.37</td>
<td>-7.19</td>
</tr>
<tr>
<td>Avg. RMSD of the top 25% poses (Å)(^b)</td>
<td>1.09</td>
<td>3.94</td>
</tr>
<tr>
<td>Avg. binding energy of the top 25% poses (kcal/mol)</td>
<td>-10.28</td>
<td>-7.17</td>
</tr>
<tr>
<td>Pose with the lowest RMSD (Å) and its energy rank indicated in brackets</td>
<td>1.07 (1)</td>
<td>2.41 (16)</td>
</tr>
<tr>
<td>No. of docking solutions</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**Table 2.7**: Surflex docking of **1 C-con** and **2 N-con**

\(^a\)Pose with the lowest binding energy; \(^b\)top 25% poses with lowest binding energy; \(^c\)docking with 5 different starting position; PP: Pyrimidyl Portion; CP: Carboranyl Portion.

Docking of **2 N-crys** using default parameters without fragment placement was unsatisfactory resulting in an avg. RMSD of the top 25% poses of 3.25 Å. The
“multistart5” option without fragment placement improved the docking by a factor of 1.65 (1.97 Å). The difference in RMSDs of the top ranked poses was insignificant (0.741 vs. 0.755 Å). These values improved even further for docking with fragment placement. When PP was selected as a placed fragment, the corresponding avg. RMSD and the top ranking pose RMSD decreased to 0.731 and 0.671 Å, respectively. The corresponding values for fragment placement of CPs were 1.26 Å and 0.463 Å. This is in contrast to FlexX, where BFP of the nido-CP led to a significant deterioration of docking quality. Overall, the Surflex docking results for 2 N-con showed similar patterns as those for 2 N-crys (see Table 2.7).

In preliminary studies, we also explored the Molecular Operating Environment (MOE) program for docking of carboranyl antifolates into hDHFR. Docking was carried out in standard mode. MOE is based on simulated annealing, which calculates the grid based interaction energy between the docked ligand and the receptor/enzyme.\textsuperscript{116,117} MOE has parameter sets for sp2 and sp3 boron atoms but not for hexavalent boron. If at all, the docking performance of MOE seemed to be comparable with that of FlexX in automated mode. Both replacement of boron with “C.3” and the use of the boron atom types provided by MOE produced similar docking results.

**2.3.6. Binding energy calculations of the docked poses (Scoring)**

Enzymatic studies indicated that antifolate 2.1 is approximately ten times more potent than 2.2 as an inhibitor of hDHFR,\textsuperscript{66} which is consistent with the binding energies that were predicted by all four docking programs for 2.1 and 2.2 (Figure 2.9). However, these free energy values should be viewed with caution. Scoring functions are in general not
very accurate with significant differences between various docking algorithms\textsuperscript{118} and hydrophobic interactions were calculated without applying appropriate force fields for hexavalent boron.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{binding_energy_graph.png}
\caption{Average total binding energies of docked 1 C-crys (black bar) and 2 N-crys (grey bar) with hDHFR.}
\end{figure}

For AutoDock and Glide, Mulliken charges at AM1 level were considered for 1 C-crys and 2 N-crys. For Surflex, FlexX and Glide, the avg. binding energies for top 25\% poses was considered. For AutoDock, the avg. binding energies of all poses were considered.

2.4. Conclusions and future outlook

In this chapter we report for the first time docking studies with compounds containing the negatively charged \textit{nido}-carboranyl cluster. Using 1 C-crys and 2 N-crys as examples, the docking performances of Autodock, Glide, FlexX, and Surflex are
summarized in Figure 2.10 A-B. Under optimized conditions (e.g. rigid docking and fragment placements), docking of 1 C-crys with Autodock and Glide were comparably good followed by Surflex and then FlexX. In the case of 2 N-crys, Autodock, Glide, and Surflex showed comparably good docking followed by FlexX. Overall, a comparative evaluation of 1 C-con and 2 N-con produced similar results (Figure 2.11 A-B)

Differences in geometries and partial atom charges resulting from different data sources and/or optimization methods/calculations did not impact the docking performances of Autodock and Glide significantly. A careful selection of placement algorithm and/or fragment placement improved the docking quality of both FlexX and Surflex. The scoring values generated by all four programs were in accordance with experimental data.

Carborane cages are rigid, hydrophobic entities and, intuitively, the docking of these scaffolds could be visualized as the docking of a single hydrophobic entity into the hydrophobic pocket of hDHFR. From this perspective, it is difficult to envision how the interactions of this single entity with its amino acid environment are changed during the docking process by replacing hexavalent boron either with carbon or sp3 boron. Although the reported docking protocols could be validated successfully, there is certainly further need for the development of accurate forcefield parameters, probably for the entire cage structures rather than for individual atom types.

A drawback of the present study was that the carboranyl antifolates used for docking only contained two rotatable bonds at the methylene bridge between the carboranyl- and the pyrimidyl portions. Consequently, the torsional degrees of freedom were limited in the present study and evaluating more complex systems may be necessary to confirm that
the reported docking protocols are also applicable to other biomolecules that contain carborane clusters. Unfortunately, such systems are currently not available.

Figure 2.10 A-B: Comparative docking evaluation by Autodock, Surflex, FlexX, and Glide for 1 C-crys (A), 2 N-crys (B). continued on the next page.

See the next page (Figure 2.11 A-B) for the detailed description of this figure.
Figure 2.11 A-B: Comparative docking evaluation for 1 C-con (C), 2 N-con (D) by Autodock, Surflex, FlexX, and Glide.

The black bars indicate the percentage of the total poses with RMSD below 2 Å and the grey bars indicate the percentage of the total poses with RMSDs below 4 Å. Data used for this presentation are based on Mulliken charges obtained at HF/6-31+G* level for Autodock and Glide, rigid docking mode for Glide, and BFP of PP with PA1 for FlexX. For Surflex docking, default settings without fragment placement were used for 1 C-crys and 1 C-con and PP placement combined with multistart 5 option for 2 N-crys and 2 N-con.
3.1) Monoradioiodination of carboranyl thymidine analogues for the imaging and treatment of cancer: A case study of B-I bond formation with a controlled level of iodination and regioselectivity

3.2) A general study of the deglycosylation and anomerization of N3-substituted thymidine- and uridine derivatives caused by electrophilic iodine (I$^+$)

Contributors other than myself:

3) Antonio Toppino (Graduate Student, Dipartimento di Chimica Generale e Chimica Organica, Universita’ degli Studi di Torino, Torino, Italy)
4) Naum Sayfullin (BSPS student, College of Pharmacy, The Ohio State University)
5) Trang Nguyen (Pharm.D. student, College of Pharmacy, The Ohio State University)
3.1. Monoradioiodination of carboranyl thymidine analogues for the imaging and treatment of cancer: A case study of B-I bond formation with a controlled level of iodination and regioselectivity

3.1.1. Introduction

Radiolabeled nucleosides, in particular 2′-deoxyuridine labeled with various radioisotopes of iodine, have attracted considerable attention in the field of cancer therapy and imaging. A major problem associated with these radioiodinated nucleosides, which usually possess carbon linked radioiodine, is their susceptibility to dehalogenation, which can lead to the accumulation of significant quantities of radioiodine in the thyroid and the stomach. In the case of 5-iodo-2′-deoxyuridine, enzymatic cleavage at the N-glycosidic bond by nucleoside phosphorylases occurs within few minutes and is followed by rapid dehalogenation of the resulting 5-iodouracil.

Radiodiagnosics/therapeutic with iodine linked to boron atoms may be superior to those having carbon-iodine bonds because the B-I bond is presumably less susceptible to enzymatic and hydrolytic cleavage than the C-I bond. No enzymatic system has been reported that cleaves the boron-iodine bonds in vivo. Also, the boron-iodine bond is stronger than the carbon-iodine bond (220 kJ/mol versus 209 kJ/mol) and iodine attached to a boron atom of the carborane cluster (B-iodocarboranes) displays an extraordinary chemical stability under various chemical reaction conditions. Tolmachev et al. carried out in vivo stability studies with dextran conjugated to 125I-labeled undecahydro-closo-dodecaborate (B_{12}H_{11}I^2). They demonstrated that the
accumulation of the radioactivity released into the thyroid from this conjugate was negligible. Indeed, all of its degradation products were rapidly eliminated without accumulation in any organ.\textsuperscript{138} Therefore, boron clusters have been investigated intensively over the past decade as “prosthetic groups” for radiohalogenation in order to overcome the undesired release of radiohalogen.\textsuperscript{138,141-150}

The radiohalogen isotopes that are useful for therapeutic applications and imaging (e.g. PET [Positron Emission Tomography and SPECT [Single Photon Emission Computed Tomography]]) are listed in Table 3.1.\textsuperscript{132} In PET imaging, positrons are emitted by radionuclides, such as $^{18}$F, $^{122}$I, $^{124}$I, $^{75}$Br, and $^{76}$Br, which then interact with a negatron (electron) in an annihilation process. This produces two coincident 511 keV photons, which are detected simultaneously in a detector ring. Radionuclides that are useful in SPECT imaging, such as $^{123}$I, emit photons in high abundance and have high enough energy (>100 keV) to readily escape the body to be detected.\textsuperscript{132} Radiohalogens of interest for therapeutic applications are $^{125}$I, $^{131}$I, $^{77}$Br, and $^{211}$At, which are either $\beta^-$- or $\alpha$-particle emitting isotopes. The choice of radiohalogens employed for a particular need is dictated by several factors, such as availability, half-life, physical characteristics (decay, positron-$\gamma$-emission), and ease of radiolabeling chemistry.\textsuperscript{132} The most frequently encountered isotope in PET imaging is $^{18}$F because of its high percentage of $\beta^+$ emission combined with its relatively low energy whereas $^{131}$I is the most commonly used in therapy.\textsuperscript{132} For laboratory studies and for radioimmunoassay, $^{125}$I is a frequently used radionuclide since it has a relatively longer half-life (~ 60 days), which gives researchers adequate time to carry out a range of studies with this isotope.
Auger electron emitters, such as $^{125}$I, are advantageous compared with other radionuclides described above because of their potential use in both imaging and radiotherapy.$^{127,135}$ Iodine-125 decays primarily by internal conversion following electron capture, which generates extremely low energy (<1 keV) Auger electrons. When incorporated into the DNA of dividing mammalian cells, the decay of this isotope leads to DNA double strand breaks and an exponential reduction in cell survival. The isotope also leads to a lower oxygen enhancement ratio compared with x-ray exposure. The extracellular and even the cytosolic decay of $^{125}$I produces no lethal effects.$^{127,135}$

As already discussed in Chapter 1, the prototype 3CTA, N5, (Figure 1.5 and Figure 3.1) is an excellent substrate of hTK1$^{20,151,152}$ and has shown resistance to degradation by thymidine phosphorylase. Also, its monophosphate form is not a substrate of 5′-nucleotidase.$^{39}$ Based on these finding and the discussion above, cage-radioiodinated N5 (henceforth referred to as N5-I, (3.2, Figure 3.1) may be superior to conventional radiolabeled nucleosides, such 5-iodo-2′-deoxyuridine. Radioiodinated 3.1 could be used.
in cell culture and biodistribution/pharmacokinetic experiments in rodent/tumor models as well as in the radiotherapy and imaging of cancer.\textsuperscript{87,132,156} It could also be used to monitor the real time biodistribution of 3.1 in clinical BNCT.

\textbf{Figure 3.1}: The 2D structures of 3.1 and 3.2 along with:

(A) The molecular surface of N5 (3.1)

(B) The molecular surface of N5-I (3.2)

(C) The molecular surface of an unsubstituted o-carborane

(D) The molecular surface of monoiodinated o-carborane
Human TK1 tolerates extensive modification at N-3 position of dThd,\textsuperscript{151,157-160} as discussed in Chapter 1, Sections 1.4 and 1.6.1, and it can be hypothesized that the introduction of single iodine to the carborane cage of a 3CTA, such as 3.1, will not affect its phosphorylation by hTK1 because the carborane cage in these compounds presumably is located outside of the substrate binding site (see Chapter 1, Figure 1.7).\textsuperscript{16}

The differences in the surface areas and the molecular volume between 3.1 and 3.2 are relatively small with 575.53 Å\textsuperscript{2} vs. 608.24 Å\textsuperscript{2} and 1297.05 Å\textsuperscript{3} vs. 1340.1 Å\textsuperscript{3}, respectively (Figure 3.1). Therefore, the introduction of a single iodine atom at the carborane cluster may not affect the hTK1 substrate characteristics of 3.1, and consequently its tumor localizing properties, significantly.

Lipinski’s “Rule of Five” states that passive oral absorption through the intestinal epithelium becomes unlikely if a molecule has a MW $\geq$ 500 Da, $\geq$ 5 H-bond donors, $\geq$ 10 H-bond acceptors, and a log P $\geq$ 5.\textsuperscript{161,162} Principally, these rules also apply for the passive diffusion of a drug through the membrane of a cancer cell. Preliminary studies in our laboratories indicated that 3CTAs are not substrates for cell membrane nucleoside transporters.\textsuperscript{92} In addition to their hTK1 substrate characteristics, the physicochemical properties of carborane cage radiiodinated 3CTAs may therefore play a pivotal role in their tumor accumulation.\textsuperscript{39,40,158,163}

Introduction of iodine to 3.1 to obtain 3.2, will increase the MW and this may pose an obstacle for cell membrane permeability. However, the MW of 3.2 is still fairly low with 580.24 Da, which may be an acceptable value considering that the H-bond acceptor/donor numbers (5/2) and the log P value (+ 3.62) for 3.2 are within the limits of Lipinski’s “Rule of Five”. The Log P value of 3.2 was calculated with ChemDraw Ultra,
Cambridge, MA, using a procedure previously described by our group. Based on the discussion above, we did embark on the synthesis of N5-I (3.2).

3.1.2. Synthetic Strategies for the iodination of N5.

![Structure of o-carborane with atoms numbers](image)

**Figure 3.2:** Structure of o-carborane with atoms numbers. The carbon-bound hydrogen atoms are referred to as carboranyl C-Hs throughout this chapter. The substitution of carboranyl C-Hs by any other group (e.g. alkyl) will be described as “C-substitution” and the substitution of boron-bound hydrogens by any other group (e.g. halogen) will be described as a “B-substitution”.

Theoretically, radioiodination of the *closo-o*-carborane cage can be achieved by three methods. The first method employs a “*closo*-cluster reconstruction” from *nido-o*-carborane via treatment with *n*-butyl lithium (*n*-BuLi) and BI$_3$ to generate iodinated *closo-o*-carboranes. Thus the B-iodocarborane synthesized in this way will have iodine either at B3 or at B6 (see Scheme 3.1).
Scheme 3.1: General reaction scheme for the synthesis of 3-iodo-o-carborane via “closo-cluster reconstruction” from nido-o-carborane (7,8-dicarba-nido-undecaborane). Structures are enlarged to highlight the change in the numbering of carbon and boron atoms of from the nido- to the closo-form.

The second method is the direct introduction of iodine to the closo-o-carborane by means of a Friedel Craft-type (electrophilic) halogenation strategy. The electrophilic attack of positively charged halogens on closo-o-carborane is regioselective, based on the charge distribution present in the cluster.\textsuperscript{74,166,167} Therefore, electrophilic halogenation occurs first at the most electron rich boron atoms 9 and 12, which are furthest away from both carbon atoms followed by the boron atoms 8 and 10 (see Scheme 3.2). The boron atoms 3-7 and 11 are directly linked to the carbon atoms, and thus, less prone to undergo electrophilic halogenation because of their relatively high electron deficiencies.\textsuperscript{74,166,167}

In summary, the high electron density at the boron atoms 8-10 and 12 makes them most susceptible to electrophilic substitution in presence of Lewis acids and I\textsubscript{2} or iodine monochloride (ICl) resulting into the formation of mono- or di-, tri- or tetraiodination,
depending on applied the reaction conditions and quantities (stoichiometry) of halogenating species.

**Scheme 3.2**: Direct iodination of \(o\)-carborane using \(I_2\) and \(AlCl_3\). Only the monoiiodinated (9-iodo-o-carborane, compound **3.6, Scheme 3.5**) and the tetraiodinated (8, 9, 10, 12-tetraiodo-o-carborane) product are shown. Structures are enlarged to show the numbering of carbon and boron atoms of the cluster.

Both methods are, however, not suitable for the synthesis of \(^{125}\text{I}\)-labeled **3.2** because \(^{125}\text{I}_2\), \(^{125}\text{ICl}\), and \(^{125}\text{B} \text{I}_3\) are not readily available from commercial sources.

The third method is based on a halogen-exchange strategy. Stanko and Iroshnikova reported the first radioiodination of “cold” \(o\)-, \(m\)-, and \(p\)-iodocarboranes via an isotopic exchange reaction with \(Na^{131}\text{I}\) in presence of iron (II) sulfate as the catalyst.\(^{168}\) Later, Marshall *et al.* reported the first Pd-catalyzed halogen exchange reaction.\(^{169}\) In this case, 9-iodo-\(m\)-iodocarborane was reacted with \(NaBr\) in the presence of Pd-catalyst to obtain 9-iodo-\(m\)-bromocarborane (**Scheme 3.3**). Subsequently, Sjoeberg and coworkers adopted
this strategy for the synthesis of several radiohalogenated \textit{closo}-carboranes (\textbf{Scheme 3.3}).\textsuperscript{170-172} The advantage of this method is that commercially available Na\textsuperscript{125}I can be used. This strategy was therefore chosen for the synthesis of \textsuperscript{125}I-labeled \textbf{3.2}.

\begin{center}
\includegraphics[width=0.7\textwidth]{scheme3.3.png}
\end{center}

\textbf{Scheme 3.3}: The halogen/isotope exchange reaction using “cold” B-halocarboranes.\textsuperscript{168, 170, 171}

For the synthesis of \textsuperscript{125}I-labeled \textbf{3.2} by means of isotope exchange reaction, its “cold” precursor, \textsuperscript{127}I-\textbf{3.2}, has to be synthesized first. This can be accomplished by three methods: 1) Direct iodination of carborane cage of \textbf{3.1} using either I\textsubscript{2} or ICl in presence of a Lewis acid, 2) direct iodination of unsubstituted \textit{closo-o}-carborane using the procedure described above, followed by the attachment of the iodinated \textit{o}-carborane to the N3 position of dThd via a pentylene spacer, or 3) deboronation of \textit{o}-carborane using TBAF or a strong Brønsted base (see \textbf{Chapter 2, Section 1}), followed by reconstruction with BI\textsubscript{3} in presence of \textit{n}-BuLi (see \textbf{Scheme 3.1}) and attachment of the iodinated \textit{o}-carborane to the N3 position of dThd via a pentylene spacer.

The first strategy may result in iodination of \textbf{3.1} with varying substitution levels ranging from mono- to tetraiodination (\textbf{Section 3.1.3.1}). The second strategy will result
only in monoiiodination of 3.1. As will be discussed in Section 3.1.3.2, however, this method will generate a mixture of geometric isomers of 3.2 with iodine substitution at either B9 or B12 of the carborane cluster. The third strategy will also result in monoiiodinated 3.1 with iodine at either the B3 or B6 position, forming an epimeric mixture of 3.2.74,165,173-178 Thus, the level and the regioselectivity of iodination can be controlled by selecting one of the methods described above. We decided to focus on strategies 1 and 2 because the 3rd strategy requires an additional reaction step compared with strategy 2.74,165,173,174,178

3.1.3. Chemistry

3.1.3.1. Direct iodination of preformed N5 (3.1) [Strategy 1]

A number of different reaction conditions were explored to optimize the synthesis of 3.2 via direct iodination of 3.1. These are summarized in Table 3.2. When 3.1 was reacted with 1 equivalent of ICl in the presence of 10 equivalents of AlCl₃ in dichloromethane at 0 °C for 2 h, (Entry 1, Table 3.2) ~ 40 % of the desired monoiiodinated 3.2 was produced along with approximately 10 % of deglycosylated nucleosides (compound 3.5, Scheme 3.4), ~ 45 % of unreacted 3.1, and a minor quantity (< 5%) of 3.1 that was diiodinated at the carborane cluster. Nucleobases were separated from all β-nucleosides by silica gel column chromatography followed by isolation of 3.2 from the nucleoside mixture via semi-preparative HPLC (see experimental Section [Chapter 5] for details).
When 3.1 was reacted with 5 equivalents of ICl and 10 equivalents of AlCl₃ at 40 °C in dichloromethane for 48 h (Entry 5, Table 3.2), ~ 85 % of 3.3 (tetraiodinated 3.1) was produced along with ~15 % triiodinated 3.1, ~ 30% deglycosylated forms of 3.1, and 2 % anomerized forms of 3.1 (compound 3.4, Scheme 3.4). Nucleobases were separated from α- and β-nucleosides by silica gel column chromatography followed by isolation of 3.3 from the nucleoside mixture via semi-preparative HPLC (see Chapter 5 for details). Surprisingly, when the reaction was carried out with I₂ instead of ICl using the same conditions (Entry 6, Table 3.2,) significantly higher quantities of nucleobases and α-nucleosides were produced. Reaction conditions shown in Entries 2-4 of Table 3.2, which were harsher than those in Entry 1 and milder than those in Entry 5, produced varying quantities of nucleobases, α-nucleosides, and β-nucleosides with varying levels of iodination.

Only in the cases of the reactions described in Entries 1 and 5 of Table 3.2, products 3.2 and 3.3 were isolated and thoroughly characterized by HRMS as well as ¹H-, ¹³C-, and ¹¹B NMR (see Chapter 5 for details). All nucleosides had similar R_f values (0.34, dichloromethane/methanol, 10:1) independent of their level of iodination and configuration at the anomeric carbon. Almost identical R_f values (0.72, dichloromethane: methanol, 15:1) were also found for all nucleobases despite of different iodination levels. Therefore, isolation of individual products was not possible by normal silica gel chromatography. If at all, the nucleosides were crudely separated from nucleobases because the difference in R_f values between both groups of compounds was sufficient.
Scheme 3.4: Direct iodination of 1 using ICl or I₂ in presence or absence of 10 eq. AlCl₃

A partial characterization by means of ¹H NMR and mass spectrometry was attempted in the case of the separated nucleobases from the reaction shown in Entry 5 of Table 3.2. As expected, the ¹H NMR spectrum of this mixture did not show any signals for deoxyribose protons and a HR-ESI analysis indicated that it contained at least two nucleobases, one di- and the other tetraiodinated. No attempt was undertaken to isolate and characterize individual α-nucleosides or mixtures thereof. In most cases, the product
ratios shown in Scheme 3.4 were estimated based on an analysis of spot intensities on TLC plates and/or $^1$H NMR and/or $^{11}$B NMR and/or MS data for complete product mixtures and/or the crudely separated subsets of $\alpha$-nucleosides and $\beta$-nucleosides. The level of iodination of various compounds in these mixtures could be determined effectively by the integration ratios of the signals for the carbon-bound proton of the carborane cage of individual mixture components. This phenomenon will be discussed in more detail Section 3.1.4.1 (see Figure 3.3). Another important tool was the analysis of the chemical shifts for the boron atoms of the carborane cluster, which changed dramatically as a result of the iodination level (see Figure 3.4 and Section 3.1.4.2). Finally, a distinction between $\alpha$-nucleosides and $\beta$-nucleosides in this mixtures was possible based on the specific signal patterns of their C1′-protons, as shown in Figure 3.15.

Several additional reaction conditions for direct iodination of 3.1 were explored, as shown in Entries 7-11 of Table 3.2. All of which produced unfavorable results, which prompted us to abstain from a detailed analysis of iodination levels. The use of ICl without Lewis acid at room temperature for 1 h in dichloromethane proved to be detrimental to the nucleoside framework causing approximately 85 % deglycosylation and 15 % anomerization (Entry 7, Table 3.2).
<table>
<thead>
<tr>
<th>Entry</th>
<th>Temp</th>
<th>Reagents (Eq)</th>
<th>Time</th>
<th>% base&lt;sup&gt;a&lt;/sup&gt;</th>
<th>% α&lt;sup&gt;b&lt;/sup&gt;</th>
<th>% β&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Estimated ratio of all iodinated and non-iodinated α- and β-forms of 3.1. &lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mono</td>
<td>Di</td>
<td>Tri</td>
<td>Tetra</td>
</tr>
<tr>
<td>1</td>
<td>0º</td>
<td>ICl (1) AlCl&lt;sub&gt;3&lt;/sub&gt;(10)</td>
<td>2</td>
<td>10</td>
<td>0</td>
<td>90</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>25º</td>
<td>I&lt;sub&gt;2&lt;/sub&gt; (1) AlCl&lt;sub&gt;3&lt;/sub&gt;(10)</td>
<td>2</td>
<td>15</td>
<td>30</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>3</td>
<td>40º</td>
<td>I&lt;sub&gt;2&lt;/sub&gt; (1) AlCl&lt;sub&gt;3&lt;/sub&gt;(10)</td>
<td>2</td>
<td>20</td>
<td>24</td>
<td>56</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>40º</td>
<td>ICl (1) AlCl&lt;sub&gt;3&lt;/sub&gt;(10)</td>
<td>24</td>
<td>30</td>
<td>39</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>40º</td>
<td>ICl (5) AlCl&lt;sub&gt;3&lt;/sub&gt;(10)</td>
<td>48</td>
<td>30</td>
<td>2</td>
<td>68</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>40º</td>
<td>I&lt;sub&gt;2&lt;/sub&gt; (5) AlCl&lt;sub&gt;3&lt;/sub&gt;(10)</td>
<td>48</td>
<td>60</td>
<td>36</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>25º</td>
<td>ICl (1)</td>
<td>1</td>
<td>85</td>
<td>15</td>
<td>0</td>
<td>ND</td>
</tr>
<tr>
<td>8</td>
<td>25º</td>
<td>ICl (1) K&lt;sub&gt;2&lt;/sub&gt;CO&lt;sub&gt;3&lt;/sub&gt; (1.5)</td>
<td>1</td>
<td>70</td>
<td>27</td>
<td>3</td>
<td>ND</td>
</tr>
<tr>
<td>9&lt;sup&gt;e&lt;/sup&gt;</td>
<td>25º</td>
<td>ICl (1) Et&lt;sub&gt;3&lt;/sub&gt;N (1.5)</td>
<td>1</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>10</td>
<td>85º</td>
<td>I&lt;sub&gt;2&lt;/sub&gt; (1) CAN (0.5)</td>
<td>2</td>
<td>70</td>
<td>25</td>
<td>5</td>
<td>ND</td>
</tr>
<tr>
<td>11</td>
<td>85º</td>
<td>I&lt;sub&gt;2&lt;/sub&gt; (1) CAN (0.5)</td>
<td>24</td>
<td>90</td>
<td>10</td>
<td>0</td>
<td>ND</td>
</tr>
</tbody>
</table>

**Table 3.2:** Effect of various reaction conditions on the level of iodination of 3.1

<sup>a</sup> % Base represents the percent amount of deglycosylated product (3.5) that was obtained. The quantitation of % base is primarily based on TLC analysis.

<sup>b</sup> % α represents the percent amount of anomerized product (3.4) that was obtained. The quantitation of % α is based on the integration of the NMR signals for the C1΄ proton.

<sup>c</sup> % β represents the percent amount of all isolated β-nucleosides. The quantitation of % β is based on integration of NMR signals for C1΄ proton.

<sup>d</sup> The % level of iodination of 3.1 was based on a combined evaluation of <sup>1</sup>H-NMR- and mass spectra.

<sup>e</sup> Reaction led to to *nido-o*-carborane formation.

ND: Not determined
We hypothesized that the formation of minor amounts of HCl, as it can be found e.g. as an impurity in dichloromethane, in the reaction mixtures may have been responsible for nucleoside deglycosylation/anomerization. This hypothesis proved to be invalid since the presence of the inorganic base K$_2$CO$_3$ in the reaction mixture did not significantly affect the outcome of the reaction (Entry 8, Table 3.2). Therefore, it seems most plausible that the highly reactive I$^+$, generated either from ICl or from I$_2$, and not H$^+$ was responsible for the deglycosylation/anomerization of the nucleosides. Quenching the acidity of the reaction medium with the stronger organic base, triethylamine (Entry 9, Table 3.2) apparently caused the formation of negatively charged nido-carboranyl compounds. This was demonstrated by TLC analysis, which revealed a new intensively dark spot with significantly smaller R$_f$ value than those of closo-carboranyl nucleobases and nucleosides.

The results shown in Table 3.2, Entries 1-8, related to the Scheme 3.4 indicated that nucleoside deglycosylation/anomerization in the presence of AlCl$_3$ is less extensive than in its absence. The exact role of AlCl$_3$ in reducing deglycosylation/anomerization is not known. However, ICl supposedly forms a complex (I$^+$ AL$^{+}$Cl$_4^-$) with AlCl$_3$. The attack of this complex at the glycoside bond may be sterically disfavored compared with a naked I$^+$ produced in situ from ICl alone. As will be discussed in Section 3.2.4.1, however, AlCl$_3$ alone in the absence of ICl or I$_2$ also causes substantial deglycosylation and anomerization of N3-substituted pyrimidine nucleosides.

Ceric ammonium nitrate (CAN) was successfully used as an in situ oxidant to generate electrophilic iodine species (I$^+$) from I$_2$ for the iodination of variety of compounds, including uracil/uridine derivatives. Therefore, we explored the effect
of the in situ generation of I$^+$ using I$_2$/CAN in acetonitrile on the iodination of 3.1 (Entries 10-11, Table 3.2). Unfortunately, both reactions produced significant deglycosylations and also some anomerization.

In summary, the ICl/AlCl$_3$ combination was optimal for the iodination of 3.1. The degree of iodination could be controlled to some extent by adjusting stoichiometry, temperature, and reaction time appropriately. However, even under optimized reaction conditions (Entry 1, Table 3.2), deglycosylation, diiodination of the carborane cluster, and substantial quantities of unreacted 3.1 generated a complex reaction mixture that necessitated the purification of 3.2 by semi-preparative HPLC. This is a major disadvantage of the direct iodination of 3.1. Although not reported here, the direct iodination of two other 3CTAs (N5-2OH and N3$^{20}$) produced similar results as those obtained for N5 (3.1).

AlCl$_3$ seems to play a pivotal role in this reaction, rendering the carborane cage more accessible to an attack by I$^+$ while preventing the attack of I$^+$ at the glycosidic bond to a some degree. To the best of our knowledge, the attack by I$^+$ on the glycoside bond of pyrimidine nucleosides has not been reported previously. Therefore, we decided to study this phenomenon in more detail at the example of the model compound N3-butylthymidine, as will be discussed in Section 3.2. The rationale for choosing N3-butylthymidine was that the produced nucleobase and the anomerized nucleoside would be sufficiently lipophilic to be purified by normal silica gel column chromatography for a detailed subsequent characterization. This would not be the case for thymine and dThd, which are insoluble in almost all solvents and very hydrophilic, respectively, and therefore difficult to analyze by TLC.
3.1.3.2. Iodination of unsubstituted o-carborane, followed by the attachment of the iodinated o-carborane to the dThd scaffold (Strategy 2)

In 1985, Andrews et al. reported a simple and efficient synthesis for 9-iodo-o-carborane (3.6) [see Schemes 3.2 and 3.5] using o-carborane and I₂ as starting materials in the presence of a catalytic amount of AlCl₃ under reflux condition using dichloromethane or carbon tetrachloride as the solvent. An attempt to explore this strategy for the direct iodination of 3.1 resulted in a mixture of products with varying levels of iodination (see Section 3.1.3.1). In contrast, the reaction of o-carborane with ICl in the absence of any other Lewis acid under reflux condition in dichloromethane furnished compound 3.6 in 70 % yield without producing significant amounts of products with higher level of iodination, even when an excess of ICl was used (Scheme 3.6). To our knowledge, this finding has not been reported previously. The spectral data obtained for 9-iodo-o-carborane synthesized with ICl were in accordance with those previously reported for 9-iodo-o-carborane. It should be noted that in the absence of Lewis acid, I₂ did not cause any iodination of o-carborane (data not shown), whereas both ICl and I₂ were far more reactive in the presence AlCl₃ causing even tetraiodination of 3.1, as discussed in Section 3.1.3.1.

Scheme 3.5 shows the general synthetic strategy for the synthesis of N5-I (3.2 a / 3.2 b) starting from 3.6. Both CH groups of 3.6 are susceptible to lithiation by n-BuLi. Therefore, one of these groups may be protected (R) to prevent disubstitution during reaction step 1. The terminal groups (X) of the linker employed in step 2 could be standard leaving groups. However, one X could also be a protective group (e.g. -OTBDPS or -OTBDMS) to prevent disubstitution at the linker with two carboranyl
moieties. Reaction step 2 is a classical N3-alkylation of dThd\textsuperscript{16,20,36,91,92,158-160} furnishing protected or deprotected N5-I (3.2 a / 3.2 b).

![Synthetic strategy for 3.2 a / 3.2 b (N5-I)](image)

**Scheme 3.5**: Synthetic strategy for 3.2 a / 3.2 b (N5-I)

A possible problem associated with this strategy could be the reaction of \( n\)-BuLi with the iodine of 3.6. Halogens attached to the boron in non-cluster molecules can undergo lithium-halogen exchange.\textsuperscript{184-186} To the best of our knowledge, however, such reactions have never been observed with B-halocarbonanes.\textsuperscript{140} Also, the nucleophilic displacement of the iodine atom of 3.6 (or any halogen attached to any type of carborane) by a nucleophile (e.g. \( \text{Bu}^- \) from \( \text{BuLi} \) or \( \text{F}^- \) from TBAF) may be unlikely because a nucleophile
approaching from the side opposite of the halogen atom will be sterically hindered by the icosahedral geometry of the cluster.\textsuperscript{140}

**Scheme 3.6** shows our initial approach for the synthesis **N5-I (3.2 a / 3.2 b)** according the general strategy shown in **Scheme 3.5**. Selective monoiiodination at B9 of \( o \)-carborane was accomplished with ICl in refluxing dichloromethane. The reaction of **3.6** with \( n \)-BuLi, and subsequently, with TBDMSI in THF under reflux conditions resulted in the formation of a mixture of the geometric isomers **3.7 a** and **3.7 b** with iodine either at B9 or at B12, in 66 % overall yield.\textsuperscript{187,188} Exclusive monosilylation of **3.6** occurs because a TBDMS group at one carbon atom of the cluster is sufficiently bulky to prevent silylation at the adjacent carbon.\textsuperscript{187,188} As anticipated, no nucleophilic displacement at the B-I bond was observed. \( R_f \) values of 0.47 and 0.36 (pentane: Et\(_2\)O, 25:1) were found for **3.7 a** and **3.7 b**, respectively, and it was possible to separate small quantities of this mixture by normal phase silica gel flash chromatography for X-ray crystallographic studies (see **Section 3.1.4.3** for details). Unfortunately, the column chromatographic separation of larger quantities of **3.7 a / 3.7 b** proved to be extremely tedious. The same was true for all other compound mixtures described in **Schemes 3.7, 3.8, 3.9, and 3.10** because of very close \( R_f \) values. Therefore, we abstained from the separation of isomers in the synthesis of **3.2 a / 3.2 b**.
Scheme 3.6: Synthesis of 3.2 a and 3.2 b using 1, 5-diodopentane

Reagents: a) ICl, DCM, 40°C, 5 h.; b) n-BuLi, TBDMSCl, THF, 66 °C, 12 h.; c) n-BuLi, 1,5-diodopentane, THF, rt, 24 h; d) dThd, K₂CO₃, DMF/acetone (1:1), 40°C, 3.5 h; e) TBAF, THF, -78°C to rt, 40 min.
As already discussed in Chapter 1, Section 1.6.1 and Sections 3.1.1, the carborane
cage of 3CTAs seems to be located outside of hTK1 and it was found that even
significant differences in cage substitution patterns\textsuperscript{91} did not alter 3CTA phosphorylation
by hTK1. Therefore, the regioselectivity of iodination (B9 vs B12) will most likely not
affect hTK1 substrate characteristics, as was confirmed in preliminary phosphoryl
transfer assays with N5, N5-I, and N5-Br (see Section 3.1.7).

Treatment of 3.7 a / 3.7 b with n-BuLi and 1, 5-diiodopentane in THF under reflux
afforded 3.8 a / 3.8 b in approximately 10 % yield following column chromatographic
purification. The presence of numerous unidentified side products impeded the
purification significantly. Attempts to improve the yield of 3.8 a / 3.8 b by carrying the
reaction out in diethyl ether: benzene (2:1)\textsuperscript{189} at room temperature failed. Purified 3.8 a / 3.8 b was reacted with dThd in presence of K\textsubscript{2}CO\textsubscript{3} in DMF: acetone (1:1) at 40 ºC to
furnish 3.9 a / 3.9 b in 61 % yield. However, the applied reaction conditions also led to
the partial removal of the TBDMS protective group furnishing approximately 20 % of 3.2
a / 3.2 b directly. Deprotection of remaining TBDMS protected 3.9 a / 3.9 b (isolated in
~ 52 % yield) was carried out with TBAF in THF affording a second batch of 3.2 a / 3.2
b in 45% yield.

Because of the low yield and difficult purification of 3.8 a / 3.8 b encountered during
the first synthetic method described above, the reaction of 3.7 a / 3.7 b with a
monoproected electrophile, 1-(tert-butyldiphenylsilyloxy)-5-pentyl tosylate,\textsuperscript{159,190} was
explored in a second approach, which is shown in Scheme 3.7. Reaction of 3.7 a / 3.7 b
with n-BuLi, and subsequently with 1-(tert-butyldiphenylsilyloxy)-5-pentyl tosylate, in
diethyl ether: benzene (2:1) under reflux afforded 3.10 a / 3.10 b in approx.70% yield.
Scheme 3.7: Synthesis and the deprotection of 3.10 a / 3.10 b.

Reagents: a) n-BuLi, 1-(tert-butyldiphenylsilyloxy)-5-pentyl tosylate, diethyl ether: benzene (2:1), 24 h. b) TBAF, THF, -78°C to rt, 3.5 h.

Theoretically, deprotection of both silyl protecting groups in 3.10 a / 3.10 b to produce 3.12 a / 3.12 b should be possible by using TBAF in THF.\textsuperscript{187,188,191} The hydroxyl function in 3.12 a / 3.12 b could be converted easily into a leaving group such as a tosylate, which then could be reacted with dThd in presence of K\textsubscript{2}CO\textsubscript{3} to furnish the 3.2 a / 3.2 b.
Scheme 3.8: Synthesis and deprotection of 3.14a / 3.14b

Reagents: a) n-BuLi, 5-(tert-butyldimethylsilyloxy)pentyl 4-methylbenzenesulfonate (3.13), 192-195, THF, 66°C, 12 h; b) TBAF, THF, -78°C to rt, 30 min, c) 10% methanolic HCl, rt, 30 min,

Unfortunately, the reaction of TBAF in THF for 30 min at 0 ºC and subsequently 3 h at room temperature did not completely remove the O-TBDPS group of 3.10a / 3.10b whereas the C-TBDMS group at the carborane was completely removed to produce a mixture of 3.11a / 3.11b and 3.12a / 3.12b (Scheme 3.7). Longer reaction times, higher temperatures and/or higher equivalents (2.2-2.5 eq) of TBAF were not explored.
because this could lead to the degradation of the closo-o-carborane cage to nido-o-carborane (see Chapter 2, Section 2.1).\textsuperscript{74,76}

Since cleavage of an O-TBDMS group with TBAF proceeds more readily than that of an O-TBDPS group,\textsuperscript{191} we carried out the synthesis of 3.14 a / 3.14 b, as shown in Scheme 3.8. 1-(tert-Butyldimethylsilyloxy)-5-pentyl tosylate (3.13) was initially synthesized from commercially available 1-(tert-butyldimethylsilyloxy)-5-pentanol in 75 % yield.\textsuperscript{192-195}

This was followed by the reaction of 3.7 a / 3.7 b with n-BuLi and 3.13 in refluxing THF overnight to afford 3.14 a / 3.14 b in 66% yield. As observed in the case of 3.10 a / 3.10 b, the reaction of 3.14 a / 3.14 b with TBAF in THF resulted only in the partial cleavage of the O-TBDMS in 3.14 a / 3.14 b, whereas the C-TBDMS group was removed completely. Acidic hydrolysis of silyl ethers has been reported to be more effective than cleavage with TBAF.\textsuperscript{191} Surprisingly, however, the reaction of 3.14 a / 3.14 b with 10% methanolic-HCl for 30 min at room temperature led to the complete removal of the O-TBDMS group whereas the C-TBDMS remained unaffected (3.16 a / 3.16 b, 55 % yield). Therefore, we decided to carry out the desilylation of 3.14 a / 3.14 b with TBAF/THF for 30 min at 0 °C followed by addition of 10% methanolic-HCl and continued stirring for 30 min at room temperature. These reaction conditions led to the complete removal of both silyl protective groups from both 3.10 a/ 3.10 b (Scheme 3.7) and 3.14 a / 3.14 b (Scheme 3.8) producing 3.12 a / 3.12 b. The yield for the conversion of 3.14 a / 3.14 b into 3.12 a / 3.12 b was 76 %. 

84
Scheme 3.9: Optimized synthesis of 3.2 a and 3.2 b.

Reagents: a) ICl, DCM, 40°C, 5 h.; b) n-BuLi, TBDMSCl, THF, 66°C, 12 h.; c) n-BuLi, 5-((tert-butyldimethylsilyloxy)pentyl 4-methylbenzenesulfonate; THF, 66°C, 12 h; d) TBAF, THF, -78°C to 4°C, 30 min; ii) 10% methanolic HCl, 4°C, 30 min; e) TsCl, Et₃N, DMAP, 0°C to rt, 7 h; f) dThd, K₂CO₃, DMF/acetone (1:1), 40°C, 2.5 h.
Finally, 3.12 a / 3.12 b, were converted into the corresponding tosylates, 3.17 a / 3.17 b, in 70 % yield (Scheme 3.9), which were then subjected to treatment with dThd in the presence of K₂CO₃ in DMF: acetone (1:1) to afford 3.2 a /3.2 b in 60 % yield after purification by column chromatography (see Scheme 3.9 for all optimized reaction steps).

3.1.4. NMR spectroscopic- and X-ray crystallographic analysis of various halogenated species described in Sections 3.1.3.

3.1.4.1. ¹H NMR analysis

The ¹H NMR spectrum of o-closo-carborane displays overlapping singlets for both carboranyl C-Hs (3.69 ppm in acetone-d₆) due to the overall symmetry of the cluster (see Figure 3.2, for the structure of o-carborane). However, monoiiodination of o-carborane at B9, as in 3.6 (Scheme 3.6), renders the carboranyl C-Hs either in meta- or in para-position to iodine. Therefore, the ¹H NMR of 3.6 displays two separate singlets at 4.74 ppm and 4.95 ppm in acetone-d₆, as shown in Figures 3.5 and 3.8 A.¹⁸² Similarly, the ¹H NMR spectrum of 3.1 shows a singlet for the single carboranyl C-H at 4.74 ppm (see Figure 3.3 A) whereas the ¹H NMR spectrum of 3.2 a / 3.2 b displays two singlets for this proton at 4.86 and 5.06 ppm (see Figures 3.3 B, and 3.9). Both signals are slightly downfield compared to that of the carboranyl C-H of 3.1. In contrast to 3.6, however, 3.2 is a mixture of geometric isomers (3.2 a / 3.2 b) as a result the unsymetric substitution at the carboranyl carbon atoms, as clearly shown in Scheme 3.6. It should be noted that monoiiodination of 3.1 both via strategy 1 and strategy 2 should result in a geometric mixture of 3.2.
Figure 3.3: The change in the $^1$H NMR chemical shifts of the carboranyl C-H of 3.1 (highlighted in red) as a result of various levels of iodination.

Diiodination of 3.1 (disubstitution) at B9 and B12 leads to formation of a single compound, which shows one singlet at 5.30 ppm for the single carboranyl C-Hs peak in its $^1$H NMR spectrum (see Figure 3.3 C). Analogously, tri- (B8, B9, B12) and tetraiodination (B8, B9, B10, B12) of 3.1 presumably resulted in the formation of single compounds with $^1$H NMR signals for their carboranyl C-Hs at 5.52 ppm and 5.74 ppm, respectively (see Figure 3.3 D and E). Similar results were obtained by Teixidor for different $\alpha$-carboranyl compounds with varying iodination levels. The $^1$H NMR resonances of the carboranyl C-Hs as a result of mono-, di-, tri, and tetra iodination
apparently indicate a gradual deshielding, which appears to be due to the electron withdrawing effect of the halogen substituent transmitted through the \(\sigma\)-bonds of the carborane cage.\(^{196,197}\) These distinct changes in chemical shifts of carboranyl C-Hs as a result of the halogenation level was used to monitor the reactions carried for the direct iodination 3.1, as described in Section 3.1.3.1.

3.1.4.2. \(^{11}\)B-NMR analysis.

The proton-coupled \(^{11}\)B-NMR of \(o\)-carborane shows two overlapping doublets for B9 and B12 at -2.52 ppm and also for B8 and B10 at -9.36 ppm (Figure 3.4 A). In 3.6 (Figure 3.4 B), the signal for iodine-substituted B9 has collapsed to a singlet because of the absence of a coupling partner (H) and it has also shifted significantly upfield (-16.52 ppm). The signal for B12 remains a doublet approximately at the same position as in \(o\)-carborane (-1.12 ppm). The proton-coupled \(^{11}\)B-NMR of 3.1 shows two doublets for B9 and B12 at -2.19 ppm and -5.46 ppm, respectively (Figure 3.4 D). This is due to the fact that these two boron atoms are either at meta- and para-position to the carbon atom of the carborane cluster that is substituted with the thymidylpentyl substituent. Both doublets are still visible in the proton-coupled \(^{11}\)B NMR spectrum of the isomeric mixture of 3.2 a / 3.2. b at -17.24 and -18.96 ppm (Figure 3.4 E). However, one isomer was iodinated at B9 and the other at B12. This is the reason why there are two singlets present at -2.20 and -5.45 ppm for B9 and B12, respectively.
Figure 3.4: Change in the chemical shifts in the proton coupled $^{11}$B NMR spectra of $\sigma$-closo-carborane and N5 (3.1) as a result of iodination.
The proton-coupled $^{11}$B-NMR spectrum of \textbf{3.3} displays a single peak (four overlapping singlets) at -17.84 ppm (\textbf{Figure 3.4 F}), which corresponds to the iodinated boron atoms 8, 9, 10, and 12, whereas the original signals for these boron atoms in \textit{o}-carborane are missing.

The changes in the chemical shifts observed in $^{11}$B NMR as a result of a halogen substitution at carborane clusters are of a more complex nature than those in $^1$H NMR, as has been reviewed extensively by Stanislav Heřmánek.\textsuperscript{198} Briefly, these changes are related to the nature of the substituent, i.e inductive electron donating or withdrawing effects (+I or –I), resonance effects (+R), conjugative diamagnetic susceptibility, and anisotropic effects. Heřmánek \textit{et al} stated that for heavy atoms, such as iodine, the magnetic anisotropy plays an important role because the B-I bond can generate a secondary magnetic field in presence of an externally applied magnetic field. This secondary magnetic field opposes the external magnetic field leading to increased shielding of the boron atom. Shielding intensifies with increasing orbital radius of the halogen (F< Cl< Br <<I) and is very pronounced for iodine.

The magnetic anisotropy induced change in $^{11}$B NMR chemical shift patterns of halogenated carboranes proved to be useful in the NMR spectroscopic analysis of halogen exchange reactions, which will be discussed in detail in the following \textbf{Section 3.1.5}. This phenomenon is briefly explained at the example of the proton-decoupled $^{11}$B NMR spectra of 9-iodo-$\textit{o}$-carborane (3.6) and 9-bromo-$\textit{o}$-carborane (3.18, see \textbf{Scheme 3.11} for synthesis) in \textbf{Figure 3.5}. 

\textit{\footnotesize\textsuperscript{90}}
Figure 3.5: Change in the $^{11}$B NMR (proton-decoupled) and $^1$H NMR chemical shifts of $o$-carborane as a result of substitutions with either iodine or bromine. The $^1$H NMR spectrum shows only the signal for the carboranyl C-Hs.

The signal for the heavily shielded iodinated boron of 9-iodo-$o$-carborane (3.6) appears upfield to the hydrogen substituted borons at -16.53 ppm whereas the corresponding brominated boron in 3.18 is deshielded and appears slightly downfield (-0.28 ppm) of the hydrogen substituted boron atoms. Also, in contrast to the $^1$H NMR spectrum of 3.6, which shows two separate singlets for the two carboranyl C-Hs of the carborane cluster, the $^1$H NMR spectrum of the 3.18 showed a single signal for both hydrogens. Apparently the two singlets for carboranyl C-Hs merged into a single signal despite different distances to the bromine substituent of the cluster. Paradoxically, this signal is positioned slightly upfield compared to those of 3.6.
3.1.4.3. X-ray crystallographic analysis.

Figure 3.6: A) Carboranyl C-H signals in the $^1$H NMR spectra of purified 3.7 a, purified 3.7 b, and the mixture 3.7 a and 3.7 b. B) X-ray crystal structure of 3.7 a.

The formation of geometrical isomers upon substitution of one of the carboranyl carbon atoms of 3.6 was also explored at the example of 3.7 a / 3.7 b, in which case a separation was possible (see Section 3.1.3.2) Initial crystallization attempts were only successful in the case of 3.7 a ($R_f 0.47$) but not for 3.7 b ($R_f 0.36$). The x-ray structure of 3.7 a is shown in Figure 3.6.A. It identified this compound as the isomer with the iodine and the silyl group in meta-orientation. The $^1$H NMR for this compound shows a single signal for the carboranyl C-H of the cluster at 4.76 ppm (Compound 3.7 a, Figure 3.6 A)
whereas the $^1$H-NMR for the corresponding “para-isomer” displays this signal at 4.55 ppm (Compound 3.7 b, Figure 3.6 A).

3.1.5. Halogen exchange reaction

As already mentioned in Section 3.1.1, iodine attached to a boron atom of a carborane cluster has proven to be very stable. The sensitivity of the B-halocarboranes toward nucleophilic attack, however, can be enhanced by converting them to the corresponding carboranyl (phenyl) iodonium salts (B-I activation). Also, several palladium catalyzed reactions were reported in which the B-halogen bond of B-halocarboranes was replaced with B-C bonds (alkylation / alkynylation / arylation / heteroarylation), B-O bonds (etheration), B-NH$_2$ bonds (amination), B-NHCOR bonds (amidation), and B-X bonds (halogen exchange/isotope exchange). To the best of our knowledge, however, the halogen exchange at a B-halocarborane has never been explored in the presence of a biomolecule that is structurally as complex as dThd.

Therefore, we explored and optimized the reaction conditions for $^{127}$I-$^{125}$I exchange at 3.2 a / 3.2 b successively using initially C-unsubstituted 3.6, then C-pentyl substituted o-carborane (3.20 a / 3.20 b), and finally 3.2 a / 3.2 b (Scheme 3.11). Cold NaBr was used as the halogen source instead of Na$^{125}$I during this optimization phase.
Scheme 3.10: Synthesis of 3.20 a / 3.20 b

Reagents: a) n-BuLi, 1-iodopentane, THF, 66°C, 18 h; b) TBAF, THF, -78°C to 0°C, 30 min.

Compounds 3.20 a / 3.20 b were synthesized in 66% yield by reacting 3.7 a / 3.7 b with n-BuLi, and subsequently, iodopentane followed by removal of the TBDMS protective group with TBAF (Scheme 3.10). Compound 3.6 and 3.20 a / 3.20 b were subjected to halogen exchange using 10 equivalents of NaBr in presence of various quantities of Hermann’s catalyst (10-50 mol %) in DMF at 110 °C for 1 h (Scheme 3.11). Reaction progress was monitored by $^1$H NMR spectroscopy and the brominated products were characterized by $^1$H NMR, $^{11}$B NMR, and mass spectrometry (see Figure 3.7 and 3.8).
Scheme 3.11: Bromine-iodine exchange reaction in presence of Hermann’s catalyst.

Reagents: a) NaBr, Hermann’s catalyst, DMF, 110 °C, 1 h.
**Figure 3.7**: Change in the \(^1\text{H} \text{NMR} \) \(^{11}\text{B} \text{NMR} \) (decoupled) chemical shifts of 1-\(\text{o}\)-carboranyl-pentane as a result of substitutions with either iodine or bromine. The \(^1\text{H} \text{NMR} \) spectrum shows only the signal for the carboranyl C-Hs.

In order to simulate and explore the interference of the nucleoside scaffold on the outcome of the halogen exchange, the halogen exchange reactions of \(3.6\) and \(3.20\ a / 3.20\ b\) were repeated with several modifications. The reaction of \(3.6\) with NaBr was carried out in presence of 1 equivalent of N3-butylthymidine (\(3.24\)) (see **Section 3.2**) in DMF at 110 °C for 1 h and the reaction of \(3.20\ a / 3.20\ b\) with NaBr was carried out using butanol instead of DMF as a solvent. The \(^1\text{H} \text{NMR} \) analyses of both reactions indicated either complete (\(3.6\)) or ~ 50% (\(3.20\ a / 3.20\ b\)) halogen exchange (see **Figure 3.8**).
Figure 3.8: $^1$H NMR spectroscopy as a tool to monitor the halogen exchange reactions. Both examples are from reactions in which the halogen exchange was only about 50% complete. The change in the chemical shifts of the carboranyl C-Hs was followed to monitor the course of the reaction. A) The reaction mixture was refluxed for 4 h in presence of 5 mol % Hermann’s catalyst in DMF. B) The reaction was carried out in butanol in presence of 10 mol % Hermann’s catalyst for 25 min.

Encouraged by the successful test reactions with the surrogates 3.6 and 3.20 a / 3.20 b, we proceeded with the reaction of 3.2 a / 3.2 b using 10 equivalents of NaBr in presence of Hermann’s catalyst (25 mol %) in DMF at 110 ºC for 1 h (see Scheme 3.11). This resulted in 100 % exchange to afford N5-Br (3.22 a / 3.22 b) in 52 % yield following column chromatographic purification, as indicated by $^1$H NMR, $^{11}$B NMR, and mass spectrometry. Boron-11- and $^1$H-NMR spectral data of 3.2 a / 3.2 b and 3.22 a / 3.22 b are shown in Figure 3.9.
<table>
<thead>
<tr>
<th>Compounds</th>
<th>$^{11}$B NMR</th>
<th>$^1$H NMR</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Compound Image" /></td>
<td><img src="image2" alt="B9-I/B12-I" /></td>
<td><img src="image3" alt="H" /></td>
</tr>
<tr>
<td><img src="image4" alt="Compound Image" /></td>
<td><img src="image5" alt="B9-Br/B12-Br" /></td>
<td><img src="image6" alt="H" /></td>
</tr>
</tbody>
</table>

Figure 3.9: The $^{11}$B (proton-decoupled) and $^1$H NMR spectra of 3.2 a / 3.2 b, and 3.22 a / 3.22 b as a result of substitutions with either iodine or bromine. The $^1$H NMR spectrum shows only the signal for the carboranyl C-Hs.
3.1.6. Isotope exchange reaction of N5-I (3.2 a / 3.2 b) with Na\textsuperscript{125}I

Commercial Na\textsuperscript{125}I (Perkin Elmer) is available in form of an aqueous alkaline solution (pH 8-11). Strong bases such as methoxide ion, ammonia, alkylamines, and piperidine are known to degrade closo-o-carboranes into the corresponding nido-carboranes.\textsuperscript{74,76}

\begin{align*}
\text{HO} & \quad \text{OH} \\
\text{O} & \quad \text{O} \\
\text{HO} & \quad \text{OH} \\
\text{N} & \quad \text{N} \\
\text{O} & \quad \text{O} \\
\text{H} & \quad \text{I}
\end{align*}

3.2 a / 3.2 b

\begin{align*}
\text{HO} & \quad \text{OH} \\
\text{O} & \quad \text{O} \\
\text{HO} & \quad \text{OH} \\
\text{N} & \quad \text{N} \\
\text{O} & \quad \text{O} \\
\text{H} & \quad \text{I}
\end{align*}

3.23 a / 3.23 b

\textbf{Scheme 3.12}: Radiohalogen exchange reaction of 3.2 a / 3.2 b with Na\textsuperscript{125}I to obtain 3.23 a / 3.23 b

In order to measure the possible extent of degradation of 3.2 a / 3.2 b during halogen exchange reaction using an alkaline solution of Na\textsuperscript{125}I, a test reaction was carried out with 3.2 a / 3.2 b using 0.5 eq NaBr and 5 mol % of NaOH in DMF applying the conditions indicated in Scheme 3.11. Fortunately, \textsuperscript{1}H NMR and \textsuperscript{11}B NMR spectroscopy did not indicate significant nido-production during this pilot reaction.
Figure 3.10: Radio-TLC analysis of the isotope exchange reaction of 3.2 a / 3.2 b with Na$^{125}$I in presence of Hermann’s catalyst: A) Radio-TLC of the crude reaction mixture indicating that three $^{125}$I-containing species are present. These are Na$^{125}$I, located near the origin of the TLC plate, N5-$^{125}$I (3.23.a / 3.23.b), which is the expected product of the reaction, and Hermann’s catalyst near the solvent front. Since the halogen/isotope exchange reaction is carried out sequentially starting with oxidative addition of the halogenating species at the Pd atom, Hermann’s catalyst may also be labeled with $^{125}$I and becomes visible on radio TLC plate. B) Radio TLC of N5-$^{125}$I (3.23.a / 3.23.b), purified by preparative TLC.
Finally, the isotope exchange reaction of \(3.2\ a / 3.2\ b\) was carried out with Na\(^{125}\)I in presence of 10 mol % of Hermann’s catalyst in DMF at 110 °C for 1 h. The reaction was monitored by radioTLC analysis (Figure 3.10) and the product (N5\(^{125}\)I, 3.23 a / 3.23 b) was purified by semi-preparative HPLC (RP-18, solvent system: acetonitrile: water, 1:1, isocratic elution) Both, the analytical UV chromatogram and radioanalytical trace of 3.23 a / 3.23 b displayed two peaks (RP18 analytical HPLC retention time = 26.5 min, acetonitrile: water, 1:1, isocratic elution), which is consistent with the presence of two N5\(^{125}\)I isomers (see Figure. 3.11)

\[\text{Figure 3.11: HPLC analysis for isotope exchange reaction of N5\(^{127}\)I (3.2 a / 3.2 b) with Na\(^{125}\)I to yield N5\(^{125}\)I (3.23 a / 3.23 b) [RP18 analytical HPLC retention time = 26.5 min, acetonitrile: water, 1:1, isocratic elution].}\]
3.1.7. Summary and future directions

To the best of our knowledge, this is the first report of halogen exchange at a B-halocarborane, which is conjugated to a biomolecule such as dThd. Future efforts will focus on the scale up of N5-125I and the exact determination of its radiochemical yield and specific activity. Preliminary phosphoryl transfer assays with N5 (3.1) N5-I (3.2 a / 3.2 b) and N5-Br (3.22.a / 3.22 b) resulted in comparable phosphorylation rates for all compounds (see Chapter 5, Section 5.5.2, for detailed description of experimental procedure). The results of these initial comparative enzyme assays will be substantiated with more detailed studies to validate the hypothesis that the effect of halogen substitution on the phosphorylation of N5 (3.1) by hTK1 is neglectable. The biodistribution profile of N5-125I (3.23 a / 3.23 b) will be evaluated in rats bearing intracerebral F98 or RG2 glioma. 41

In general, the synthetic technology developed at the examples of N5-Br and N5-125I may pave the way for synthesis of a wide range of carborane cage radiohalogenated therapeutics and diagnostics.
3.2. A General Study of the deglycosylation and anomerization of N3-substituted thymidine and uridine derivatives caused by electrophilic iodine (I⁺)

3.2.1. Introduction

As discussed in detail in, Section 3.1.3.1 the major limitations of a direct iodination of 3.1 was found to be a considerable deglycosylation and anomerization combined with difficult-to-control iodination levels. Since these finding have not been reported previously, we decided to explore the effect of iodinating reagents such as ICl and I₂ on several protected and unprotected non-carboranyl model thymidine- and uridine analogues in more detail. The products of such reactions, α-nucleosides and (N3-substituted) pyrimidine bases, may have important applications, as will be discussed in following Sections 3.2.2 and 3.2.3.

3.2.2. Importance / Significance of α-nucleosides, α-nucleotides and α-oligonucleotides in medicinal chemistry

α-Nucleosides were used as building blocks for α-oligodeoxynucleotides for hybridization with complementary natural DNA or RNA as antisense molecules. In contrast to their β-oligodeoxynucleotides counterparts, however, they were resistant towards nucleases.221-229 Also, duplexes of α-oligodeoxynucleotides with RNA were not substrates for RNase H.227,229-234 Analogues of α-oligodeoxynucleotides formed stable duplexes with ssRNAs and stable triplexes with dsDNAs.233 Recently, α-nucleosides were found to be selective inhibitors of thymidylate kinase from Mycobacterium tuberculosis.235 Importantly, these analogues did not inhibit human thymidylate kinase
even at concentrations as high as 1 mM. So far, methods for the synthesis of \( \alpha \)-deoxynucleosides have primarily focused on the epimerization of commercially available \( \beta \)-deoxynucleosides. 

3.2.3. Importance / Significance of N3-substituted thymine nucleobases and nucleosides in medicinal chemistry

N3-substituted dThd analogues have attracted considerable attention in recent years (see Chapter 1, Section 1.6.1 and Chapter 3, Section 3.1.1). N3-substitution can be exploited for the attachment of therapeutic or diagnostic entities. The syntheses of such dThd analogues generally starts with dThd, which is subjected to N3-alkylation with alkyl halides or tosylates in the presence of weak bases, such as \( \text{K}_2\text{CO}_3 \), or Mitsunobu reaction with various alcohols in the presence of e.g. PPh3/DEAD (Figure 3.12, Strategy 1). The advantage of both methodologies is that the reaction conditions are simple, efficient, and stereospecific, resulting only in the formation of N3-substituted \( \beta \)-anomers.
Figure 3.12: Strategies for the synthesis of various 3-substituted thymine/uracil based nucleosides.

Alternatively, natural/unnatural pyrimidine bases that are not N3-substituted could be condensed with protected natural/unnatural carbohydrates in presence of a Lewis acid. After removal of protective groups, such pyrimidine nucleosides could be subjected to N3-alkylation (Figure 3.11, Strategy 2.1). A third strategy is based on the condensation of N3-substituted pyrimidine bases with protected carbohydrates, which would furnish the desired N3-substituted nucleosides following deprotection (Figure 3.11, Strategy 2.2). The advantage of the latter two strategies is that a wide variety of natural or
unnatural pyrimidine bases can be condensed with natural or unnatural carbohydrates to furnish the desired N3-substituted nucleosides (see also Chapter 4). To the best of our knowledge, Strategy 2.2 has not been explored widely, possibly because methods for the synthesis of N3-substituted pyrimidine bases are not straightforward.

3.2.4. Result and Discussion

3.2.4.1. Investigation of factors impacting deglycosylation and anomerization of β-pyrimidine nucleosides by iodine species.

Scheme 3.13. The synthesis of substituted 3-butyl thymidine and its reaction with I₂.

Reagents: a) Butyl bromide, K₂CO₃, DMF: Acetone (1:1), 50 ºC, overnight; b) I₂, dichloromethane, 40 ºC, 96 h.
The synthesis of 3-butyldrimidine (3.24, Scheme 3.13) is more convenient and less expensive than that of N5 (3.1). The N3-butyl group in 3.24 renders, the corresponding nucleobase (3-butyldrimine, 3.25), and anomerized 3-butyldrimidine (α-anomer, 3.26) still sufficiently lipophilic for purification by normal silica gel column chromatography, which is necessary for proper characterization. Corresponding analogues with smaller N3-substituent or no N3-substituent at all may have to be purified by reversed phase chromatography because of an increased hydrophilic character, which is more tedious and expensive. Therefore, we have chosen 3.24 as surrogate for N5 (3.1) for a continued study of the mechanism of the deglycosylation and anomerization of N3-substituted pyrimidine nucleosides by iodinating agents, such ICl and I2. Another objective was to vary conditions for the reaction shown in Scheme 3.13 in order to 1) suppress deglycosylation and anomerization, 2) maximize the yield of N3-substituted nucleobase, and 3) maximize the yield of α-N3-substituted pyrimidine nucleoside analogue.

We subjected 3.24 to a variety of reactions conditions, as shown in Table 3.3, which can be considered as an extension of Table 3.2 (Section 3.1.3.1). Overall, the action of both I2 and ICl caused anomerization and deglycosylation irrespective of the nature of the solvent (dichloromethane vs acetonitrile). ICl alone caused significant deglycosylation and anomerization at room temperature (Table 3.3, Entries 1, 2, and 8), whereas I2 was less reactive at this temperature (Table 3.3, Entries 3, 5, and 9). On the other hand, deglycosylation with I2 accelerated at elevated temperatures apparently at the expense of anomerization (Table 3.3, Entry 5 vs. 6 and Entry 9 vs. 10).

Based on the reaction conditions described Entries 1-3 and 5-9, three culprits theoretically could be responsible for the observed deglycosylation and anomerization: 1)
Acidic residues in the reaction medium (H\(^{+}\)), 2) electrophilic iodine (I\(^{+}\)), or 3) nucleophilic iodine (I\(^{-}\)). The possible roles of these factors will be discussed in the following.

1) Acidity of the reaction medium.

Nucleosides are known to undergo deglycosylation under acidic conditions.\(^{253,254}\) The material safety data sheet (MSDS) of ICl (Mallinckrodt Baker Inc.) indicates that it decomposes to yield hydrogen chloride (HCl) and hydrogen iodide (HI) when exposed to heat, moisture, water or air. I\(_2\) is known to generate HI during iodination reactions, which increases the acidity of reaction media that are not completely anhydrous.\(^{255}\) Reaction mixtures were therefore basified with inorganic bases in order to quench any residual acidity (Table 3.3, Entries 4 and 12). The presence of potassium carbonate had no significant impact on the activity of ICl (Table 3.3, Entry 4). Similar results were found when N\(_5\) (3.1.) was reacted with ICl in the presence of potassium carbonate (Table 3.2, Section 3.1.3.1). The reactivity of I\(_2\) in presence of sodium bicarbonate, however, appeared to be significantly reduced (Table 3.3, Entry 12). Therefore, it appears that ICl is much more effective than I\(_2\) in causing deglycosylation and anomerization even under non-acidic conditions. On the other hand, we cannot exclude the possibility that in the case of I\(_2\) acidic species are contributing factors to deglycosylation/anomerization.

When the reactions were carried out with toluene sulfonic acid (Table 3.3, Entry 15) and AlCl\(_3\) (Table 3.3, Entries 13 and 14), as representatives of potent Brønsted- and Lewis acids, respectively, in the absence of any iodine species, significant deglycosylation and anomerization was observed. In the case of AlCl\(_3\), deglycosylation
may be initiated by a nucleophilic attack of the free electron pair of N1 at aluminum followed by opening of the glycoside bond.\textsuperscript{256}

<table>
<thead>
<tr>
<th>Solv.</th>
<th>Entry</th>
<th>Temp.</th>
<th>Reagents (Eq.)</th>
<th>Time (h)</th>
<th>% Outcome</th>
<th>Deglycosylation</th>
<th>Anomerization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>25º</td>
<td>ICl (1)</td>
<td>40 min</td>
<td></td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>DCM</td>
<td>2</td>
<td>25º</td>
<td>ICl (1)</td>
<td>18</td>
<td></td>
<td>75</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>25º</td>
<td>ICl (1)</td>
<td>40 min</td>
<td>No rxn</td>
<td>No rxn</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>25º</td>
<td>ICl (1) K$_2$CO$_3$ (1)</td>
<td>12</td>
<td></td>
<td>60</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>25º</td>
<td>I$_2$ (1)</td>
<td>48</td>
<td></td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>6*</td>
<td>40º</td>
<td>I$_2$ (1)</td>
<td>96</td>
<td></td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>40º</td>
<td>KI (1.2)</td>
<td>12</td>
<td>No rxn</td>
<td>No rxn</td>
<td></td>
</tr>
<tr>
<td>MeCN</td>
<td>8</td>
<td>25º</td>
<td>ICl (1)</td>
<td>18</td>
<td></td>
<td>75</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>25º</td>
<td>I$_2$ (1)</td>
<td>19</td>
<td></td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>85º</td>
<td>I$_2$ (1)</td>
<td>24</td>
<td></td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>85º</td>
<td>NaI (1.2)</td>
<td>48</td>
<td>No rxn</td>
<td>No rxn</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>85º</td>
<td>I$_2$ (1.2) NaHCO$_3$(0.5)</td>
<td>24</td>
<td></td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>25º</td>
<td>AlCl$_3$ (1)</td>
<td>48</td>
<td></td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>85º</td>
<td>AlCl$_3$ (1)</td>
<td>24</td>
<td></td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>25º</td>
<td>TsOH (0.5)</td>
<td>12</td>
<td></td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>50 % THF</td>
<td>16</td>
<td>25º</td>
<td>Chloramine-T (1.5) NaI (1.5)</td>
<td>24</td>
<td>No rxn</td>
<td>No rxn</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.3:** Effect of various reaction conditions on the deglycosylation and anomerization of 3.24.

*Reaction conditions similar to those outlined in Entry 6 were used to produce 3.26 and 3.27 for characterization by NMR and MS (see Chapter 5). In general, the percentages given for deglycosylation and anomerization were estimated based on TLC patterns.

The rate of the reaction was found to be somewhat slower at for AlCl$_3$ (Table 3.3, Entry 13 vs Entry 15). However, it increased significantly at elevated temperature
Overall, the reactivities of toluene sulfonic acid and AlCl$_3$ seemed to be lower than that of ICl. It should be noted that the capacity of AlCl$_3$ alone to cause deglycosylation / anomerization seems to be comparable to that of AlCl$_3$ in the presence of ICl or I$_2$ (Table 3.2). This finding suggests that the exact species responsible for deglycosylation / anomerization in reactions that contain both ICl/I$_2$ and AlCl$_3$ remains the subject of exploration.

2) Nucleophilic iodine (I$^-$)

In order to explore the potential role of the negatively charged nucleophilic iodine species (I$^-$) on the deglycosylation of 3.24, two reactions were carried with KI (Table 3.3, Entry 7) and NaI (Table 3.3, Entry 11) alone in dichloromethane and acetonitrile, respectively, under reflux conditions. No deglycosylation or anomerization was observed.

3) Electrophilic iodine (I$^+$)

Besides the use of Lewis acids, such as aluminum chloride (Section 3.1.3.1), two other strategies can be used to generate I$^+$: 1) ICl appears to be an effective source of I$^+$. The fact that KI / NaI did not cause any deglycosylation/anomerization of 3.24 and that ICl was equally effective in deglycosylation/anomerization of 3.24 in the presence and absence of base, provides strong evidence that acidic species are not involved in this reaction and that I$^-$ must be the active species. 2) The generation of I$^+$ in situ using e.g. NaI and oxidizing agents such as Chloramine-T. Such conditions have been exploited for the iodination of varies nucleophilic arenes, alkenes, and alkynes.$^{255}$ In these cases, electrophilic iodine reacts with double bonds, which serve as the nucleophiles. However,
in situ production of I⁺ using Chloramine-T failed to cause deglycosylation/ anomerization of 3.24 (Table 3.3, Entry 16). A reason could be that 3.24 does not have a nucleophilic double bond, which may be a necessary requirement for iodination under such conditions.

3.2.4.2. Hypothetical mechanism of deglycosylation and anomerization of 3.24

![Diagram of hypothetical mechanism]

Figure 3.13: Hypothetical mechanism, shown at the example of the initial attack of I⁺ on C-2 carbonyl oxygen, for the deglycosylation and anomerization of pyrimidine nucleosides by ICl. R = any alkyl-type substituent.
A hypothetical mechanism for the action ICl on the glycoside bond of 3.24 is shown in Figure 3.13. Because of the presence of partial charges in ICl, the attack of the electrophilic I\(^+\) on the nucleophilic oxygen at the 2 position, seems possible (intermediate A, Figure 3.13). A series of resonance delocalization then leads to the release of a nucleobase, and 2′-deoxyribose in the form of an oxonium cation (deglycosylation). The released nucleobase could then re-attack at deoxyribose preferentially from the α-face leading to the formation of the α-nucleoside (anomerization). Re-attack from the α-face may be facilitated due to the presence of the 5′-hydroxymethyl group located on the β-face of deoxyribose.

According to the mechanism of the acidic hydrolysis of the glycoside bonds in the nucleosides,\(^{257}\) I\(^+\) could theoretically also attack the deoxyribofuranose ring oxygen or the C-4 carbonyl oxygen, which would subsequently produce the 3.25 and 3.26 via similar mechanisms.

3.2.4.3. The effects of various modifications at the ribose portion of 3.24 on the deglycosylation and anomerization by ICl

Methyl- and benzyloxycarbonyl (Cbz) groups were introduced at the 3′- and 5′ hydroxyls of 3.24 (Scheme 3.14) in order to study the effects of substituents on the deglycosylation/anomerization by ICl. Compound 3.27 (3′, 5′-dimethyl-3-butythymidine) was synthesized in 80 % yield by reaction of 3.24 with NaH in THF followed by the addition methyl iodide.\(^{258,259}\) Similarly, compound 3.28 (3′, 5′-di-O-(carbobenzyloxy)-3-butythymidine) was synthesized in 75 % yield using benzyl chloroformate in dichloromethane in the presence of DMAP.\(^{260}\)
**Scheme 3.14:** The synthesis of substituted N3-butyl thymidine

Reagents: MeI, NaH, THF, sonic bath activation, 0 ° to rt, 40 min; b) CbzCl, DMAP, dichloromethane, rt, 17 h.

Neither the reaction of 3.27 nor that of 3.28 with ICl in dichloromethane at room temperature produced any deglycosylation (Table 3.4, Entries 1 and 6). In the case of 3.27, no deglycosylation and anomerization was observed after 6 h reflux with ICl in dichloromethane (Table 3.4, Entry 3) whereas about 25% deglycosylation and 5-10% anomerization of was observed when the reflux time was extended to 48 h (Table 3.4, Entry 2). The level of deglycosylation/anomerization of 3.27 further increased with reflux in acetonitrile (Table 3.4, Entry 4).

The reaction of 3.28 with ICl under reflux in dichloromethane did not produce any deglycosylation (Table 3.4, Entry 7) whereas anomerization was observed at reflux- and at room temperature (Table 3.4, Entries 7 and 6). As already mentioned in Section 3.1.3.1, a distinction between α-nucleosides and β-nucleosides in the reaction mixtures listed in Table 3.2 was made based on the specific ¹H NMR signal pattern of their C1’ protons, as shown in Figure 3.13 at the example of 3.24 and 3.27. The same is true for all
reaction mixtures listed in Tables 3.3, 3.4, and 3.5 of Section 3.2.4.3. In the case of the reaction of 3.28 with ICl, a characteristic signal for a C-1’ proton of a α-nucleoside was indeed observed in the reaction mixture. However, in contrast to all other nucleosides discussed in Section 3.2, compound 3.28 contains two aryl groups at the 3’- and 5’-positions, which are potentially also susceptible to iodination by ICl. It is therefore conceivable that the observed signal stems from 3.28 with iodinated Cbz groups. Nevertheless, it is unlikely that iodination of these groups in 3.28 would significantly affect the 1H NMR chemical shift of its C1’ proton.

<table>
<thead>
<tr>
<th>Entry</th>
<th>R,R’</th>
<th>Solv</th>
<th>Temp</th>
<th>Reagents (Eq)</th>
<th>Time h</th>
<th>Deglycosylation</th>
<th>Anomerization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Me</td>
<td>DCM</td>
<td>25º</td>
<td>ICl (1)</td>
<td>24</td>
<td>No rxn</td>
<td>No rxn</td>
</tr>
<tr>
<td>2</td>
<td>Me</td>
<td>DCM</td>
<td>40º</td>
<td>ICl (1)</td>
<td>48</td>
<td>25</td>
<td>5-10</td>
</tr>
<tr>
<td>3</td>
<td>Me</td>
<td>DCM</td>
<td>40º</td>
<td>ICl (1)</td>
<td>6</td>
<td>No rxn</td>
<td>No rxn</td>
</tr>
<tr>
<td>4</td>
<td>Me</td>
<td>MeCN</td>
<td>85º</td>
<td>ICl (1)</td>
<td>30</td>
<td>33</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>Me</td>
<td>MeCN</td>
<td>85º</td>
<td>AlCl₃ (1)</td>
<td>12</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>Cbz</td>
<td>DCM</td>
<td>25º</td>
<td>ICl (1)</td>
<td>6</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>Cbz</td>
<td>DCM</td>
<td>40º</td>
<td>ICl (1)</td>
<td>12</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>Cbz</td>
<td>DCM</td>
<td>25º</td>
<td>AlCl₃ (10)</td>
<td>1</td>
<td>5-10</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3.4: Effect of various reaction conditions on the deglycosylation and anomerization of 3.27 and 3.28.

*The reactions were analyzed by TLC using 3.24 and 3.25 as references. The ratio between α- and β-nucleosides was estimated based on 1H NMR data (see text related to Table 3.4. for further discussion. The CBz groups were removed under these reaction conditions.

A comparison of the results presented in Tables 3.3 and Table 3.4 indicates that the introduction of Me- or Cbz-groups at the 3’- and 5’-hydroxyls decreased the
deglycosylation / anomerization of 3.24 significantly. The exact mechanism for this finding is unknown. Free hydroxyl groups at the 3′- and 5′-positions may contribute in some way to the stabilization of the intermediate oxonium cation (Figure 3.14 shows such stabilization by 2′ hydroxyl group) whereas substituted hydroxyl groups may have lost this capacity.

AlCl₃ alone did not seem to have a significant impact on the deglycosylation/anomerization of 3′, 5′-disubstituted 3.24 (Table 3.4, Entries 5 and 8). However, the reaction of 3.28 with AlCl₃ at room temperature in dichloromethane completely removed the Cbz groups within one hour of reaction (Table 3.4, Entry 8). The Cbz group has been used as a protective group for ribose hydroxyl functions in pyrimidine nucleoside synthesis and their removal has been accomplished by hydrogenolysis. To the best of our knowledge, this is the first example of deprotection of an O-Cbz-protected pyrimidine nucleoside by the action of AlCl₃.

<table>
<thead>
<tr>
<th>R′,R″</th>
<th>Solvent</th>
<th>Temp.</th>
<th>Reagents</th>
<th>Time</th>
<th>Outcome¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>DCM</td>
<td>25º</td>
<td>ICl (1)</td>
<td>72</td>
<td>Iodination</td>
</tr>
<tr>
<td></td>
<td>DCM</td>
<td>40º</td>
<td>ICl (1)</td>
<td>24</td>
<td>Iodination</td>
</tr>
<tr>
<td></td>
<td>MeCN</td>
<td>85º</td>
<td>TsOH (0.5)</td>
<td>60</td>
<td>No rxn</td>
</tr>
<tr>
<td></td>
<td>MeCN</td>
<td>85º</td>
<td>ICl (3)</td>
<td>48</td>
<td>Iodination</td>
</tr>
<tr>
<td></td>
<td>MeCN</td>
<td>25º</td>
<td>AlCl₃ (1)</td>
<td>24</td>
<td>No rxn</td>
</tr>
<tr>
<td></td>
<td>MeCN</td>
<td>85º</td>
<td>AlCl₃ (1)</td>
<td>24</td>
<td>No rxn</td>
</tr>
<tr>
<td>Me</td>
<td>DCM</td>
<td>25º</td>
<td>ICl (1)</td>
<td>20</td>
<td>Iodination</td>
</tr>
</tbody>
</table>

Table 3.5: The effect of various reaction conditions on compound 3.29 and 3.30.

¹Iodinated products were analyzed by ¹H NMR, ¹³C NMR and mass spectrometry.
In order to explore the effect of ICl on the deglycosylation / anomerization of ribonucleosides, N3-butyluridine (3.29) and its 2’’, 3’, 5’’-trimethyl derivative (3.30) were initially synthesized as shown in Scheme 3.15. According to the results of various reactions listed in Table 3.5, neither 3.29 nor 3.30 appeared to be susceptible to deglycosylation / anomerization by ICl, toluene sulfonic acid, or AlCl₃. However, the action of ICl resulted in the production of compounds 3.31 and 3.32, which are the 5-iodo analogues of 3.29 and 3.30, respectively.

\[
\begin{align*}
\text{HO} & \quad \text{HO} \\
\text{OH} & \quad \text{OH} \\
\text{OH} & \quad \text{OH} \\
\text{N} & \quad \text{N} \\
\text{O} & \quad \text{O} \\
\text{O} & \quad \text{O} \\
\text{N} & \quad \text{N} \\
\text{O} & \quad \text{O} \\
\text{OH} & \quad \text{OH} \\
\end{align*}
\]

Scheme 3.15: Synthesis of N3-substituted uridines and their reaction with ICl.

Reagents: a) Butyl bromide, K₂CO₃, DMF: acetone (1:1), 50 °C, overnight; b) Methyl iodide, NaH, THF, sonic activation; c) ICl, dichloromethane, 40 °C, 48 h.
A hypothetical mechanism for the higher stability of the glycosidic bonds in 3.29 and 3.30 compared with that of 3.24 against ICl is shown in Figure 3.14. According to this mechanism, 2’-hydroxy/alkoxy substituents in 3.29-3.30 stabilizes the intermediate B, after the attack of I+, which is then converted back to 3.29-3.30 via an intermediate C. However, in the case of deoxyribopyrimidines, such as 3.24, the ring oxygen of deoxyribose takes part in the stabilization of intermediate A, as shown in Figure 3.13, leading to deglycosylation/anomerization.

**Figure 3.14:** Hypothetical mechanism for the resistance of the glycoside bond in ribonucleosides towards cleavage by ICl.
3.2.5. Applications for the developed deglycosylation and anomerization strategies.

Pyrimidine nucleobases, such as thymine and uracil, possess two nucleophilic nitrogen atoms at the N1 and the N3 positions. Both nitrogens react with electrophiles approximately to the same extent and regioselective N-substitution is not straightforward. Methods for the selective protection of one of the nitrogens have been developed and these generally drive the course of the intended substitution towards the desired position. However, the introduction and removal of the protective groups are two additional steps in any synthetic sequence and this is disadvantageous.

Scheme 3.16: ICl mediated deglycosylation for the synthesis of N3-substituted thymine nucleobases.

Reagents: a) RX, K₂CO₃, DMF: acetone (1:1), 50 °C; b) ICl, acetonitrile, reflux.

The synthesis of N3-substituted thymine/uracil nucleobases, as shown in Scheme 3.16, via ICl-based deglycosylation of the N3-substituted deoxyribopyrimidine nucleosides that have been derived from commercially available deoxyribopyrimidine nucleosides may be advantageous compared with previous strategies, which employ the introduction N1-protective groups because the deoxyribose portion serves here as the N1-protective group.
As discussed in Section 3.2.4.3, the presence of 3’ and 5’-Cbz protective groups at 3.24 significantly decreased the deglycosylation whereas anomerization was not affected. Therefore, the reaction of Cbz protected β-pyrimidine deoxyribonucleosides with ICl may have the potential to be developed into a general strategy for the synthesis of the corresponding α-pyrimidine deoxynucleosides.

3.2.6. NMR analysis of compound 3.24 and 3.26

The assignment of the α-configuration of 3.26 was primarily based on the differences between $^1$H NMR chemical shifts and multiplicities of signals for the deoxyribofuranose protons of 3.26 and 3.24 (Figure 3.15).

![Figure 3.15: Difference in $^1$H NMR chemical shifts and splitting patterns for Cl’-H between 3.24 and 3.26](image)

In particular, the spectrum of 3.26 showed a doublet of doublet at 6.04 ppm for C1’-H whereas that of 3.24 showed an apparent triplet at 6.30 ppm. These signal patterns are very characteristic for α- and β-nucleosides and are due to the change in configuration at the anomeric carbon.236,240,266

In addition, 13C chemical shifts related to the deoxyribofuranose ring of 3.24 and 3.26 differed significantly. For compound 3.24, the carbon chemical shifts for C1’, C3’, C4’, and C5’ in CD3OD were observed at 87.12 ppm, 72.13 ppm, 88.89 ppm, and 62.80 ppm, respectively. The corresponding signals for 3.26 in CD3OD were observed at 79.75 ppm, 68.39 ppm, 67.86 ppm, and 67.38 ppm respectively.

The complete assignment of the carbon chemical shifts and the carbon-hydrogen correlations for 3.24 and 3.26 was based on a combined analysis of 1H NMR-, 13C NMR-, 13C DEPT-, and 2D HSQC spectra (see Appendix, Figure A1-A8 for detailed proton and carbon assignments).
CHAPTER 4

The synthesis of novel N3-substituted D- and L-dioxolanyl pyrimidine nucleoside analogues to utilize thymidine kinase as a molecular target for the treatment of Epstein-Barr virus (EBV) associated cancers

Contributors other than myself:

6) Elena Sjuvarsson (Ph.D, Research Associate, Department of Anatomy, Physiology, and Biochemistry, Swedish University of Agricultural Sciences, Uppsala)

7) Fengting Yan (Ph.D, Post-Doctoral Researcher, Comprehensive Cancer Center, The Ohio State University)

8) Hitesh K. Agarwal (Ph.D, Post-Doctoral Researcher, Division of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, The Ohio State University)

9) Naum Sayfullin (Student Research Assistant, BS Program, College of Pharmacy, The Ohio State University)
4.1. Introduction

The Epstein-Barr virus (EBV) is a member of the human herpesvirus family and infects approximately 90% of the world’s population. Oral transmission through saliva is the most common form of transmission leading to primary infection, which is often asymptomatic but can occasionally result in infectious mononucleosis (IM). The oropharyngeal epithelial cells are the initial target of infection and production of infectious virus in the host. Infectious virions shed from epithelial cells, gain access to submucosal lymphoid tissue where B-cells are then infected. It is within this B cell compartment where EBV maintains a persistent latent infection for the lifetime of the host (See Figure 4.1). Common EBV associated diseases can be benign or malignant. Benign diseases include IM and oral hairy leukoplakia (OHP) whereas malignant diseases encompass a broad spectrum of pathology including Burkitt’s lymphoma (BL), Hodgkin’s disease (HD), nasopharyngeal carcinoma (NPC), gastric carcinoma, X-linked lymphoproliferative syndrome (XLPS), and lymphoproliferative diseases. The association of EBV with these malignancies and their traditional treatment approaches including chemotherapy, surgery, and radiotherapy have been reviewed elsewhere.

4.2. EBV life cycle

As mentioned in Section 4.1, EBV is primarily transmitted to the oropharyngeal epithelial cells via saliva. Once inside the epithelial lining of the oropharynx, where primary lytic replication occurs, EBV soon gains access to the submucosal lymphoid tissue infecting both naïve and memory B-lymphocytes. Figure 4.1 a
illustrates the sequence of events following primary infection by EBV. The virus enters B-lymphocytes via the binding of the viral envelope glycoproteins gp350 and gp42 to the CD21 receptor and human leukocyte antigen (HLA) Class II molecule, respectively, on the surface of B cells. EBV establishes latent infection within the B-lymphocyte compartment where it will persist for the life of the infected individual. After infection, EBV has the potential to express latent gene product (latency III program) that can differentiate naïve B-cell into activated B-cell blasts and finally memory B-cells (See Tables 4.1 and 4.2). Fortunately, the rapidly proliferating B-cell blasts express several highly immunogenic latency III gene products that are effectively recognized and removed from the circulation by a highly efficient memory cytotoxic T-cell response. However, some of the latency III B-lymphocytes escape T-cell attack by downregulating antigen expression and establishing a stable reservoir of ‘resting’ viral-genome in memory B cells (latency 0). The viral lytic cycle becomes re-activated when these latency 0 B-cells differentiate into the plasma cells. Within this type of cells virus replication causes membrane rupture followed by re-infection of epithelial cells, where lytic replication is re-established (Figure 4.1b).
Figure 4.1: EBV life cycle. Reproduced with permission from Nature Publishing group
EBV has three different latent gene transcription patterns (latency programs) and different malignancies associated with EBV express distinct latent gene transcription patterns (See Table 4.1). These latency programs are as follows:

**Latency I program**

The type I latency program has not been shown to induce proliferation. In addition, the EBNA1 expressed during this period is not recognized by the EBV specific CD8 (+) T-cells in the absence of any other antigens or co-stimulatory factors. The strategy of EBV to down regulate the expression of all other antigenic signals is the survival strategy for its persistence and it does not endanger the life of the virus-carrying host.

<table>
<thead>
<tr>
<th>Latency program</th>
<th>Gene expression</th>
<th>Malignancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency 0</td>
<td>EBERs</td>
<td></td>
</tr>
<tr>
<td>Latency I</td>
<td>EBNA 1, EBERs</td>
<td>Burkitt’s lymphoma</td>
</tr>
<tr>
<td>Latency II</td>
<td>EBNA 1, LMP 1, LMP 2, EBERs</td>
<td>Hodgkin lymphoma, NK/T-cell lymphoma</td>
</tr>
<tr>
<td>Latency III</td>
<td>EBNA 1, 2, 3A, 3B, 3C LMP 1, LMP 2, EBERs</td>
<td>Lymphoproliferative disease</td>
</tr>
</tbody>
</table>

Table 4.1: EBV latency program and latency associated genes and malignancies.

**Latency II program**

This latency program is restricted to the expression of EBNA1, LMP1, and LMP2. This program was initially discovered within the cells of NPC and
later also identified in the malignant cells of Hodgkin’s and NK/T-cell lymphoma patients. Because of the absence of EBNA2, these cells cannot be induced to proliferate on their own unless additional cellular or growth promoting changes occur in the cell environment.  

<table>
<thead>
<tr>
<th>Latency gene products</th>
<th>Proposed functions/ characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBNA1</td>
<td>Required for the replication and maintenance of the episomal EBV genome.</td>
</tr>
<tr>
<td>EBNA2</td>
<td>EBNA2 transactivates LMP1 and LMP2 genes.</td>
</tr>
<tr>
<td>EBNA 3A, 3B, and 3C</td>
<td>EBNA3A and EBNA3C are essential for B cell transformation. EBNA3C upregulates LMP1 gene expression.</td>
</tr>
<tr>
<td>LMP 1, 2A, 2B</td>
<td>Protection of B-lymphocytes against apoptosis. LMP1 – an important mediator of number of signaling pathways. LMP 2A, 2B are required for B-cell transformation</td>
</tr>
<tr>
<td>EBERs</td>
<td>Non polyadenylated RNAs that inhibit interferon induced apoptosis, promote proliferation.</td>
</tr>
<tr>
<td>BART</td>
<td>Bam H1A fragment, present in almost all EBV associated diseases.</td>
</tr>
</tbody>
</table>

Table 4.2: Latency associated viral gene products and their proposed function/characteristics.

Latency III growth program

As indicated in Table 4.1, latency III is characterized by the expression of six nuclear antigens along with the virus associated latent membrane proteins, LMP1, LMP2a, LMP2b, EBERs, and BART. This latency pattern is expressed only in EBV
infected B-lymphocytes. One of the prominent features of this growth program is the constitutive activation of several physiologic pathways involved with growth and differentiation of normal B-cells. For example, LMP1 drives constitutive activation of the NF-κB, Akt and JAK/STAT pathways, resulting in a state referred to as immortalization. Cells expressing latency III are the most prominent during the initial phase of B-cell infection and EBV specific primary CTL response develops in response to this infection. However, in patients with immunodeficiency such as the recipients of organ transplantation or patients with acquired or primary immunodeficiency (AIDS, XLPS) also express latency III program and consequently are at high risk for the development of the lymphoproliferative diseases.268,272,279,283

4.3. Association of EBV with B-cell malignancies

Figure 4.2 represents a model linking the life cycle of EBV to the development of the EBV associated B-cell malignancies. As shown in Figure 4.1, EBV usually infects naïve and/or memory B cells. How these cells evolve to acquire malignant features seen in the lymphomas is poorly understood. Hodgkin’s disease, which is a B-cell lymphoma, arises from an EBV infected cell that is blocked at the germinal center stage in its growth development. This ultimately leads to the expression of the Latency II program in the Reed Sternberg cell, the malignant tumor cell of Hodgkin’s disease. Burkitt’s lymphoma, on the other hand, results from a germinal center cell that is entering the memory compartment but is constitutively proliferating because of an activated c-myc oncogene.
Figure 4.2: The sequence of events in the life cycle of EBV development inside B-cells that may be linked with the development of malignancies in B-cells.\textsuperscript{272,287}

The events which occur in healthy carriers are highlighted with black arrows whereas the probable events leading to malignancies are indicated in grey. CTL: Cytotoxic T cells.

BL is characterized by three different variants, which are endemic, sporadic, and immunodeficiency related.\textsuperscript{267,268,271,272,276} Out of these three variants, the endemic form appears to be most closely associated with EBV infection and occurs in equatorial Africa and Central America in the same geographic regions as malaria. The errors caused by faulty DNA repair machinery of an antigen specific clone driven by chronic immune stimulation of malarial infection is believed to be a contributing factor to genomic instability and translocation of the $c$-$myc$ protooncogene.\textsuperscript{288,289} The BL cell has a limited EBV gene repertoire that expresses EBNA1 and LMP2a.\textsuperscript{272,283,288} It should be noted that for both HD and BL the sequence of events leading to the development of malignancy most likely involve multiple genetic and epigenetic factors that eventually result in evolution of a transformed B cell clone.\textsuperscript{268,271,272,288,289}
Post transplant lymphoproliferative disease (PTLD) and other lymphoproliferative disease arise as a direct consequence of immune suppression. Loss of efficient adaptive cellular immunity results in unchecked EBV gene expression, the consequences of which involve permissive expression of all 9 latent genes.\textsuperscript{272,290} While many, if not all, of the latent genes are highly immunogenic, the medically induced T cell deficiency allows latency III EBV infected B-cells to escape immune surveillance and undergo unchecked proliferation.\textsuperscript{272,290} Other factors including genetic susceptibility may also play a role, however, more definitive studies are needed.\textsuperscript{290-292}

The model described in Figure 4.2 does not include malignancies in which the tumors arise in non B lymphoid cell types. Such tumors may result from epithelial cell infection with EBV leading to the oncogenic expression of the latent genes such as LMP1.\textsuperscript{268} Examples of such scenarios include tumors such as nasopharyngeal carcinoma, gastric carcinoma, and NK/T-cell lymphoma. It was observed that in these cases, there is a high degree of correlation between the disease and EBV apart from the factors such as genetic and/or environmental predisposition.\textsuperscript{293-295}

4.4. The lytic phase and associated genes and gene products

During the lytic cycle, EBV genes are expressed in a sequential order in an immediate-early (IE)-, an early (E)-, and a late (L) phase (See Figure 4.3). The majority of the early gene products are responsible for EBV DNA replication whereas the late gene products are structural proteins. The promoters for the IE genes BZLF1 and BRLF1 are Zp and Rp, respectively. Both promoters cause transactivation of several E and L promoters.\textsuperscript{268,272}
In the mid 1990’s, it was discovered that a high expression level of BZLF1 in a latently infected B-cell line was essential to trigger the switch from the latent to the lytic cycle of EBV. During the reactivation of the lytic phase, BZLF1 initially activates Rp, followed by activation of many early (E) lytic genes by both BZLF1 and BRLF1 gene products. BZLF1 and BRLF1 also activate the expression of their respective promoters (Zp and Rp) as well as several transcription factors. Eventually, this leads to the activation of all lytic genes and their gene products resulting in the production of infectious virus particles.\textsuperscript{268,272}

![Lytic cycle associated genes](image)

**Figure 4.3**: Important lytic cycle associated genes

In latently infected B-cells, the promoters Zp and Rp are inactive. This transcriptional repression enables EBV to maintain the latent phase. Epigenetic modifications of viral DNA, including DNA methylation and histone deacetylation, or the inhibition of the
transcription factor binding to DNA may be employed by EBV to achieve this transcriptional repression.\textsuperscript{268,272,296-300} CpG methylation also appears to be responsible for the silencing of EBV antigens, such as EBNA2, EBNA3A, EBNA3B, EBNA3C, LMP1, the immediate-early lytic antigens such as BZLF1, BRLF1, and kinases expressed during the viral lytic cycle. This hypermethylation of genes can be reversed with DNA methyltransferase inhibitors.\textsuperscript{268,272,296-300}

The activation of the Zp requires cis-acting cyclic AMP-responsive elements (CRE) and MEF2D binding motifs (See Figure 4.4). Two types of cis-acting motifs, namely ZI (ZIA, ZIB, and ZID) and ZII are crucial for the activation of Zp, and hence, the activation/induction of the lytic cycle of EBV (See Figure 4.4).\textsuperscript{268,272,296-300} These cis acting motifs are DNA sequences that regulate the expression of genes located on the same strand. In the latent cycle of infection, ZI binds to the cellular protein MEF2D. It was postulated that MEF2D interacts with histone deacetylating complexes (HDACs). This represses the transcription of Zp leaving EBV in the latent phase.\textsuperscript{268,272,296-300}

The phosphorylation state of MEF2D may change as a result of B-cell activation leading to the initiation of the lytic phase. Another way to initiate the transcription of Zp, and thus activate the lytic phase, is the phosphorylation of HDACs, which disrupts the interaction between MEF2D and HDACs.\textsuperscript{268,272,296-300}

ZII is also an important motif for the induction of BZLF1 transcription (See Figure 4.3).\textsuperscript{301} The ZII motif binds CREB, ATF-1, and the ATF-2/c-jun heterodimer. Cellular kinases, especially the c-jun N-terminal kinase (JNK) and the stress Map kinase p38, are involved in the activation of cellular transcription factors that bind to the ZII motif. This
is important because many of the stimuli that induce the EBV lytic cycle are also known to activate JNK and stress Map kinase p38 (see Sections 4.4).

**Figure 4.4**: The EBV early promoter Zp.\(^{301}\) Reproduced with permission from Elsevier Inc.

Based on the knowledge of the epigenetics associated with the maintenance of the EBV latent phase and the induction of the EBV lytic phase, two approaches could be used to achieve the latter. One approach would be to express EBV IE gene products by administration of heterologous promoters (gene delivery), the other would be lytic induction by means of expressing of viral IE gene products using externally administered drug(s).\(^{282}\) As will be discussed in the following **Section 4.5**, the viral lytic cycle may become activated following numerous stimuli. Experimentally, it has been shown that the lytic induction can be induced in latently EBV infected BL cells (latency I), EBV positive gastric carcinoma cells (latency II) and in lymphoblastoid cells (latency III).\(^{283,302}\) Thus it is possible to induce the lytic phase at any stage of the latent infection.\(^{283,302}\) The induction of the viral lytic phase in EBV associated malignancies is important in a therapeutic context because this could lead to both direct anti tumor activity (e.g. through
cell membrane rupture) and enhanced tumor cell sensitivity to other drugs, such as nucleoside prodrugs, that could target both viral and cellular proteins.

4.5. Novel therapies for EBV associated malignancies

As recently reviewed by Kenney et al.\textsuperscript{282} and Pagano et al.,\textsuperscript{303} novel treatment strategies and drugs for EBV associated diseases include suicide gene therapy, vaccination, small interfering RNA (siRNA), antisense RNA, hydroxyurea, Foscarnet, phosphonoacetic acid, and Maribavir (see Figure 4.5 for structures). Most of these strategies and therapeutic agents do not address the issue of the induction of the lytic form of infection.

Intentional induction of the EBV lytic phase as a means of treating EBV associated cancers can be accomplished with several chemotherapeutic approaches. DNA methyltransferase inhibitors, such as azacytidine and 2’-deoxyazacytidine, have proven to be very effective in this context.\textsuperscript{297,304} HDAC inhibitors, such as valproic acid\textsuperscript{298} and arginine butyrate,\textsuperscript{300,305} can also induce EBV lytic phase. DNA demethylation and histone acetylation achieved with the use of these drugs also reactivates cellular tumor suppressor genes\textsuperscript{297,298,300,305} and the combined effects may offer a profound therapeutic benefit.

Chemotherapeutic agents such as TPA (Phorbol ester),\textsuperscript{268,282,306} Doxorubicin,\textsuperscript{306} Gemcitabine,\textsuperscript{306} Methotrexate,\textsuperscript{306} Bortezomib,\textsuperscript{51} and Taxol have shown to activate the promoters (Zp and Rp) responsible for the induction EBV lytic phase in EBV positive epithelial cells and/or B-cell tumors.\textsuperscript{268,272,280,282,307} These agents also have been associated with the activation of at least three different signal transduction pathways, including the stress MAP kinase p38, PI3 kinase, and protein kinase C, which may lead to
the production of ZII activating factors (see Section 4.4). However, the use of chemotherapeutic agents such as HDAC inhibitor to induce the lytic phase also lead to the induction of survival pathways in EBV+ cells, which may drive chemo/radio resistance.\(^{308}\)

Recently, it was also shown that high cellular levels of NF-κB inhibit the activation EBV lytic phase promoters as well as viral protein synthesis and replication.\(^{51,309-313}\) Consequently, the inhibition of NF-κB in latently infected B-lymphocytes \textit{in vitro} led to lytic protein synthesis.\(^{51,309-313}\) It is therefore likely that deactivation of NF-κB signal transduction is at least partially responsible for lytic EBV induction following the treatment with NF-κB inhibiting chemotherapeutics such as Bortezomib.\(^{51,309-313}\)

Two virally encoded kinases, namely EBV thymidine kinase (ORF: BXLF1, EBVTK) and the CMV (UL97) protein kinase homologue, BGLF4, are expressed during the lytic phase.\(^{280,314,315}\) The substrate characteristics of EBVTK have already been discussed in Chapter 1, Section 1.5. Indeed, hTK1 negative/EBVTK positive 143b osteosarcoma cells as well as EBVTK positive melanoma cells proved to be susceptible to treatment by AZT.\(^{315,316}\) The prodrug Ganciclovir (GCV, Figure 1.3) was monophosphorylated by recombinant BGLF4.\(^{317}\) As in the case of AZT (See Chapter 1, Figure 1.3 for structures), GCV monophosphate is converted by cellular mono- and diphosphate kinases to the cytotoxic triphosphate form, which inhibits both viral and cellular DNA polymerase. This results in the premature termination of the cellular DNA synthesis and cancer cell death.\(^{280,314,315}\) Additionally, phosphorylated GCV can induce “bystander” killing to the adjacent cells.\(^{280,314,315}\) The induction of lytic EBV infection in EBV positive LCL-1 tumors by Gemcitabine (See Chapter 1, Figure 1.3 for structures)
and Doxorubicin conferred sensitivity to the cytotoxic effects of GCV.\textsuperscript{280} GCV was also effective in SNU-719 EBV-positive gastric cancer cells following prior exposure to the lytic phase inducer 5-aza-2’-deoxycytidine.\textsuperscript{304} The combined use of AZT/GCV has been effectively used to treat patients with primary central nervous system lymphomas.\textsuperscript{318}

\textbf{4.6. Problems associated with current therapies directed towards the lytic induction in EBV positive tumors.}

As discussed in \textbf{Section 4.3.1}, DNA methyl transferase inhibitors, HDAC inhibitors, and various chemotherapeutic agents have proven to be inducers of the lytic phase of EBV and especially GCV appeared to be very effective in combination with some of these agents in preclinical studies with EBV positive tumors. Therefore, it seems feasible to develop such a “combination therapy” to the clinical level. However this strategy currently suffers from several challenges, which are discussed in the following.

Viral DNA polymerase inhibitors or nucleoside/protein kinase substrates/inhibitors that have been explored clinically and/or experimentally against EBV associated malignancies with or without the induction of the lytic cycle comprise of acyclic purine nucleoside analogues such as Acyclovir\textsuperscript{319,320}, 9-(1,3-dihydroxy-2-propoxymethyl) guanine (DHPG),\textsuperscript{321} GCV,\textsuperscript{32,269,281,300} and Penciclovir,\textsuperscript{322} pyrimidine nucleoside analogues such as AZT,\textsuperscript{315,316} L-FMAU,\textsuperscript{323-325} D-FMAU,\textsuperscript{323,326,327} D-FIAU,\textsuperscript{323,327} L-I-OddU,\textsuperscript{328,329} Brivudine,\textsuperscript{323,327} and acyclic nucleotide analogues such as Cidofovir\textsuperscript{330,331} and Adefovir\textsuperscript{330,331} (\textbf{Figure 4.5}).

However, none of these agents have been approved specifically for the treatment of EBV associated diseases, and, except for the L-dioxolanyl nucleosides L-I-OddU (see
Section 4.6), none of them show any specificity for the EBV-kinase/DNA polymerase system and/or were specifically designed for it. As discussed in detail in Chapter 1, Section 1.5, γ-herpesviral TKs have evolved separately from α- and β-herpes virus TKs, and thus, their substrate specificities differ significantly.\(^{32}\) Similar differences in substrate specificities have been observed in the case BGLF4 and the protein kinases from α- and β-herpes viruses, such as pUL97, pR97, and pUL13.\(^{317}\) This may explain the moderate success of nucleoside analogues (NAs), such as AZT and GCV, in the experimental therapy of EBV associated malignancies, since these agents were specifically developed to target the kinase systems of HIV and α-/β-herpes viruses, respectively. Another major limitation for the development of EBV-specific NAs is the lack of crystal structures of EBVTK, BGLF4, and EBV DNA polymerase for structure-based design.

These shortcomings provide the rationale for the studies described in Chapter 4, which deal with the development of a homology model of EBVTK and the design and synthesis of novel dioxolane nucleosides that selectively target EBVTK for an improved “combination therapy” of EBV associated malignancies. Preliminary biological studies related to these novel EBVTK-targeting agents will also be described.
Figure 4.5: Nucleosides and non-nucleosides agents used for studies with EBV associated malignancies.
4.7. Homology Modeling of EBVTK

As discussed in Section 4.6, there is an urgent need to explore the interaction of small molecules with the active site of EBVTK to develop EBVTK specific substrates and/or inhibitors. Since X-ray crystal- or NMR-based 3D-structures of EBVTK are not available, the development of an EBVTK homology model may facilitate the structure-based design of EBVTK substrates/inhibitors. The crystal structure of Equine Herpes Virus type-4 thymidine kinase EHV4TK (PDB ID # 1P6X) should be suitable as a template for the development of a homology model of EBVTK. According to basic local alignment search tool for proteins (BLASTP) hosted on the website http://blast.ncbi.nlm.nih.gov/Blast.cgi, both proteins have an amino acid sequence homology of 30 %.

Crystal structures of (EHV4TK) were reported in a complex with i) dThd and ADP (PDB # 1P72), referred to as EQ1 henceforth, ii) dThd and SO$_4^{2-}$ (PDB # 1P6X), referred to as EQ2 henceforth, iii) a bisubstrate inhibitor in which adenosine and dThd are linked between their 5’-positions through 4 phosphate groups (TP$_4$A, PDB # 1P73), referred to as EQ3 henceforth, and iv) a bisubstrate inhibitor in which adenosine and dThd are linked between their 5’-positions through 5 phosphate groups (TP$_5$A, PDB # 1P75), referred to as EQ4 henceforth. All structures were crystallized as dimers with a very characteristic difference in the “lid” conformations between the individual monomers in both EQ1 and EQ2. The structural difference between the “lid” conformations of the aligned monomers of EQ1 is shown in Figure 4.6.
Figure 4.6: EHV4TK (EQ1, PDB # 1P672). Overlay of the two monomers of EQ1 having dThd and ADP in the substrate- and catalytic binding site, respectively. The monomer with the closed lid is depicted in green and the monomer with the open lid in red.

4.7.1. The substrate binding site and the catalytic site of EHV4TK

Figure 4.7 illustrates the binding interactions of dThd and ADP within the EQ1 crystal structure. The thymine base forms two hydrogen bonds with a Gln 102, which is conserved among the family of herpes viral TKs. It is also sandwiched via π stacking interactions between Phe 105 and Phe 148. The 3′- and the 5′-hydroxyl group of the 2′-deoxyribose sugar of dThd interact via H-bonding and/or ionic interactions with Tyr 78/Glu 201/Arg 196, and Glu 60/Arg 139, respectively. It was speculated that Glu 60
may act as a base during the phosphoryl transfer reaction from ATP to dThd because of its close proximity (2.94 Å) to the 5’-hydroxyl group.\textsuperscript{332} The binding of ADP near the catalytic site (P-loop) is stabilized through H-bonding/ionic interactions of the β-phosphate of ADP with of Arg 196 and π stacking with Arg 192.

\textbf{Figure 4.7:} The amino acid binding interactions of dThd and ADP in EQ1.
The lid region is a highly flexible and conserved motif in many nucleoside kinases and nucleoside monophosphate kinases (see also Chapter 1, Section 1.8). \textsuperscript{333-335} EHV4TK has been shown to have dual nucleoside kinase and nucleoside monophosphate kinase activity. \textsuperscript{332} The lid region alters its confirmation from an open state to a closed state upon ATP binding. \textsuperscript{332} This conformational change moves one or more of the arginine residues (Arg 196, Arg 192, Arg 139) closer to ATP, and thus, makes the latter available for the phosphate transfer reaction. Therefore, the “closed lid” form is the catalytically active conformation.

In the case of EHV4TK, all monomers of all reported crystal structures possess either ADP, bisubstrate inhibitors (TP\textsubscript{4}A, TP\textsubscript{5}A), or SO\textsubscript{4}\textsuperscript{2-} in the ATP binding site. Therefore, it was unexpected to observe the open lid conformations in one of the monomers of EQ1 and EQ2. It was hypothesized that the “open lid” conformation in these monomers may be due to a crystallization artifact. \textsuperscript{332} The availability of two conformational forms of EHV4TK allows for the preparation of the two corresponding homology models of EBVTK, which permits a more extensive and detailed investigation of the binding modes of novel EBVTK substrates/inhibitors.

### 4.7.2. EBVTK homology models

Homology models of EBVTK were built with SWISS-MODEL \textsuperscript{336-339} using the crystal structure of EHV4TK (PDB ID # 1P6X/ EQ1) as the template. As mentioned earlier, the crystal structure of EQ1 is a dimer with one monomer having an open and the other a closed lid conformation. Both monomers were chosen to generate the corresponding open and closed lid conformations of EBVTK.
Figure 4.8: Amino acid sequence alignment of EBVTK on EHV4TK.

Amino acid residues important for dThd binding interactions are highlighted in red for both EBVTK and EHV4TK. Amino acid residues important for interactions with the β-phosphate of ADP are highlighted in blue.

Based on the sequence alignment of EBVTK on EHV4TK shown in Figure 4.8, it is obvious that the homology is especially very high for amino acids that are essential for the interaction in the substrate binding- and the catalytic site. This high local homology between EHV4TK and EBVTK was a major reason for the selection of EHV4TK as the
template for the homology model of EBVTK even though the overall amino acid sequence homology between both proteins is moderate with about 30%.

As shown in Figures 4.9, the obtained homology models of EBVTK reproduced the characteristic difference in the “lid” regions observed in EQ1 whereas the remaining portions had similar structural features.

**Figure 4.9**: Side view of the aligned “open lid” and the “closed lid” forms of the EBVTK homology models.

Open form = Red, Closed form = Green
4.8. Design of substrate and/or inhibitors of EBVTK

The design of novel improved EBVTK substrates/inhibitors was based on following two findings:

1) Povey et al. reported in 1995\textsuperscript{245} that N3-methyl thymidine (N3-Me-Thd) was phosphorylated to a two fold higher extent than the endogenous substrate dThd in lysates of SF-9 cells, which were transfected to express high levels of EBVTK. To the best of our knowledge, this is the only report that indicates that substitution at the N3-position of the endogenous substrate dThd was not only tolerated by EBVTK but even facilitated phosphorylation.

2) Chu et al. evaluated \textit{in vitro} the anti-EBV activity of several 5-halo (I, Br, Cl, F) and 5-trifluoromethyl substituted β-L-dioxolane uracil analogues in the high-yield EBV-producing cell line H1, which has EBVTK activity.\textsuperscript{329} The anti-EBV activity was determined in terms of percent inhibition of EBV DNA replication (EC\textsubscript{50} in µM) in EBV infected H1 cells compared with an untreated control. It was found that the anti-EBV activity for these analogues decreased in the following order: 5-iodo-L-dioxolanyl uracil [L-I-OddU] (EC\textsubscript{50} = 0.03 µM), 5-bromo-L-dioxolanyl uracil (EC\textsubscript{50} = 0.19 µM), 5-chloro-L-dioxolanyl uracil (EC\textsubscript{50} = 0.6 µM), 5-trifluoromethyl-L-dioxolanyl uracil (EC\textsubscript{50} = 4.0 µM) and 5-fluoro-L-dioxolanyl uracil (EC\textsubscript{50} = 50 µM). L-I-OddU produced approximately 80% inhibition of EBV DNA replication after 36 h exposure compared to untreated H1 cells. Interestingly, D-I-OddU also showed moderate inhibition of EBV DNA synthesis in H1 cells but, in contrast to L-I-OddU, it was also accompanied with significant cytotoxicity, indicating that this agent also interfered with cellular functions.\textsuperscript{329} Chu and coworkers later evaluated various 5-halovinyl substituted uracil
analogs as anti-EBV agents in EBV producing H1 cells.\textsuperscript{329,340} Two of these agents, 5-chlorovinyl (EC\textsubscript{50} = 0.49 \(\mu\)M) and 5-bromovinyl (EC\textsubscript{50} = 0.59 \(\mu\)M) dioxolanyl uracil, showed moderate capacity to inhibit EBV DNA replication.\textsuperscript{340}

Kira and Chu \textit{et al.} demonstrated that treatment of H1 cells with L-I-OddU resulted in the inhibition of the replication of linear viral DNA\textsuperscript{328} indicating that the virus was actively replicating and that the cells were in a lytic phase. Linear EBV genomes indicate viral replication, whereas circular genomes are indicative of latently infected cells.\textsuperscript{272,283} L-I-OddU did not produce any inhibition of DNA replication in L5 cells, which were supposedly latently infected with EBV and did not have any EBVTK activity.\textsuperscript{328} This study also indicated that L-I-OddU was phosphorylated about 2.5 times more effectively than dThd by purified EBVTK protein. In contrast, L-I-OddU was not a substrate of hTK1.\textsuperscript{328} The \(K_m\) of L-I-OddU with EBVTK was estimated to be 5.5 \(\mu\)M, which was similar to that obtained with the endogenous substrate dThd but about five fold higher than that obtained with L-FMAU. The relative \(V_{\text{max}}\) was found to be seven fold higher than that of dThd.\textsuperscript{328} In addition to significant amounts of L-I-OddU monophosphate, minor quantities of L-I-OddU di- and triphosphate were presumably also detected in H1 cells that were exposed to L-I-OddU. Diphosphorylation could have been performed by EBVTK or cellular thymidylate kinase whereas cellular NDPK is most likely responsible for triphosphorylation. Since only very small amounts of intracellular L-I-OddU triphosphate were detected, Kira and Chu \textit{et al.} speculated that inhibition of EBV polymerase by L-I-OddU triphosphate may not have been responsible for decreased EBV DNA replication and that either the mono- or diphosphate may have been the active form of L-I-OddU.\textsuperscript{328}
So-called bicyclic pyrimidine nucleoside analogues (BCNAs) are very interesting in the context of this speculation. These agents were identified as highly selective substrates for VZVTK and potent inhibitors of VZV replication in HEL cells. Viral toxicity of BCNAs was depended on VZVTK activity.\textsuperscript{341-344} It is found that BCNAs were converted into their monophosphates and also to their diphosphates by VZVTK because of the additional thymidylate kinase activity of VZVTK.\textsuperscript{341-344} However, BCNA diphosphate was not converted into BCNA triphosphate in human erythrocytes indicating that it is not a substrate of cellular NDPK.\textsuperscript{345} Balzarini \textit{et al.} therefore stated that it is not known whether the active metabolite of BCNAs is the mono-, di-, or triphosphate and that ribonucleotide reductase or even enzymes not associated with viral DNA synthesis rather than VZV DNA polymerase may be the molecular target of phosphorylated BCNA. Similarly, the anticancer prodrug Gemcitabine (Chapter 1, Figure 1.3) exerts its cytotoxicity to a significant extent through its diphosphate form as a ribonucleoside reductase inhibitor.\textsuperscript{346-348} Also, L-FMAU was identified as a potent inhibitor of EBV replication. However, the triphosphate form of L-FMAU was not a substrate of EBV DNA polymerase, and it was not incorporated into EBV DNA. The mechanism of action of L-FMAU triphosphate appeared to be more consistent with that of an allosteric at EBV DNA polymerase.\textsuperscript{325,326}

Overall, it seems that the long-standing paradigm that antiviral and anticancer nucleoside prodrugs are active as triphosphates by polymerase inhibition and incorporation into DNA is not valid any more and that nucleoside kinases become increasingly important as molecular targets.
Based on these findings, we decided to design and synthesize a small focused library of known and novel D- and L-dioxolanyl uracil analogues (Figure 4.10) and to evaluate their inhibitor/substrate characteristics in enzyme assays with EBVTK and hTK1. The known compounds, shown in red color in Figure 4.10, were previously evaluated in vitro by Chu and coworkers. The novel compounds, which are primarily several N3-substituted D- and L-dioxolanyl uracil analogues, are shown in blue in Figure 4.10. The intention of the planned biological studies is limited and focuses initially only on the exploration of the combined effect of the dioxolane moiety and the N3-substitution on the monophosphorylation of these types of compounds by hTK1 and EBVTK. Based on the prior discussions in Chapter 4, we hypothesize that N3 substituted L-dioxolanyl uracil analogues will be selective substrates/inhibitors of EBVTK and that they are superior to L-1-OddU in this capacity! If this should indeed be the case, the design and synthesis of a larger compound library and extended in vitro and in vivo studies are warranted to explore if these agents are suitable components in a clinical “combination therapy” (Section 4.4.2) of EBV associated malignancies.
Figure 4.10: Proposed novel and existing compounds to be evaluated in EBVTK enzyme assays.

The proposed novel nucleosides are shown in blue and known compounds are shown in red. Only N3-Me-Thd and L-I-OddU were found to be good substrates of EBVTK.
In order to substantiate our design strategy, we carried out a preliminary comparative docking study with the “closed” EBVTK homology model using dThd and L-N3-methyl-dioxolanyl thymine (compound 4.72, Figure 4.10) as substrates (Figures 4.11 A & B). Initially, Surflex\textsuperscript{99,100,115} was used to identify the binding site of dThd in the “closed” homology model of EBVTK (see Chapter 5 for experimental details and Chapter 2 for a general discussion of Surflex algorithms and docking). This was followed by the docking of dThd and 4.72. The quality of the homology model generated by SWISS-MODEL and the identification of the active site by Surflex proved to be credible as both operations reproduced the binding pattern observed for the interaction of dThd with the active site of EHV4TK (see Figures 4.7 & 4.11 A). Also, dThd and 4.72 showed very similar binding patterns in the active site of the EBVTK homology model indicating that the latter should be a good substrate (see Figures 4.11 A & B).
Figure 4.11: Comparison of the docking patterns of the dThd (A) and L-N3-methyl-dioxolanyl thymine (B, compound 4.72) in the active site of the “closed” homology model of EBVTK.

For comparison see also Figure 4.7, which illustrates the spatial orientation of Glu 60 and Gln 102 in relation to dThd in the crystal structure of EHV4TK.
4.9. Chemistry

4.9.1. Synthetic Strategies

In order to synthesize the proposed known and novel D- and L-dioxolane uracil analogues shown in Figure 4.10, we decided to employ traditional Vorbrüggen-type glycosylation/condensation reactions\textsuperscript{329,349,351-353} as the key step in their synthesis, as shown in Figure 4.12.

Figure 4.12: Synthetic Strategy for the synthesis of D- and L-dioxolanyl nucleoside analogues.
Initial treatment of various substituted and unsubstituted thymine and uracil nucleobases with hexamethyldisilazane (HMDS) resulted in the protection of the carbonyl oxygens of the nucleobases with trimethylsilyl (TMS) groups (Figure 4.12 A). This was followed by the condensation of these silylated nucleobases with the protected D- or L-dioxolanyl acetate and subsequent deprotection of all protective groups to furnish the D- and L-dioxolanyl uracil analogues as target products (see Figure 4.12 B).

4.9.2. Synthesis of N3-substituted nucleobases

N3-alkylated thymine and uracil nucleobases were synthesized by using two different protective group strategies. The first method employed the selective N3-alkylation of commercially available deoxyribose nucleosides followed by deglycosylation with I$_2$, as discussed in detail in Chapter 3. The second method utilized the selective N1 Boc-protection of thymine and uracil followed by N3-alkylation and removal of the Boc protective group. C-5 substituted nucleobases, such as thymine (4.4), 5-iodouracil (4.6), and 5-bromovinyl uracil (4.13), were commercially available.

4.9.2.1. Synthesis of N3-substituted thymine analogues using the deglycosylation strategy (see also Chapter 3).

Thymidine (4.1) was reacted with methyl iodide in presence of 3 equivalents of K$_2$CO$_3$ at 50 °C for 3 h to afford N3-methyl thymidine (4.2) in 85 % yield after purification (Scheme 4.1 A). Compound 4.2 was then subjected to the deglycosylation reaction with I$_2$ in acetonitrile under reflux conditions for 24 h to produce N3-methyl thymine 4.3 in 61 % yield. This methodology is simple, convenient, and so far
unreported, general method for the synthesis of N3-alkylated thymines as it utilizes inexpensive dThd as the starting material and requires only two reaction steps. It should be noted, however, that the same deglycosylation conditions were not successful in the synthesis of N3-alkylated uracils.

4.9.2.2. Synthesis of N3-substituted uracil and thymine analogues using the N1-Boc protection strategy

A regioselective N1-Boc protection was employed for the synthesis of the N3-alkylated thymine/uracil bases 4.10, 4.11, and 4.12 (Scheme 4.1 B).<sup>263</sup> The nucleobases 4.4 - 4.6 were reacted with di-tert-butyldicarbonate in the presence of DMAP in anhydrous acetonitrile at room temperature for 2 h to afford the corresponding N1-Boc-thymine (4.7), N1-Boc-uracil (4.8), and N1-Boc-5-iodouracil (4.9), all in approximately 50 % yield after purification. Compounds 4.7 was reacted with ethyl bromide in the presence of NaH in DMF at room temperature for 2 h followed by in situ removal of the Boc-protective group with 0.5 eq K<sub>2</sub>CO<sub>3</sub> in methanol at room temperature to afford 4.10 in 70 % yield. Similar procedures were applied for the syntheses of 4.11 (70 % yield) and 4.12 (75% yield) starting from 4.8 and 4.9, respectively. In contrast to the deglycosylation procedure (Section 4.9.2.1), this method requires three reaction steps.

Reagents: A) i) MeI, K$_2$CO$_3$, DMF:acetone (1:1), 50 °C, 3 h; ii) I$_2$, acetonitrile, 82 °C, 24 h; B) i) di-tert-butyldicarbonate, DMAP, acetonitrile, rt, 2 h; ii) alkyl halides, NaH, DMF, 2 h; iii) K$_2$CO$_3$, MeOH, rt 2 h.

The regioselectivity of Boc substitution (N1 vs N3) was established by 2D-NMR spectroscopy using HMBC (Heteronuclear Multiple Bond Correlation Experiment) techniques. The hydrogen attached to N3 of 4.7 did not couple with the carbons C5 and C6 (see Appendix, Figure A9), which was the expected outcome as opposed to a compound with a hydrogen at N1 and the Boc group at N3. In the later case the hydrogen at N1 should have coupled with C5 and C6 because of the closer proximity.
As an indirect way to determine the regioselectivity of Boc substitution, we also carried out 2D experiment with 3-butylthymine (compound 3.25, Chapter 3), which was, however, synthesized according to strategy B shown in Scheme 4.1. The HMBC spectrum of 3.25 showed coupling of the \( \alpha \)-methylene hydrogens next to N3 with C2 and C4, as indicated in Figure A10 of the Appendix with arrows. In the corresponding N1-butylthymine, these would not have coupled with both C2 and C4. Furthermore, the HMBC spectra of 3.26, which was synthesized using strategy A shown in Scheme 4.1, was identical to that of HMBC spectra of 3.26 obtained using strategy B of Scheme 4.1. This proved unequivocally that the alkyl groups in the compounds synthesized using strategy B in Scheme 4.1 are indeed located at N3 and that the initial Boc protection must have occurred at N1.


The synthesis of 4.23 was achieved in 10 steps using commercially available D-mannose as the starting material according to a previously described procedure.\(^{351}\) Minor modifications were especially necessary for the synthesis of D-1,6-anhydromannose (4.15).\(^{354}\) Briefly, the 6-hydroxyl group of D-mannose was tosylated using p-toluensulfonyl chloride (p-TsCl) in pyridine at 0 °C for 24 h to afford the corresponding tosylate 4.14 in 55 % yield after purification. In the next step, the intramolecular nucleophilic attack of the anomeric hydroxyl group of 4.14 on the tosyl functionality under basic conditions caused ring closure to furnish 4.15.
The previously described procedure \textsuperscript{351,355} for the synthesis of 4.15 calls for the use of 3 N aqueous NaOH to adjust the pH to 9. After 90 min reaction time, neutralization (pH 7) of the reaction medium is carried out with 3 N HCl to furnish 4.15 in poor yields (~25-30 %). Therefore the synthesis of 4.15 was carried out by adopting a different procedure, which involves the isolation of the intermediate tosylate, 4.14, followed by cyclization with 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in ethanol for 24 h at room temperature to afford 4.15 in 70 % yield.\textsuperscript{354} The 2, 3-geminal hydroxyl functions of 4.15 were protected as acetonide by reaction with dimethoxypropane and a catalytic amount of p-TsOH at room temperature for 18 h to furnish compound 4.16. The 4-hydroxyl group of 4.16 was benzyolated without prior purification using benzoyl chloride at 0 °C for 45 min to furnish 4.17 in 70 % overall yield from 4.15. The acetonide group of 4.17 was then selectively deprotected using diluted H\textsubscript{2}SO\textsubscript{4} (0.004 %) in 60% aqueous dioxane at 80 °C for 15 h to give the 2,3-geminal diol 4.18 in 85 % yield.
Scheme 4.2: Synthesis of the D-dioxolanyl acetate 4.23 for condensation with various nucleobases.\textsuperscript{351}
Reagents: a) p-TsCl, pyridine, 0 ºC, 24 h; b) DBU, EtOH, rt, 24 h; c) 1,2-dimethoxypropane, p-TsOH, acetone, rt, 18 h; d) BzCl, pyridine, 0 ºC, time; e) dil. H$_2$SO$_4$ (0.004 %), 60 % dioxane, 80 ºC, 15 h; f) NaIO$_4$, NaBH$_4$, EtOH, 0 ºC to rt, 1.1 h; g) TBDPSCI, imidazole, DMF, rt, 2.5 h; h) NaOMe, MeOH, rt, 2 h; i) NaIO$_4$, RuO$_2$, H$_2$O, rt, 12 h; j) Pb(OAc)$_4$, pyridine, ethyl acetate, rt, 15 h.

Treatment of **4.18** with NaIO$_4$ for 1 h in 95 % ethanol at room temperature followed by reduction with NaBH$_4$ at 5ºC for 10 min afforded the dioxolane **4.19** in 70 % yield. During this reaction the secondary benzoyl group in **4.18** migrates to the primary hydroxyl group. The hydroxymethyl group at the 6´-position of **4.19** was then selectively protected with tert-butyldiphenylsilyl (TBDPS) in presence of imidazole in DMF for 2.5 h to yield **4.20** in 81 % yield. Removal of the benzoyl group of **4.20** was achieved by treatment with NaOMe in methanol at room temperature for 2 h to produce the **4.21** in 85 % yield. Oxidation of **4.21** to the carboxylic acid **4.22** was accomplished by treatment with NaIO$_4$ in presence of a catalytic amount of RuO$_2$ at room temperature overnight to afford **4.22** in 70 % yield. Oxidative decarboxylation of **4.22** with Pb(OAc)$_4$ in the presence of pyridine in ethyl acetate at room temperature for 15 h afforded the dioxolane acetate, **4.23** in 50% yield.
Scheme 4.3: Condensation of silylated nucleobases with D-dioxolanyl acetate 4.23.\textsuperscript{351}

Reagents: a) i) HMDS, cat. (NH\textsubscript{4})\textsubscript{2}CO\textsubscript{3}, reflux, 12 h; ii) 4.23, TMSOTf, dichloroethane, rt, 3 h.

* The synthesis of 4.29 A/B was unsuccessful.
4.9.4. Glycosylation/condensation of the D-dioxolanyl acetate with nucleobases and the deprotection of the TBDPS.

The synthesis of D-dioxolanyl uracil and -thymine analogues was accomplished by condensing 4.23 with the silylated forms of the thymine or uracil analogues 4.3, 4.4, 4.6, 4.10, 4.11, 4.13 in presence of TMSOTf in dichloroethane at room temperature for 3 h to obtain mixtures of the TBDPS protected α- and β-nucleosides 4.24 A / B - 4.28 A / B and 4.30 A / B in a 1:1.5 ratio in ~ 60-70 % yields (Scheme 4.3). Unexpectedly, the silylated form of nucleobase 4.12 did not undergo glycosylation with 4.23 to furnish 4.29 A / B despite of the fact that both its non-methylated (4.6) and its non-iodinated (4.11) counterpart reacted smoothly with 4.23. The reason for the lack of reactivity of 4.12 under these reaction conditions is unknown. The α-isomers 4.24 A-4.27 A and 4.30 A were separated from the corresponding β-isomers 4.24 B-4.27 B and 4.30 B by silica gel column chromatography. Isomers 4.28 A and 4.28 B could not be separated by normal silica gel column chromatography. The assignments of α- and β-configurations were based on the coupling constants and chemical shifts reported previously for the C1΄- and C4΄-protons of the obtained nucleosides.329,340,351,352

Compounds 4.24 A / B - 4.27 A / B, 4.28 A / B (as a mixture) and 4.30 A / B were deprotected by treatment with TBAF in THF at room temperature for 1 h to furnish target compounds 4.31-4.38, 4.39 A / B (as a mixture), 4.42, and 4.43, respectively in 75-90 % yields (Scheme 4.4).
Scheme 4.4: Removal of the TBDPS group from 4.24 A/B - 4.28 A/B and 4.30 A/B.

Reagents: a) 1M TBAF in THF, rt, 1 h.

4.9.5. Synthesis of L-(2S, 4S/R)-4-acetoxy-2-[(benzoyloxy)-methyl] dioxolane (“L-dioxolanyl acetate”/4.54) for the condensation with nucleobases

The synthesis of L-dioxolane acetate 4.54 was achieved in 9 steps using commercially available L-gulonic acid-γ-lactone (4.44) as the starting material. 352
Scheme 4.5: Synthesis of the L-dioxolanyl acetate 4.54 for condensation with various nucleobases.\textsuperscript{352}

Reagents: a) Acetone / CuSO\textsubscript{4}, cat H\textsubscript{2}SO\textsubscript{4}, rt, 48 h; b) diisobutylaluminium hydride (DIBAL-H), toluene, -78 °C to rt, 2 h; c) 0.5 N HCl, reflux, 20 h; d) NaIO\textsubscript{4}, NaBH\textsubscript{4}, MeOH-H\textsubscript{2}O, 0 °C, 35 min; e) p-TsOH, acetone, rt, 6 h; f) BzCl, pyridine, 0 °C, 2 h; g) cat p-TsOH, MeOH, rt, 2 h; h) NaIO\textsubscript{4}, cat RuO\textsubscript{2}, CCl\textsubscript{4}-CH\textsubscript{3}CN-H\textsubscript{2}O (1:1:1.5), rt, 5 h; i) Pb(OAc)\textsubscript{4}, pyridine, THF, rt, 45 min.
The two geminal hydroxyl groups of 4.44 were protected as an acetonide using acetone/CuSO$_4$ in presence of cat H$_2$SO$_4$ at room temperature for 48 h to afford 4.45 in approx. 55 % yield. The ketone (lactone) functionality of 4.45 was reduced using DIBAL-H (diisobutyl aluminium hydride) in toluene at -78 ºC followed by stirring at room temperature for 2 h to furnish 4.46 in 75 % yield. This compound was then subjected to treatment with 0.5 N HCl under reflux overnight to produce L-1,6-anhydrogulopyranose (4.47) in ~ 30 % yield. L-gulose (4.48) was obtained in ~ 50% yield as a side product in this reaction, which was again subjected to the treatment with 0.5 N HCl under reflux condition overnight to produce a second batch of 4.47 in 25% yield. The combined yields of the reaction from 4.46 to 4.47 were ~ 55 %.

![Diagram](diagram.png)

**Scheme 4.6:** Step-by-step formation of 4.49 from 4.47.

The one pot oxidation-reduction reaction of 4.47 to produce 4.49 was carried out according to the procedure by Chu *et al.*[^352] Briefly, an aqueous solution of NaIO$_4$ was added at 0 ºC dropwise for 10 min to a solution of 4.47 in methanol, which was followed after 15 min stirring by the addition of NaBH$_4$. This reaction involves the *in situ* generation of dialdehyde 4.55 (Scheme 4.6) by NaIO$_4$ and the subsequent reduction of 4.55 to produce 4.49 as part of a very complex mixture of compounds. According to the
original procedure described by Chu et al.,\textsuperscript{351} acetonide-protection of 4.49 is carried out directly without isolating this compound to produce 4.50. However attempts to protect the geminal hydroxyl groups of 4.49 without prior isolation proved to be tedious in our hands resulting in very low yields. Briefly, the geminal hydroxyl groups of 4.49 were protected as acetonide using acetone in presence of cat. p-TsOH to produce 4.50 in 30 % overall yield from 4.47 compared to 62 % reported previously by Chu et al.\textsuperscript{351} Compound 4.50 was subjected to benzylation by treatment with benzoyl chloride in pyridine: dichloromethane (1:2) for 2 h to afford 4.51 in ~70 % yield. The acetonide protective group of 4.51 was removed using catalytic amounts of p-TsOH in MeOH to yield 4.52 in ~80 % yield. Oxidation of 4.52 to the carboxylic acid 4.53 was accomplished in ~ 70 % yield using NaIO\textsubscript{4} in presence of a catalytic amount of RuO\textsubscript{2} in a biphasic mixture of CCl\textsubscript{4}-CH\textsubscript{3}CN-H\textsubscript{2}O (1:1:1.5) as solvents. Oxidative decarboxylation of 4.53 with Pb (OAc)\textsubscript{4} and pyridine in anhydrous THF afforded the key intermediate 4.54 in ~ 50 % yield. The major problem with the reaction sequence shown in Scheme 4.5 was the poor yield for the conversion of 4.47 into 4.49. There are several possible explanations for the occurrence of the complex mixture that contained apparently only limited amounts of 4.49. A major culprit, however, may have been the intermediate dialdehyde 4.55 (Scheme 4.6), which may undergo nucleophilic attack at a carbonyl carbon by the free hydroxyl group present in the molecule. Protecting this hydroxyl group prior to the sequential oxidation-reduction step could prevent such an intramolecular attack. As shown in the proposed alternative reaction Scheme 4.7 for the synthesis of the L-dioxolane acetate 4.54, however, the protection of this hydroxyl group requires a different synthetic procedure,\textsuperscript{356} including the acetonide-protection of the 2, 3-geminal
hydroxyl groups of 4.47 to yield 4.56 followed by the protection of the 4-hydroxyl group of 4.56 with a suitable protective group (e.g. benzoyl or a silyl). Indeed, a similar protective group strategy was successfully used for the synthesis of the D-dioxolanyl acetate 4.23 (Scheme 4.2).

Scheme 4.7: Proposed alternative synthetic route for the L-dioxolanyl acetate 4.63 with a tert-butyldiphenylsilyl protective group.
4.9.6. Glycosylation/condensation of the L-dioxolanyl acetate (4.54) with nucleobases and the deprotection of the benzoyl group.

\[
\begin{align*}
&\text{R} \quad \text{N} \quad \text{O} \quad \text{R'} \\
&\text{4.3} : R = \text{CH}_3, R' = \text{CH}_3 \\
&\text{4.4} : R = \text{CH}_3, R' = \text{H} \\
&\text{4.6} : R = \text{I}, R' = \text{H} \\
&\text{4.10} : R = \text{CH}_3, R' = \text{CH}_2\text{CH}_3 \\
&\text{4.11} : R = \text{H}, R' = \text{CH}_3 \\
&\text{4.13} : R = \text{CH}=\text{CH-Br}, R' = \text{H}
\end{align*}
\]

4.3, 4.4, 4.6, 4.10, 4.11, 4.13

\[
\begin{align*}
&\text{4.64 A} : R = \text{CH}_3, R' = \text{CH}_3 \\
&\text{4.65 A} : R = \text{CH}_3, R' = \text{H} \\
&\text{4.66 A} : R = \text{I}, R' = \text{H} \\
&\text{4.67 A} : R = \text{CH}_3, R' = \text{CH}_2\text{CH}_3 \\
&\text{4.68 A} : R = \text{H}, R' = \text{CH}_3 \\
&\text{4.70 A} : R = \text{CH}=\text{CH-Br}, R' = \text{H}
\end{align*}
\]

4.64 A, 4.65 A, 4.66 A, 4.67 A, 4.68 A, 4.70 A

\[
\begin{align*}
&\text{4.64 B} : R = \text{CH}_3, R' = \text{CH}_3 \\
&\text{4.65 B} : R = \text{CH}_3, R' = \text{H} \\
&\text{4.66 B} : R = \text{I}, R' = \text{H} \\
&\text{4.67 B} : R = \text{CH}_3, R' = \text{CH}_2\text{CH}_3 \\
&\text{4.68 B} : R = \text{H}, R' = \text{CH}_3 \\
&\text{4.70 B} : R = \text{CH}=\text{CH-Br}, R' = \text{H}
\end{align*}
\]

4.64 B, 4.65 B, 4.66 B, 4.67 B, 4.68 B, 4.70 B

**Scheme 4.8**: Condensation of silylated nucleobases with the L-dioxolanyl acetate 4.54.

Reagents: i) HMDS, cat. (NH$_4$)$_2$CO$_3$, reflux, 125 °C, 12 h; ii) 4.54, TMSOTf, dichloroethane, rt, 3 h.
The synthesis of L-dioxolanyl uracil and thymine analogues was accomplished by condensing 4.54 with the silylated forms of the thymine or uracil analogs 4.3, 4.4, 4.6, 4.10, 4.11, 4.13 according to the procedure described in Section 4.6.3 for the synthesis of the corresponding TBDMS-protected D-isomer (4.23) to obtain mixtures of the 5'-benzoylated α- and β-nucleosides 4.64-4.68 A / B and 4.70 A / B approximately in a 1:1.5 ratio (~60 - 70 % yield) (Scheme 4.8). Because of the failed attempt to condense 4.12 with 4.23 (see Scheme 4.6.3), the glycosylation reaction of 4.12 with 4.54 was not carried out. All α- and β-anomers, except for 4.68 A and 4.68 B, were separated by silica gel column chromatography to give the pure α-isomers 4.64 A-4.67 A and 4.70 A as well as the pure β isomers 4.64 B-4.67 B and 4.70 B. The assignment of α- and β-isomers was based on the coupling constants and chemical shifts reported for the H-1’ and H-4’ protons of the dioxolane sugar.329,340,351,352

Nucleosides 4.64-4.67 A / B, 4.68 A / B (as a mixture) and 4.70 A / B were deprotected by treatment with ammonia in methanol at room temperature to afford 4.71-4.78, 4.79 A / B (as a mixture), 4.82, and 4.83, respectively, in 60-70 % yields (Scheme 4.9).
Scheme 4.9: Removal of the benzoyl group group from 4.64 A / B - 4.68 A / B and 4.70 A / B.

Reagents: i) NH3, MeOH, rt, 48 h.
4.10. Biology

4.10.1. Preliminary enzyme assays with the hTK1 protein (carried out in the laboratory of Dr. Staffan Eriksson in the Department of Anatomy, Physiology, and Biochemistry at the Swedish University of Agricultural Science [Biomedical Centre] in Uppsala, Sweden).

Initial hTK1 phosphoryl transfer assays (see Chapter 5, Section 5.5.2 for experimental details) were carried out with compound 4.34, (Figure 4.10 and Scheme 4.4) and L-FMAU (Figure 4.5). Both compounds were generously provided by Dr. C. K. Chu (College of Pharmacy, The University of Georgia, Athens, Georgia) for these studies. The fairly low relative phosphorylation rate (rPR) of 4.34 (20.3 %) compared with a rPR of 100% for dThd indicated that this compound was not as well tolerated by hTK1 as a substrate as L-FMAU (rPR = 67%) (Figure 4.13).

Interestingly, when 4.34 and L-FMAU were evaluated in phosphoryltransfer assays with Bacillus Anthracis thymidine kinase (BaTK), the obtained rPRs were 100.6 % and 91.9 %, respectively (Figure 4.13). These initial phosphorylation studies identified 4.34 as an excellent BaTK specific substrate since it was about 5 times less effectively phosphorylated by hTK1.

Very preliminary hTK1 phosphoryl transfer assays with 4.35, 4.36, 4.75, and 4.76 (Schemes 4.4 and 4.9) indicated that the β-D-nucleoside (4.36) in this series is probably the best hTK1 substrate.
4.10.2. Molecular biology studies (carried out in Dr. Baiocchi’s laboratory at the Ohio State University Comprehensive Cancer Center).

4.10.2.1. Expression and purification of hTK1

In order to carry out phosphoryl transfers assay with the target compounds described in Sections 4.9.4 and 4.9.6 at The Ohio State University, two batches of purified recombinant hTK1 were generously provided Dr. Eriksson. Unfortunately, activity measurements with both batches using $[^3]H$ dThd (Chapter 5, Section 5.5.1) did not produce $[^3]H$ dThd-monophosphate. In addition, phosphoryl transfer assay with dThd and $^{32}$P-ATP (Chapter 5, Section 5.5.2) did not generate detectable amounts of the $^{32}$P-
labeled dThd monophosphate. The outcome of both experiments led us to the conclusion that the sensitive hTK1 protein had lost its activity during the transport from Sweden to the US. Therefore, we decided to express and purify hTK1 in our own laboratory using an hTK1 expressing *E.coli* strain, which was generously provided by Dr. Eriksson.

![Image of gel electrophoresis](image)

**Figure 4.14:** The plasmid vector digestion using Bgl II and BamH1 enzymes for the confirmation of hTK1 gene insertion. M: Molecular weight marker; CP: Undigested (unsequenced) plasmid; DP: Digested (sequenced plasmid) band at 705 bp is an indication of the correct hTK1 gene insertion.

The transfection of the hTK1 cDNA into the plasmid vectors had been performed in Dr. Eriksson’s laboratory as described previously.\(^{20,39}\) The bacterial host BL21 (DE3) pLysS with transformed vector pET-14b + hTK1 was incubated on a LB plate. A single
colony of the bacteria was transferred to fresh LB media (10 mL). The presence of hTK1 gene in the bacterial colony was confirmed by sequencing (digesting) the plasmid using BamH1 and Bgl II enzymes and examination of the digested plasmid by gel electrophoresis (Figure 4.14).

**Figure 4.15**: Western blot of hTK1 expression showing that the induction is at its peak at a concentration of 0.2 mM IPTG after 6 h of induction.

A single bacterial colony from the same batch was incubated in LB media (20 mL) with subsequent induction of the hTK1 protein using varying concentrations of isopropyl β-D-1-thiogalactopyranoside (IPTG) as the inducing agent (0.1 mM to 0.4 mM) for 6 h. Protein formation was confirmed by western blot using a anti-His antibody (GE Healthcare, NJ) (Figure 4.15).

### 4.10.2.2. Expression and purification of EBVTK

The DNA sequence of EBV thymidine kinase flanked by Bgl II and EcoRI was cloned into the pBAD expression vector (pBAD-EBVTK) in Dr.Baiocchi’s laboratory. This vector was transformed into competent Top 10 *E. coli* cells and plated onto a LB agar plate with 100 µg/mL of Ampicillin. A single colony of the bacteria was transferred
to fresh LB media (10 mL) and the plasmid was purified. The insertion of the EBVTK gene in the bacterial vector was verified by sequencing. The same single colony of pBAD-EBVTK was used for the optimization of the expression of EBVTK. The amount of induction of EBVTK by the bacterial colonies was optimized by using varying concentrations of the inducing agent L-arabinose (L-ARA), ranging from 0.00002 % to 0.08 % of at 30°C for 4 h, 6 h, 10 h and overnight (> 12 h) [Figure 4.16]. The cells were then lysed by sonication at 150 s.

![Western blot of EBVTK expression showing that the induction is at its peak at 0.004 % of L-ARA after 8 h of induction.](image)

**Figure 4.16:** Western blot of EBVTK expression showing that the induction is at its peak at 0.004 % of L-ARA after 8 h of induction.

The EBVTK protein was purified by applying the crude lysate to a His-trap chelating column using Fast Protein Liquid Chromatography (FPLC). EBVTK (230 mAU) was eluted at 200 mM imidazole concentration. The SDS-PAGE gel of FPLC-purified EBVTK demonstrated enrichment of EBVTK (69 KDa) after His-trap purification. Protein formation was confirmed by western blot using an anti-His antibody (GE Healthcare, NJ) (see Figure 4.17)
**Figure 4.17**: SDS-PAGE of EBVTK after FPLC purification showing the enrichment of EBVTK (69 KDa) in the bands T10 and T9. M: Molecular weight marker, CP: Unpurified (crude) protein fraction; FT: last FPLC column fraction.

### 4.11. Future work

Future work in this area includes the expression and purification of hTK1 and EBVTK in larger quantities to evaluate the substrate characteristics of the target compounds described in **Sections 4.9.4 and 4.9.6**. Since L-I-OddU already proved to be effective in selectively inhibiting EBV DNA replication in EBV producing H1 cells,\textsuperscript{329} it would be interesting to evaluate L-I-OddU *in vitro* with EBV infected Akata- (latency I),
LCL- (B95-8, latency III) and Raji cells (latency III) after lytic induction. This could be followed by an in vivo evaluation of L-I-OddU in an animal model of EBV associated lymphoproliferative disease (Hu-PBL-SCID model) after induction of the lytic phase by any of the inducing agents described in Section 4.5. All of these cell culture and animal systems are available in Dr. Baiocchi’s laboratory.

Scheme 4.10: Proposed synthetic strategies for 4.41.

Reagents: a) MeI, K$_2$CO$_3$, DMF:acetone (1:1); b) I$_2$, ceric ammonium nitrate, acetonitrile.

As already mentioned in Sections 4.9.4 and 4.9.6, the synthesis of D-dioxolane uracil 4.29 A/B (Scheme 4.3) failed, which prevented the synthesis of target compounds 4.41.
and 4.81 (Figure 4.12). Both agents will be synthesized according to the synthetic strategies shown in Scheme 4.10 at the example of 4.41. 247,360

In addition, the separation of the anomeric mixtures 4.38 A / B and 4.79 A / B by semi-preparative HPLC will be attempted since their separation was not possible by regular silica gel column chromatography.

Since compound 4.34 (Figure 4.10 and Scheme 4.4) was identified as a selective substrate of BaTK, it would be interesting to evaluate all of the target compounds described in Sections 4.9.4 and 4.9.6 in enzymes assays with BaTK. Anthrax is a potential agent for use as a biological weapon in bio-terrorism and there is a growing need to identify agents against this deadly infectious disease. 361,362
CHAPTER 5

Experimental Section
5.1. Computational studies

5.1.1. Hardware and software


2) Dell Optiplex GX 270 desktop computer/workstation with Red Hat Enterprise Linux operating system: AutoDock Tools (ADT), Autogrid 4, Autodock 4 (Molecular Graphics Laboratory, The Scripps Research Institute, La Jolla, CA).

3) Dell optiplex 745 desktop computer/workstation with windows operating system: FlexX 3.1 (BioSolveIT GmbH, St. Augustin, Germany).

4) AMD Opteron 275 2.2GHz dual core processor desktop workstation with OpenSUSE 10.1 operating system: GLIDE v5.0 in Maestro 8.5 environment (Schrodinger LLC, Portland OR).

5) Silicon graphics O2 workstation: Sybyl 7.1 (Tripos Inc., St Louis, MO).

6) AMD opteron processor-team HPC cluster (M & A technology Inc., Carrolton, TX): Gaussian 03 (Gaussian Inc. Pittsburgh, PA) Molecular optimization and charge calculations with Gaussian as well as all docking jobs with Autodock 4 were carried out on this cluster.
5.1.2. Ligand Preparation (Chapter 2)

All Gaussian generated mol2 files of the ligands were aligned with the ligand coordinates of 1 C-prot and 2 N-prot using Sybyl before ligand preparation. Autodock does not provide parameters that recognize boron atoms. Therefore, the boron atoms were changed to “C” by modifying the pdbqt (text) files. In the case of FlexX and Surfex, boron atoms were changed to “C.3” and “C”, respectively. FlexX automatically converts boron atoms into “Du” (dummy) atoms if they are not changed to “C.3” (or “Du”) by the user. Surfex does not have “Du” atom parameters and it recognizes both boron and Du atoms as “funky atoms” and changes their atom type to “C”. Glide v 5. (OPLS2001) does not have boron and “Du” atom parameters and all the borons were replaced with “C.3” (OPLS2005 supports sp.2- and sp.3 boron, but not hexavalent boron). Following software related modifications were done as per the individual software requirements.

1) Autodock: All files generated for 1 C-prot, 2 N-prot, 1 C-crys, 2 N-crys, 1 C-con, and 2 N-con, as described in Section 2.2.4, were used for docking. All non polar hydrogens were merged before saving the file into pdbqt format. Bonds of the carborane cages that were recognized as rotatable by AutoTors in ADT were changed to non-rotatable bonds. The total number of torsions (TORSDOF) for the written output files of all ligands was set to 2.

Boron parameter implementation in AutoDock:

The van der Waals radius for desolvated boron (B3) and the associated van der Waals well-depth value were adapted from the literature. The default AutoDock atomic solvation value for carbon was used. The values used in the AutoDock parameter files were as follows:
Rii = 3.96 (Å), epsii = 0.034 (kcal/mol), volume = 32.5281 (Å^3), solpar = -0.00143, Rij-hb = 0.0, epsij_hb = 0.0, hbond = 0, rec_index = -1, map index = -1, bond index = 0

After implementing these values in the AutoDock parameter file, both AutoDock and AutoGrid were recompiled in order to reflect these changes.

2) FlexX: 1 C-crys, 2 N-crys, 1 C-con, and 2 N-con were used for docking. FlexX automatically assigned formal charges on the ligands when they were imported into the FlexX environment. For docking in “user-specified” mode, either the 2, 4-diamino-5-methyl pyrimidine portion (referred to as pyrimidyl portion [PP] throughout Chapter 2), the 1, 2-closo-dicarbadodecaboran-1-yl portion, or 7, 8-nido-dicarbaundecaboran-7-yl) portion were selected as a “base fragments”. The latter two are referred to as carboranyl portions (CPs) throughout Chapter 2.

3) Surflex: 1 C-crys, 2 N-crys, 1 C-con, and 2 N-con were used for docking. The input files generated for FlexX docking were also used for Surflex docking. As in the case of FlexX, Surflex operates with formal charges. For docking using the fragment placement method within the Surflex environment, PPs- or CPs were prepared from 2 N-crys and 2 N-con with HyperChem and imported into Surflex.

4) Glide: 1 C-crys, 2 N-crys, 1 C-con, and 2 N-con were used for docking. The input files generated for FlexX docking were used for Glide docking.

5.1.3. Protein Preparation (Chapter 2)

For docking with Autdock, FlexX, and Surflex, the protein structures in pdb format were prepared with the structure preparation tool of Sybyl. Monomers were separated from both crystal structures and the blocking groups AMI and CXC were added to the N-
and C-termini’s, respectively, for neutralization. Water molecules were removed, hydrogen atoms were added, and side chain amides and side chains bumps were fixed. In the case of Autodock, pdb files along with the NADPH (Nicotinamide adenine dinucleotide phosphate) cofactor were imported into the ADT environment, atom types were assigned, and Gasteiger charges were added. For FlexX, the NADPH cofactor was removed from the pdb files and saved as a separate mol2 file before defining the active site. The receptor description files (rdf) for docking were created with Sybyl using both hDHFR crystal structures. The amino acids within a 6.5 Å radius of 1 C-prot and 2 N-prot were selected to define the active site. The pdb, rdf, and NADPH mol2 files were then imported into the FlexX environment. For Surflex, pdb files containing NADPH were imported into the Surflex environment. Protomol files were generated using 1 C-/2 N-crys and 1 C-/2 N-con separately.

In the case of Glide, protein coordinates were pre-processed for docking using the Protein Preparation Wizard provided in the Schrodinger Maestro environment. Hydrogen atoms were added and the bond orders were assigned after deleting the monomer B as well as water molecules. Assignment of protonation states was carried out followed by hydrogen bond optimization for hydroxyl groups as well as Asn, Gln, and His residues. The hydrogen atoms were then minimized with the OPLS2001 force field. Grid calculations were performed for the protein residues of the active sites with 1 C-prot and 2 N-prot coordinates as the center with default box size (Grid box center x,y,z coordinates = 2.285766, 29.832586, -3.303069; Grid box cube size in x,y,z direction = 21.816994 Å).
5.1.4. Docking conditions/options (Chapter 2)

1) Autodock: The *Lamarckian genetic algorithm* (LGA) was selected for ligand conformational search. The docking area was defined using Autogrid 4. A 40 x 40 x 40 3-D affinity grid centered around the antifolate binding site with 0.375 Å spacing was calculated for each of the following atom types: (a) protein: A (aromatic C), C, HD, N, NA, OA, SA; (b) ligand: C, A, HD, N, NA, e (electrostatic), and d (desolvation). Additional docking parameters were: Dockings trials: 100, Population size: 100, random starting position and conformation, translation step ranges: 2.0 Å, rotation step ranges: 50, elitism: 1, mutation rate: 0.02, crossover rate: 0.8, local search rate: 0.06, and energy evaluations: 250,000. Higher energy evaluations (2,500,000) or higher population sizes (250) did not alter docking performances significantly. Final docked conformations were clustered using a tolerance of 2Å root-mean-square deviations (RMSD).

2) FlexX: Docking was performed in command line mode. The cofactor NADPH was read using a separate mol2 file as a part of the active site. Automated docking proceeded with automated base fragment selection followed by placement algorithm 3 (PA3). User-specified base fragment selection was followed by PAs 1, 2 or 3\textsuperscript{96-98} and incremental build up of the entire ligand.

3) Surflex: Default-, multistart 5-, and fragment placement (only 2 N-crys and 2 N-con) modes were explored.

4) Glide: Final calculations were performed with the Standard-Precision (SP) rigid docking protocol. Ten thousand poses were kept in the initial phase of the docking keeping the default scoring window cutoff level 100. Ligand van der Waals radii were scaled to a factor of 0.80 for non-polar atoms with partial charge cut-off level 0.15
(absolute value). A maximum of 100 conformations with the best binding energies was retained for the final analysis while discarding poses with less than 0.01 Å RMSD and 0.01 Å atomic displacement as duplicates.

5.1.5. Generation of the EBVTK homology model and docking studies with this 3D-structure (Chapter 4)

The homology model of EBVTK based on EHV4TK (PDB ID # 1P6X/ EQ1) as a template was prepared in alignment mode using SWISS-MODEL. The obtained homology model was saved in pdb format and hydrogens were added using AutoDock Tools (ADT, MGLTools, CA). After addition of hydrogen, the pdb file was imported into the Surflex environment for the protomol generation. The Multistart 5 mode was chosen for docking. The obtained docking poses were visualized with PyMoL (DeLano Scientific LLC, CA). (see Chapter 2 and Sections 5.1.1-5.1.3 for more information on Surflex, ligand preparation, and protein preparation)

5.2. General synthetic, chromatographic, and analytical procedures

Compounds 4.2, 4.7, 4.8, 4.33, 4.34, 4.35, 4.36, 4.42, 4.43, 4.73, 4.74, 4.75, 4.76, 4.82, 4.83 were synthesized according to previously reported procedures and their 1H NMR data were found to be consistent with those previously reported.

1H, 13C NMR, and 11B NMR spectra were obtained on a Bruker 400 MHz FT-NMRs in The Ohio State University Campus Chemical Instrumentation Center at The Ohio State University College of Pharmacy. Chemical shifts (δ) are reported in ppm from an internal
dichloromethane, acetone or an external BF$_3$:Et$_2$O standard. Coupling constants are reported in Hz. High Resolution – Electron Spray ionization (HR-ESI) mass spectra were obtained at The Ohio State University Campus Chemical Instrumentation Center on a Micromass Q-Tof II- and a Micromass LCT spectrometer. Electron Impact (EI) mass spectra were obtained at the University of Illinois Mass Spectrometry Laboratory (Urbana- Champaign, Illinois) on a Micromass 70-VSE Double Focusing Sector spectrometer. The 70-VSE is high resolution (>50,000), extended geometry, 8 kV mass spectrometer with GC and an effective mass range of 1-6,000 Da. The data system is the TSSPro3.0 system (run on a PC under Windows XP) by Shrader Analytical & Consulting Laboratories, Inc. HR-ESI mass spectra were also obtained at the University of Illinois Mass Spectrometry Laboratory on a Waters Q-Tof Ultima Tandem Quadrupole/Time-of-Flight spectrometer. For all carborane-containing compounds, the mass of the most intensive peak of the isotopic pattern was reported for a 80% boron-11 to 20% boron-10 distribution. Measured patterns agreed with calculated patterns.

Silica gel 60 (0.063 -0.200 mm) from Merck was used for gravity column chromatography whereas silica gel 60 (0.015-0.049 mm) from EM science was used for flash column chromatograph. Reagent-grade solvents were used for column chromatography. Precoated glass-backed TLC plates with silica gel 60 F254 (0.25-mm layer thickness) from Dynamic adsorbents (Norcross, GA) were used for TLC. General compound visualization for TLC was achieved by UV light and I$_2$ vapor. Carbohydrate-containing compounds were selectively visualized by spraying the plate with 1% H$_2$SO$_4$ and heating at 120 °C. Carborane-containing compounds were selectively visualized by spraying the plate with a 0.06% PdCl$_2$/1% HCl solution and heating at 120 °C, which
caused the slow (15-45 s) formation of a gray spot due to the reduction of Pd$^{2+}$ to Pd$_0$. RadioTLC was visualized using AMBIS Imaging Hardware/Software (AMBIS Image Acquisition and Analysis, Version 4.0; AMBIS, San Diego, CA).

Anhydrous solvents such as dimethylformamide, acetonitrile, dichloromethane, and dichloroethane were purchased directly either from Acros Organics (Morris plains, NJ) or from EMD (Gibbstown, NJ), or from Sigma Aldrich (Milwaukee, WI). Other chemicals and solvents were purchased from standard commercial suppliers. Tetrahydrofuran (THF) was distilled from sodium and benzophenone indicator under argon. Unless specified otherwise, all reactions were carried out under Argon atmosphere. Na$^{125}$I was obtained from Perkin Elmer, (Waltham, MA) as an alkaline (NaOH) solution (pH 8-11) with a specific activity of ca. 17 Ci (629 GBq)/mg. Hermann’s catalyst ($trans$-$di$-$m$-$acetatobis$[2$-$$(di$-$o$-$tolylphosphino)benzyl]$dipalladium) (II) was purchased from Alfa Aesar Ltd. (Ward Hill, MA) 5-Iodouracil, 5-bromovinyl uracil, and trifluoromethylsilyl trifluoromethanesulfonate (TMSOTf) were purchased from Sigma Aldrich (Milwaukee, WI).

For all non-radioactive materials, semi-preparative HPLC purification was performed either with a Gemini 5u C18 column (21.20 mm x 250 mm, 5 μm particle size) supplied by Phenomenex Inc. CA, USA or with a Supelco Discovery® HS C18 column (10 mm x 250 mm, 10 μm particle size) supplied by Sigma Aldrich, MO, USA, on a Hitachi HPLC system (L-2130) with a Windows based data acquisition and Hitachi Diode array detector (L-2455). A LiChrocart® 250-4 HPLC cartridge packed with LiChrospher® RP-18 stationary phase (4 mm x 250 mm, 5 μm particle size) supplied by EM Science, NJ, USA
was used for analytical reversed-phase chromatography on a Hitachi HPLC system (L-2130) with a Windows based data acquisition and Hitachi Diode array detector (L-2455).

For all radioactive materials, semi-preparative HPLC purification was performed with a Supelco Discovery® HS C18 column (10 mm x 250 mm, 10 µm particle size). A Beckman Ultrasphere® column (4.6 mm x 250 mm, 5 µm particle size) supplied by Beckmann Coulter Inc., CA, USA, was used for analytical reversed-phase chromatography. A Beckmann HPLC system with a Windows based data acquisition and Beckman System Gold UV detector module (166) and Radiomatic FLO-ONE detector series A-500 was used for data analysis.

5.3. Detailed synthetic procedures and compound characterizations

![Chemical structure]

3-[5-(9-Iodo-o-carboran-1-yl) pentyl] thymidine and 3-[5-(12-iodo-o-carboran-1-yl) pentyl] thymidine (3.2 a / 3.2 b, [N5-I]) [Strategy 1]

To a stirred solution of 3-[5-(o-carboran-1-yl) pentyl] thymidine (3.1, N5) (0.055 mmol, 25 mg) and AlCl₃ (0.55 mmol, 73 mg) in anhydrous CH₂Cl₂, ICl (1M in dichloromethane) was added dropwise (0.055 mmol, 55 µL) at 0°C. The reaction mixture was stirred for 2 h at the same temperature and then quenched with a saturated solution of
Na$_2$S$_2$O$_3$. The reaction mixture was extracted with dichloromethane, the organic layer separated, and the aqueous layer washed 2 x with 10 mL dichloromethane. The combined organic layers were washed with brine, dried over magnesium sulfate, and evaporated in vacuo. The product was crudely separated from deglycosylated nucleoside (nucleobase) and inorganic impurities (AlCl$_3$, aluminum hydroxide etc.) by silica gel column chromatography. $R_f$ 0.34 (dichloromethane: methanol, 10:1). Final purification, primarily by separation from di-iodinated product, was accomplished by preparative HPLC (acetonitrile/water 60:40) to furnish 5 mg (15%) of 3.2 a / 3.2 b as a white foam.

$^1$H-NMR (CD$_3$COCD$_3$) $\delta$. $^1$H-NMR (400 MHz), 1.32 (m, 2H, CH$_2$), 1.57 (m, 4H, CH$_2$), 1.83 (s, 3H, CH$_3$), 2.25 -2.39 (m, 4H, CH$_2$-C$_{carborane}$ and H-2’), 3.78 (d, 2H, H-5’, $J$ = 2.7 Hz), 3.86 (m, 2H, CH$_2$-N), 3.93 (m, 1H, H-3’), 4.26 (m, 1H, -OH), 4.41 (d, 1H, H-4’, $J$ = 3.2 Hz), 4.49 (s, 1H, -OH), 4.90 (s, 1H, H-C$_{carborane}$), 5.10 (s, 1H, H-C$_{carborane}$), 6.34 (t, 1H, H-1’, $J$ = 6.8 Hz), 7.83 (s, 1H, H-6). $^{13}$C NMR (CD$_3$COCD$_3$) $\delta$ 13.37, 26.77, 27.74, 37.68, 37.87, 41.13, 41.26, 62.85, 64.41, 72.15, 78.31, 86.46, 88.69, 109.88, 135.58, 151.73, 163.79. $^{11}$B {$^1$H} NMR (CD$_3$COCD$_3$) $\delta$. -17.79 (s, 1B), -16.15 (s, 1B), -11.04 (m, 12B), -7.19 (s, 2B), -3.96 (s, 1B), -0.58 (s, 1B). $^{11}$B NMR (CD$_3$COCD$_3$) $\delta$ -17.93 (s, 1B), -16.26 (s, 1B), -11.17 (m, 12B), -7.33 (d, 4B, $J$ = 156.2 Hz), -4.10 (d, 1B, $J$ = 145.7 Hz), -0.70 (d, 1B, $J$ = 153.5 Hz). HPLC retention time = 9.91 min (Lichrosphere®100, 5 µm with 1 mL flow rate, solvent system: Acetonitrile: water, 60:40, isocratic elution). MS (HR-ESI) C$_{17}$H$_{33}$B$_{10}$IN$_2$O$_5$ (M+ Na)$^+$ calcd 603.2324, found 603.2344.
Diiodinated 3.1 (N5-2I)

HPLC retention time = 11.21 min (RP-18 analytical HPLC column (Lichrosphere®100, 5 µm with 1 mL flow rate, solvent system: Acetonitrile: water, 60:40, isocratic elution). MS (HR-ESI), C_{17}H_{32}B_{10}I_{2}N_{2}O_{5} (M + Na)⁺ calcd 729.1259, found 729.1310.

Triiodinated 3.1 (N5-3I)

HPLC retention time = 13.60 min (Lichrosphere®100, 5 µm with 1 mL flow rate, solvent system: Acetonitrile: water, 60:40, isocratic elution). MS (HR-ESI) C_{17}H_{31}B_{10}I_{3}N_{2}O_{5} (M + Na)⁺ calcd 855.0226, found 855.0277.

3-[5-(8, 9, 10, 12-Tetra-iodo-o-carboran-1-yl) pentyl] thymidine (3.3, N5-I4)

To a stirred solution of 3-[5-(o-carboran-1-yl) pentyl] thymidine (3.1, N5) (0.22 mmol, 100 mg) and AlCl₃ (3.3 mmol, 440 mg) in anhydrous dichloromethane, ICl (1.1 mmol, 1.1 mL of a 1M in dichloromethane) was added dropwise at 0°C. The reaction mixture was stirred for 2 h at the same temperature then allowed to warm to rt. This was followed by two days of reflux. The reaction mixture was quenched with a saturated solution of Na₂S₂O₃ and then extracted with dichloromethane. The organic layer was separated and...
the aqueous layer was then washed 2 x with 10 mL of dichloromethane. The combined organic layers were washed with brine, dried over anhydrous MgSO4, and evaporated in vacuo. The product was crudely separated from deglycosylated nucleoside (nucleobase) and inorganic impurities (AlCl3, aluminum hydroxide etc.) by silica gel column chromatography. Rf 0.34 (dichloromethane: methanol, 10:1). Final purification, primarily by separation from tri-iodinated product was accomplished by preparative HPLC (acetonitrile/water 60:40, isocratic elution) to furnish 27 mg (20%) of N5-I4 (3.3) as a yellow solid. HPLC retention time = 13.81 min (RP-18 analytical HPLC column, Lichrosphere®100, 5 μm with 1 mL flow rate, solvent system: Acetonitrile: water, 60:40, isocratic elution). 1H-NMR (CD3COCD3) δ 1.33 (m, 2H, CH2), 1.61 (m, 4H, CH2), 1.83 (s, 3H, CH3), 2.23 (dd, 2H, H-2', J = 4.7Hz, J = 6.7 Hz,), 2.45 (m, 2H, CH2-Ccarborane), 3.78 (m, 2H, H-6), 3.88 (m, 2H, CH2-N), 3.94 (dd, 1H, H-3', J = 3.1Hz, J = 6.1 Hz), 4.27 (s, 1H, -OH), 4.42 (s, 1H, -OH), 4.49 (s, 1H, H-4'), 5.72 (s, 1H, H-Ccarborane), 6.34 (t, 1H, H-1', J = 6.8 Hz), 7.84 (s, 1H, H-6). 13C NMR (CD3COCD3) δ 13.27, 25.74, 26.56, 27.54, 37.23, 40.96, 41.16, 63.14, 63.67, 72.05, 77.66, 86.36, 88.59, 109.77, 135.50, 151.63, 163.70. 11B {1H} NMR (CD3COCD3) δ -17.84 (s, 4B, B8, B9, B10 and B12), -12.08 (s, 1B), -10.22, -8.35 (m, 10B) 11B NMR (CD3COCD3) δ -17.87 (s, 4B, B8, B9, B10 and B12), -12.09 (d, 1B, J = 205.2 Hz), -10.12, -8.33 (m, 5B). MS (HR-ESI) C17H30B10I4N2O5 (M + Na)+ calcd 977.9346, found 977.9319.

Nucleobases (3.5) [presumably di- and triiodinated at the carborane cluster]

Rf 0.72 (dichloromethane: methanol, 15:1). 1H-NMR (CD3COCD3) δ 1.34 (m, 4H, CH2), 1.59 (m, 4H, CH2), 1.83 (s, 3H, CH3), 2.37 (m, 2H, CH2-Ccarborane), 3.82 (m, 2H, N-
CH$_2$), 5.50 (br, 1H, H-C$_{carborane}$), 5.71 (br, 1H, H-C$_{carborane}$), 7.28 (s, 1H, H-6), 9.71 (s, 1H, NH). MS (HR-ESI) C$_{12}$H$_{22}$B$_{10}$I$_4$N$_2$O$_2$ (M$^+$ Na$^+$) calcd 841.8761, found 864.8494 and MS (HR-ESI) C$_{12}$H$_{22}$B$_{10}$I$_4$N$_2$O$_2$ (M$^+$ Na$^+$) calcd 613.0828, found 613.0753.

9-Iodo-o-carborane (3.6)

The amount of 2 g of o-carborane (13.8 mmol) was dissolved in anhydrous dichloromethane and ICl (27.7 mmol, 27.7 mL of a 1M in dichloromethane) was added dropwise at 25 °C. The reaction mixture was stirred for 5h at 40 °C, quenched with a saturated solution of Na$_2$S$_2$O$_3$ and extracted with dichloromethane. The organic layer was washed with brine, dried over anhydrous MgSO$_4$, and evaporated in vacuo. The residue was purified by silica gel column chromatography affording 2.7 g of 3.6 (72%). $R_f$ 0.38 (hexanes: ethyl acetate, 10:4). $^1$H-NMR (CD$_3$COCD$_3$) δ 4.74 (s, 1H, H-C$_{carborane}$), 4.95 (s, 1H, H-C$_{carborane}$). $^{13}$C NMR (CD$_3$COCD$_3$) δ 53.12, 57.77. $^{11}$B {^1}H NMR (CD$_3$COCD$_3$) δ -16.52 (s, 1B, B9), -13.67 (s, 2B, B3, B6), -12.81 (s, 2B, B4, B5), -12.07 (s, 2B, B7, B11), -7.30 (s, 2B, B8, B10), -1.12 (s, 1B, B12). $^{11}$B NMR (CD$_3$COCD$_3$) δ -16.53 (s, 1B), -13.06 (m, 6B), -7.32 (d, 2B, $J = 155.0$ Hz), -1.15 (d, 1B, $J = 151.2$ Hz). MS (HR-EI) C$_2$B$_{10}$H$_9$I (M$^+$) calcd 270.09090, found 270.09116.
1-(tert-Butyldimethylsilyl)-9-iodo-o-carborane and 1-(tert-butyldimethylsilyl)-12-iodo-o-carborane (3.7 a / 3.7 b)

To a stirred solution of 3.6 (2.1 g, 7.77 mmol) in anhydrous THF at -78 ºC was added dropwise a 2.5 M solution of n-BuLi in hexanes (3.26 mL, 3.26 mmol). The reaction mixture was allowed to stir for 30 min while being warmed to rt. The solution was then cooled to 0 ºC and tert-butyldimethylsilyl chloride (1.29 g, 8.55 mmol) in 15 mL of anhydrous THF was added dropwise but rapidly. The solution was refluxed overnight, carefully quenched with 10 mL of water, and extracted with 60 mL of diethyl ether. The layers were separated and the aqueous layer was extracted again with 2 x 30 mL of diethyl ether. The combined organic layers were dried over anhydrous MgSO₄ and concentrated in vacuo. The crude residue was purified by silica gel column chromatography affording 2 g of 3.7 a / 3.7 b (66 %) Rf 0.42 (pentane: diethyl ether 25:1) ¹H-NMR (CD₃COCD₃) δ 0.30 (s, 6H, -Si(CH₃)₂), 0.33 (s, 6H, -Si(CH₃)₂), 1.04 (s, 9H, -C(CH₃)₃), 1.07 (s, 9H, -C(CH₃)₃), 4.52 (s, 1H, H-C-carborane), 4.74 (s, 1H, H-C-carborane).

¹³C-NMR (CD₃COCD₃) δ -4.43, 20.02, 27.22, 58.68, 63.02. ¹¹B ¹H NMR (CD₃COCD₃) δ -15.84 (s, 1B, B12-I), -14.88 (s, 1B, B9-I), -11.28 (m, 12B, B3, B4, B5, B6, B7, B11), -5.65 (s, 4B, B8, B10), -0.88 (s, 1B, B12-H), 1.33 (s, 1B, B9-H). ¹¹B NMR (CD₃COCD₃), -15.93 (s, 1B), -14.96 (s, 1B), -11.41 (m, 12B), -5.74 (d, 4B, J = 150.8 Hz),
-0.98 (d, 1H, $J = 150.4$ Hz), 1.24 (d, 1H, $J = 150.7$ Hz). MS (HR-El) $\text{C}_8\text{H}_{25}\text{B}_{10}\text{Si} (\text{M})^+$ calcd 384.17738, found 384.17738.

\[ \text{I} \quad \text{Si} \]

1-(tert-Butyldimethylsilyl)-9-iodo-o-carborane (3.7 a)

A quantity of 250 mg of a mixture of compounds 3.7 a and 3.7 b were separated by silica gel flash chromatography. $R_f$ 0.47 (pentane: diethyl ether, 25:1) $^1$H-NMR (CD$_3$COCD$_3$) $\delta$ 0.33 (s, 6H, -Si(CH$_3$)$_2$), 1.09 (s, 9H, -C(CH$_3$)$_3$), 4.76 (s, 1H, H-C$_{\text{carborane}}$).

\[ \text{I} \quad \text{Si} \]

1-(tert-Butyldimethylsilyl)-12-iodo-o-carborane (3.7 b)

$R_f$ 0.36 (pentane: diethyl ether, 25:1). $^1$H-NMR (CD$_3$COCD$_3$) $\delta$ 0.31 (s, 6H, -Si(CH$_3$)$_2$), 1.04 (s, 9H, -C(CH$_3$)$_3$), 4.55 (s, 1H, H-C$_{\text{carborane}}$).
5-[(9-Iodo-o-(2-tert-butyldimethylsilyl)-carboran-1-yl)-iodopentane] and 5-[(12-Iodo-o-(2-tert-butyldimethylsilyl)-carboran-1-yl)-iodopentane (3.8 a / 3.8 b)

To a stirred solution of 3.7 a / 3.7 b (700 mg, 1.82 mmol) in 25 mL of anhydrous diethyl ether: benzene (2:1) at -78 °C was added a 2.5 M solution of n-BuLi in hexanes (0.76 mL, 0.76 mmol) dropwise. The reaction mixture was allowed to stir for 30 min while being warmed to rt, then cooled to 0 °C, and 1, 5-diodopentane (360 mg, 1.82 mmol) was added dropwise but rapidly. The solution was stirred for 24 h, quenched with 20 mL of water and extracted with 25 mL of diethyl ether. The separated aqueous layer was extracted again 2 x 30 mL of diethyl ether, the combined organic layers dried over anhydrous MgSO₄, and concentrated in vacuo. The crude residue was purified by silica gel flash chromatography affording 70 mg of 3.8 a / 3.8 b (6 %) as a brownish oil. Rf 0.72 (pentane: diethyl ether 30:1). ¹H-NMR (CD₃COCD₃) δ 0.40 (s, 6H, -Si(CH₃)₂), 0.44 (s, 6H, -Si(CH₃)₂), 1.09 (s, 9H, -C(CH₃)₃), 1.12 (s, 9H, -C(CH₃)₃), 1.44 (m, 4H, CH₂), 1.61 (m, 4H, CH₂), 1.84 (m, 4H, CH₂), 2.25 (m, 2H, CH₂-C₉carborane), 2.37 (m, 2H, CH₂-C₉carborane), 3.28 (dd, 4H, CH₂-I, J = 7.1 Hz, J = 14.4 Hz). ¹³C NMR (CD₃COCD₃) δ -2.39, 7.27, 20.93, 21.00, 27.78, 33.74, 38.02, 71.85, 79.39. ¹¹B {¹H} NMR (CD₃COCD₃) δ -15.96 (s, 1B, B12-I/B9-I), -14.33 (s, 1B, B9-I/B12-I), -9.21 (m, 12B, B3, B4, B5, B6, B7,
B(11), -5.25 (m, 4B, B8, B10), -2.41 (s, 1B, B12-H/B9-H), 2.16 (s, 1B, B9-H/B12-H). $^{11}$B 
NMR (CD$_3$COCD$_3$) $\delta$ -15.97 (s, 1B), -14.33 (s, 1B), -9.16 (m, 14B), -5.32 (dd, 2B, $J = 42.2$ Hz, $J = 154.7$ Hz), -2.42 (d, 1B, $J = 159.7$ Hz), 2.09 (d, 1B, $J = 147.6$ Hz). MS (HR- 
EI) C$_{13}$H$_{34}$B$_{10}$I$_2$Si (M)$^+$ calcd 580.1523, found 580.1521.

3-{5-[9-Iodo-o-(2-tert-butyldimethylsilyl)-carboran-1-yl] pentyl} thymidine and 3-{5-
[12-iodo-o-(2-tert-butyldimethylsilyl)-carboran-1-yl] pentyl} thymidine (3.9 a / 3.9 b)

To a solution of compound 3.8 a / 3.8 b (1.98 g, 4.66 mmol) in 40 mL of 
DMF/acetone (1:1) were added thymidine (2.82 g, 11.64 mmol) and potassium carbonate 
(2.57 g, 18.59 mmol). The solution was stirred for 2 hours at 40 °C filtered, and the 
filtrate was concentrated in vacuo. The residue was purified by silica gel column 
chromatography using dichloromethane/methanol (10:1) as the eluent to give compound 
3.9 a / 3.9 b in 1.21 g yield (52%). $R_f$ 0.51 (dichloromethane: methanol, 10:1). $^1$H-NMR 
(CD$_3$COCD$_3$) $\delta$ 0.39 (s, 6H, -Si(CH$_3$)$_2$), 0.43 (s, 6H, -Si(CH$_3$)$_2$), 1.08 (s, 9H, -C(CH$_3$)$_3$), 
1.11 (s, 9H, -C(CH$_3$)$_3$), 1.35 (m, 4H, CH$_2$), 1.64 (m, 4H, CH$_2$), 1.84 (s, 6H, CH$_3$), 2.25 (m, 
4H, H-2’), 2.34 (m, 4H, -CH$_2$-C$_{carborane}$), 3.74 (m, 4H, H-5’), 3.86 (t, 4H, CH$_2$-N, H-4’, $J$ 
= 7.2 Hz), 3.94 (dd, 1H, H-3’, $J = 2.8$ Hz, $J = 5.8$ Hz), 4.26 (br, 1H, H-C$_{carborane}$), 4.44 (br,
1H, H-C_carborane), 4.47 (m, 2H, H-4'), 6.34 (t, 2H, H-1', J = 6.8 Hz), 7.83 (s, 1H, H-6), 7.96 (s, 1H, H-6). \textsuperscript{13}C-NMR (CD\textsubscript{3}COCD\textsubscript{3}) δ -2.27, 13.4, 21.05, 26.90, 27.86, 30.81, 36.22, 38.25, 41.10, 41.29, 62.85, 71.87, 72.14, 83.88, 86.44, 86.69, 109.87, 135.56, 151.73, 163.32. \textsuperscript{11}B \{\textsuperscript{1}H\} NMR (CD\textsubscript{3}COCD\textsubscript{3}) δ -15.98 (s, 1B, B\textsubscript{12}-I/B\textsubscript{9}-I), -14.28 (s, 1B, B\textsubscript{9}-I/B\textsubscript{12}-I), -9.94 (m, 12B, B\textsubscript{3}, B\textsubscript{4}, B\textsubscript{5}, B\textsubscript{6}, B\textsubscript{7}, B\textsubscript{11}), -5.64 (s, 2B, B\textsubscript{8}, B\textsubscript{10}), -2.49 (s, 1B, B\textsubscript{12}-H/B\textsubscript{9}-H), 2.09 (s, 1B, B\textsubscript{9}-H/B\textsubscript{12}-H). \textsuperscript{11}B NMR (CD\textsubscript{3}COCD\textsubscript{3}) δ -16.14 (s, 1B), -14.42 (s, 1B), -9.80 (m, 12B), -5.75 (d, 4B, J = 153.4 Hz), -2.62 (d, 2B, J = 163.1 Hz), 1.93 (d, 2B, J = 139.8 Hz). MS (HR-ESI) C\textsubscript{23}H\textsubscript{47}B\textsubscript{10}IN\textsubscript{2}O\textsubscript{5}Si (M+H)\textsuperscript{+} calcd 695.3394, found 695.3391.

tert-Butyl{5-[9-iodo-o-(2-tert-butyldimethylsilyl)-carboran-1-yl]pentyloxy} diphenylsilane and tert-butyl {5-[12-iodo-o-(2-tert-butyldimethylsilyl)-carboran-1-yl]pentyloxy} diphenylsilane (3.10 a / 3.10 b)

Compounds 3.10 a / 3.10 b were synthesized and purified in 70% yield (1.6 g) from 1.24 g (3.87 mmol) of 3.7 a / 3.7 b by adapting the procedure described for the synthesis of 3.8 a / 3.8 b using 1.60 g (3.22 mol) of 1-(tert-butyldiphenylsilyloxy)-5-pentyl tosylate instead of 1,5-dioiiodoantane as the starting material. R\textsubscript{f} 0.8 (pentane: diethyl ether, 30:1). \textsuperscript{1}H-NMR (CD\textsubscript{3}COCD\textsubscript{3}) δ 0.38 (s, 6H, -Si(CH\textsubscript{3})\textsubscript{3}), 0.42 (s, 6H, -Si(CH\textsubscript{3})\textsubscript{2}), 1.041 (s, 9H, -C(CH\textsubscript{3})\textsubscript{3}), 1.09 (s, 9H, -C(CH\textsubscript{3})\textsubscript{3}), 1.43-1.65 (m, 12H, CH\textsubscript{2}-CH\textsubscript{2}-CH\textsubscript{2}), 2.21 (m, 2H,
5-(9-Iodo-o-carboran-1-yl)-1-pentanol and 5-(12-iodo-o-carboran-1-yl)-1-pentanol (3.12 a / 3.12 b)

To a solution of 3.14 a / 3.14 b (2.37 g, 4.05 mmol) in 40 mL of THF was added dropwise at -78 °C, 6.1 mL of a 1.0 M solution of TBAF in THF (6.1 mmol). The reaction mixture was stirred at 4 °C for 30 min, acidified by addition of 10% methanolic-HCl (20 mL), and then stirred for an additional 30 min at 4 °C. The reaction mixture was extracted with ethyl acetate (3 x 35 mL); the combined organic layers washed with water and then saturated NaHCO₃, dried over anhydrous magnesium sulfate, and concentrated in vacuo. The residue was purified by column chromatography on silica gel (hexanes: ethyl acetate, 3:7) to furnish 1.1 g of 3.12 a / 3.12 b (76%). Rf, 0.54 (hexanes: ethyl acetate, 3:7); ¹H-NMR (CD₃COCD₃) δ 1.28 -1.59 (m, 12H, CH₂-CH₂-CH₂), 2.25 (m, 2H, CH₂-C(carborane), 2.38 (m, 2H, CH₂-C(carborane), 3.50 (dd, 4H, CH₂-OH, J = 5.3 Hz, J = 10.2 Hz), 4.88 (br, 1H, H-C(carborane), 5.09 (br, 1H, H-C(carborane). ¹³C-NMR (CD₃COCD₃) δ 26.03, 33.05, 37.92, 38.10, 59.59, 62.02, 64.39. ¹¹B {¹H} NMR (CD₃COCD₃) δ -18.39 (s, 1B,
B12-I/B9-I), -16.68 (s, 1B, B9-I/B12-I), -11.71 (m, 12B, B3, B4, B5, B6, B7, B11), -7.68 (m, 4B, B8, B10), -4.42 (s, 1B, B12-H/B9-H), -1.15 (s, 1B, B9-H/B12-H). $^{11}$B NMR (CD$_3$COCD$_3$) δ -18.07 (s, 1B), -16.36 (s, 1B), -11.33 (m, 14B), -7.32 (m, 2B, $J = 152.5$ Hz), -4.11 (d, 1B, $J = 153.3$ Hz), -0.69 (d, 1B, $J = 150.4$ Hz). MS (HR-EI) C$_7$H$_{21}$B$_{10}$IO (M)$^+$ calcd 356.16407, found 356.16467.

$^\text{TsO}$\longrightarrow\text{OTBDMS}$

5-(tert-Butyldimethylsilyloxy) pentyl-4-methylbenzenesulfonate (3.13) $^{192,193}$

The amount of 9.0 g (41.2 mmol) of 5-(tert-butyldimethylsilyloxy) pentanol, triethylamine (17.22 mL, 123.61 mmol), and DMAP (0.8 g, 8.24 mmol) were dissolved in 120 mL of dichloromethane. A solution of 11.78 g (61.78 mmol) of $p$-toluenesulfonyl chloride in 40 mL of dichloromethane was added dropwise at rt. The reaction mixture was stirred at rt for 1 h and then washed with a saturated sodium bicarbonate solution. The organic phase was separated, dried over anhydrous MgSO$_4$ and evaporated in vacuo. The residue was purified by silica gel column chromatography, giving 11.50 g of light yellowish oil (75%). $R_f$ 0.55 (dichloromethane: hexanes, 5:1). $^1$H-NMR (CDCl$_3$) δ 0.01 (s, 6H, -Si(CH$_3$)$_2$), 0.86 (s, 9H, -Si(CH$_3$)$_3$), 1.34 (m, 2H, CH$_2$), 1.45 (m, 2H, CH$_2$), 1.65 (m, 2H, CH$_2$), 2.43 (s, 3H, CH$_3$), 3.55 (t, 2H, CH$_2$-OTBDMS, $J = 6.2$ Hz), 4.01 (t, 2H, CH$_2$-OTs, $J = 6.5$ Hz), 7.33 (d, 2H, Ar-H, $J = 8.1$ Hz), 7.78 (d, 2H, Ar-H, $J = 8.2$Hz). $^{13}$C NMR (CDCl$_3$) δ -5.38, 18.27, 21.59, 21.74, 25.89, 28.57, 31.99, 62.69, 70.54, 127.83, 129.77, 133.13, 144.60.
tert-Butyl-\{5-[9-iodo-o-(2-tert-butylidimethylsilyl)-carboran-1-yl] pentyloxy\} dimethylsilane and tert-butyl-\{5-[12-iodo-o-(2-tert-butylidimethylsilyl)-carboran-1-yl pentyloxy\} dimethylsilane (3.14 a / 3.14 b)

Compounds (3.14 a / 3.14 b) were synthesized and purified in 66% yield (6.5 g) from 6.52 g (16.96 mmol) of 3.7 a / 3.7 b by adapting the procedure described for the synthesis of 3.8 a / 3.8 b (except for the use of THF as a solvent and reflux conditions) using 8.21 g (22.03 mmol) of 5-(tert-butylidimethylsilyloxy) penty-4-methylbenzenesulfonate (3.13) instead of 1, 5-dioiodoantane as a starting material. \(R_f\) 0.72 (pentane: diethyl ether, 30:1).

\(^1\text{H}-\text{NMR}\) (CD\(_3\)COCD\(_3\)) \(\delta\) 0.05 (s, 6H, -Si(CH\(_3\))\(_2\)), 0.05 (s, 6H, -Si(CH\(_3\))\(_2\)), 0.39 (s, 6H, -Si(CH\(_3\))\(_2\)), 0.43 (s, 6H, -Si(CH\(_3\))\(_2\)), 0.89 (s, 9H, -C(CH\(_3\))\(_3\)), 0.90 (s, 9H, -C(CH\(_3\))\(_3\)), 1.09 (s, 9H, -C(CH\(_3\))\(_3\)), 1.12 (s, 9H, -C(CH\(_3\))\(_3\)), 1.34 - 1.67 (m, 12H, CH\(_2\)-CH\(_2\)-CH\(_2\)), 2.22 (m, 2H, CH\(_2\)-C\(_{\text{carborane}}\)), 2.34 (m, 2H, CH\(_2\)-C\(_{\text{carborane}}\)), 3.63 (dd, 4H, -CH\(_2\)-OTBDMS, \(J = 6.2\) Hz, \(J = 13.2\) Hz). \(^{13}\text{C}-\text{NMR}\) (CD\(_3\)COCD\(_3\)) \(\delta\) -5.06, -2.22, 18.85, 21.04, 26.16, 26.37, 27.89, 30.89, 33.04, 38.38, 63.22, 79.50, 83.72. \(^{11}\text{B}\) \(^{1}\text{H}\)-NMR (CD\(_3\)COCD\(_3\)) \(\delta\) -15.88 (s, 1B, B12-I/B9-I), -14.20 (s, 1B, B9-I/B12-I), -9.02 (m, 12B, B3, B4, B5, B6, B7, B11), -5.33 (m, 4B, B8, B10), -2.47 (s, 1B, B12-H/B9-H), 2.03 (s, 1B, B9-H/B12-H). \(^{11}\text{B}\) NMR (CD\(_3\)COCD\(_3\)) \(\delta\) -15.90 (s, 1B), -14.20 (s, 1B), -9.22 (m, 12B), -5.32 (m, 4B), -2.44 (d, 1B, \(J = 148.6\)Hz), 2.07 (d, 1B, \(J = 148.8\)Hz). MS (HR-EI) C\(_{19}\)H\(_{49}\)B\(_{10}\)OS\(_2\) (M-56) calcd 528.27443, found 528.27603.
5-(9-Iodo-o-carboran-1-yl)-1-pentyl tosylate and 5-(12-iodo-o-carboran-1-yl)-1-pentyl tosylate (3.17 a / 3.17 b)

To a solution of 3.12 a / 3.12 b (1.65 g, 4.63 mmol) in dichloromethane (12 mL) was added triethylamine (0.84 mL, 6.02 mmol) and a catalytic amount of 4-(dimethylamino)pyridine (DMAP) (113 mg, 0.93 mmol). A solution of p-toluenesulfonyl chloride (1.33 g, 6.95 mmol) in 13 mL of dichloromethane was added at 0°C and the reaction mixture was stirred at rt for 7 h. The reaction was quenched with a saturated NH₄Cl aqueous solution; the organic phase was washed with water and brine, dried over anhydrous magnesium sulfate, and concentrated in vacuo. The residue was purified by silica gel column chromatography to afford 2.2 g (70%) of compounds 3.17 a / 3.17 b as an oil. R_f 0.65 (hexanes: ethyl acetate, 1:1) ¹H-NMR (CD₃COCD₃) δ 1.27-1.69 (m, 12H, CH₂-CH₂-CH₂), 2.23 (m, 2H, CH₂-C-carborane), 2.36 (m, 2H, CH₂-C-carborane), 2.47 (s, 6H, CH₃), 4.02 (m, 4H, CH₂-OTs), 4.84 (br, 1H, H-C-carborane), 5.05 (br, 1H, H-C-carborane), 7.49 (d, 4H, Ar-H, J = 7.3 Hz), 7.79 (dd, 4H, Ar-H, J = 3.0 Hz, J = 8.4 Hz); ¹³C-NMR (CD₃COCD₃) δ 21.64, 25.31, 28.87, 29.24, 37.63, 59.53, 64.33, 71.09, 128.64, 130.90, 134.31, 145.82. ¹¹B {¹H} NMR (CD₃COCD₃) δ -17.89 (s, 1B, B12-I/B9-I), -16.25 (s, 1B, B9-I/B12-I), -11.44 (m, 12B, B3, B4, B5, B6, B7, B11), -7.08 (s, 4B, B8, B10), -3.79 (s, 1B, B12-H/B9-H), -0.51 (s, 1B, B9-H/B12-H). ¹¹B-NMR (CD₃COCD₃) δ -17.98 (s, 1B), -16.31 (s, 1B), -11.54 (m,
12B), -7.29 (d, 2B, $J = 155.0 \text{ Hz}$), -4.09 (d, 1B, $J = 153.3 \text{ Hz}$), -0.67 (d, 1B, $J = 150.4 \text{ Hz}$).

MS (HR-EI) C$_{14}$H$_{27}$O$_3$SIB$_{10}$ (M)$^+$ calcd 510.17292, found 510.17239.

![Chemical structure](image)

3-[5-(9-Iodo-o-carboran-1-yl) pentyl] thymidine and 3-[5-(12-iodo-o-carboran-1-yl) pentyl] thymidine (3.2 a / 3.2 b/ N5-I) [Strategy 2]

To a solution of 3.17 a and 3.17 b (1.18 g, 2.31 mmol) in DMF/acetone (1:1, 40 mL) were added thymidine (1.68 g, 6.93 mmol) and potassium carbonate (0.96 g, 6.94 mmol). The solution was stirred for 2.5 h at 40 °C and filtered. The filtrate was concentrated in vacuo, the residue added to water, and extracted ethyl acetate (3 x 50 mL). The combined organic layers were washed with brine, dried over anhydrous MgSO$_4$, and concentrated in vacuo. The residue was purified by silica gel column chromatography using dichloromethane/methanol (10:1) as the eluent to give compound 3.2 a / 3.2 b (0.8 g, 60%).
9-Bromo-o-carborane (3.18)

To a solution of 3.6 (100 mg, 0.37 mmol) in 3 mL dichloromethane was added a solution of NaBr (381 mg, 3.7 mmol) in 5 mL of water. The solvents were evaporated in vacuo and a solution of Hermann’s Catalyst (10 mol %, 170 mg) in 1 mL anhydrous DMF was added to the dry residue. The reaction mixture was then stirred at 110 °C for 1 h. The reaction mixture was evaporated, the residue added to water and extracted with dichloromethane. The organic phase was dried over anhydrous MgSO₄, concentrated in vacuo, and the residue purified by silica gel column chromatography. Rf 0.38 (hexanes:ethyl acetate, 10:4) affording 40 mg (49 %) of 3.18. ¹H-NMR (CDCl₃) δ 3.64 (s, 1H, H-C(carborane)). ¹³C NMR (CD₂COCD₃) δ 46.76, 53.24. ¹¹B {¹H} NMR (CDCl₃) δ -13.67 (m, 6B, B3, B4, B5, B6, B7, B11), -8.00 (s, 2B, B8, B10), -1.69 (s, 1B, B12), -0.56 (s, 1B, B9-Br). ¹¹B NMR (CDCl₃) -13.75 (m, 6B), -7.89 (d, 2B, J = 154.1 Hz), -1.83 (d, 1B, J = 152.1 Hz), -0.37 (s, 1B). MS (HR-EI) C₂B₁₀H₁₁Br (M) † calcd 222.1047, found 222.1048.
5-[9-Iodo-o-(2-tert-butyldimethylsilyl)-carboran-1-yl] pentane and 5-[12-iodo-o-(2-tert-butyldimethylsilyl)-carboran-1-yl] pentane (3.19 a / 3.19 b)

Compounds 3.19 a / 3.19 b were synthesized and purified in 55% yield (3.04 g) from 4.7 g (12.22 mmol) of 3.7 a / 3.7 b by adapting the procedure described for the synthesis of 3.8 a / 3.8 b (except for the use of THF as a solvent and reflux conditions) using 2.56 mL (19.56 mmol) of 1-iodopentane instead of 1, 5-diiodoantane as a starting material. Rf 0.69 (hexanes: ethyl acetate, 20:1). $^1$H NMR (CD$_3$COCD$_3$) δ 0.39 (s, 6H, -Si(CH$_3$)$_2$), 0.43 (s, 6H, -Si(CH$_3$)$_2$), 0.89 (dd, 6H, CH$_3$, $J = 6.8$ Hz, $J = 13.6$ Hz), 1.09 (s, 9H, -C(CH$_3$)$_3$), 1.12 (s, 9H, -C(CH$_3$)$_3$), 1.29-1.62 (m, 12H, CH$_2$-CH$_2$-CH$_2$), 2.20 (m, 2H, CH$_2$-C$_{carborane}$), 2.32 (m, 2H, CH$_2$-C$_{carborane}$). $^{13}$C NMR (CD$_3$COCD$_3$) δ -2.26, 14.25, 21.02, 22.96, 27.86, 30.85, 31.84, 38.23, 78.81, 83.78. $^{11}$B{$^1$H} NMR (CDCl$_3$) δ -16.73 (s, 1B, B12-I/ B9-I), -14.95 (s, 1B, B9-I/ B12-I), -10.284 (m, 12 B), -6.03 (d, 4B, B8, B10), -3.11(s, 1B, B12-H/B9-H), -2.22 (s,1B, B9-H/B12-H), 1.47 (s,1B, B9-H/B12-H). $^{11}$B NMR (CDCl$_3$) δ. -16.75 (s, 1B), -15.02 (s, 1B), -10.35 (m, 8B), -5.94 (d, 4B, $J = 141.1$ Hz), -3.22 (d, 2B, $J = 175.1$Hz), 1.31 (d, 2B, $J = 147.7$ Hz). MS (HR-EI) C$_{13}$H$_{35}$B$_{10}$Si (M)$^+$ calcd 454.25563, found 454.25545.
5-(9-Iodo-o-carboran-1-yl) pentane and 5-(12-iodo-o-carboran-1-yl) pentane (3.20 a / 3.20 b)

A stirred solution of 3.19 a / 3.19 b (1.61g, 3.54 mmol) in anhydrous THF (20 mL) was cooled to -78 °C and a 1.0 M solution of TBAF in THF (4.25 mL, 4.25 mmol) was added dropwise. The mixture was allowed to stir for 30 min at 0 °C followed by addition of 10 mL of water. The mixture solution was extracted 3 x 30 mL of diethyl ether, the combined organic layers were dried over anhydrous MgSO₄, and evaporated in vacuo. The residue was purified by silica gel column chromatography to furnish 800 mg of a brownish oil (66 %). $R_f$ 0.52 (hexanes: ethyl acetate, 9:1). $^1$H NMR (CD₃COCD₃) δ 0.89 (m, 6H), 1.29 (m, 8H), 1.54 (m, 4H), 2.24 (m, 2H), 2.37 (m, 2H), 4.86 (br, 1H, H-C_carborane), 5.06 (br, 1H, H-C_carborane). $^{13}$C NMR (CD₃COCD₃) δ 14.15, 22.86, 31.71, 37.87, 38.05, 59.63, 64.44. $^{11}$B{$^1$H}NMR (CD₃COCD₃) δ -17.93 (s, 1B, B12-I/B9-I), -16.26 (s,1B, B9-I/B12-I), -12.24—10.44 (m, 12B, B3, B4, B5, B6, B7, B11), -7.23 (d, 4B, B8, B10), -4.07 (s, 1B, B12-H/B9-H), -0.69 (s,1B, B9-H/B12-H) $^{11}$B NMR (CD₃COCD₃) δ -17.93, -16.26, -11.37 (m, 12B), -7.25 (d, 4B, $J = 153.2$ Hz), -4.12 (d, 2B, $J = 151.5$ Hz), -0.76 (d, 2B, $J = 149.7$ Hz). MS (HR-EI) C₇H₂₁B₁₀I (M) $^+$ calcd 340.16915, found 340.16891.
5-(9-Bromo-o-carboran-1-yl) pentane and 5-(12-bromo-o-carboran-1-yl) pentane
(3.21 a / 3.21 b)

Compounds 3.21 a / 3.21 b were synthesized and purified in 63% yield (54 mg) from 100 mg (0.29 mmol) 3.20 a / 3.20 b by adapting the procedure described for the synthesis of 3.18. $R_f$: 0.52 (hexanes: ethyl acetate, 9:1). $^1$H NMR (CD$_3$COCD$_3$) $\delta$ 0.88 (m, 6H), 1.28 (m, 8H), 1.56 (m, 4H), 2.35 (m, 4H), 4.83 (s, 2H, H-C$_{carborane}$). $^{13}$C NMR (CD$_3$COCD$_3$) $\delta$ 14.74, 23.44, 32.30, 37.52, 38.60, 6.18, 63.27. $^{11}$B NMR (CD$_3$COCD$_3$) $\delta$ -12.20 (m, 12B), -8.11 (d, 1H, $J$ = 149.5Hz), -4.90 (m, 1B), -0.69 (m, 4B); $^{11}$B {$^1$H} NMR (CD$_3$COCD$_3$) $\delta$ -11.80 (m, 12B, B3, B4, B5, B6, B7, B11), -8.07 (s, 2B, B8, B10), -4.88 (B12-H/B9-H), -0.39 (m, 4B, B12-Br/B9-Br and B9-H/B12-H); MS (HR-EI) C$_7$H$_{21}$B$_{10}$Br (M)$^+$ calcd 296.17672, found 293.17367.

3-[5-(9-Bromo-o-carboran-1-yl) pentyl] thymidine and 3-[5-(12-bromo-o-carboran-1-yl) pentyl] thymidine (3.22 a / 3.22 b - N5-Br)
Compounds 3.22 a / 3.22 b were synthesized and purified in 52% yield (48 mg) from 0.100 g (0.175 mmol) 3.2 a / 3.2 b by adapting the procedure described for the synthesis of 3.18 with the use of 25 mol % of the Hermann’s catalyst. Rf, 0.34 (dichloromethane: methanol, 10:1). ^1H-NMR (CD$_3$COCD$_3$) δ. 1.31 (m, 4H), 1.56 (m, 8H), 1.83 (s, 6H), 2.23 (m, 4H), 2.36 (m, 4H), 3.77 (m, 4H), 3.86 (m, 4H), 3.92 (m, 2H), 4.27 (s, 2H), 4.42 (s, 2H), 4.49 (d, 2H, J = 2.9 Hz), 4.84 (s, 2H), 6.34 (t, 2H, J = 6.7 Hz), 7.83 (s, 2H). ^11B ^1H NMR (CD$_3$COCD$_3$) δ -11.87 (m, 12B, B3, B4, B5, B6, B7, B11), -8.11 (s, 4B, B8, B10), -4.93 (s, 1B, B12-H/B9-H), -1.31 (s, 1B, B9-H/B12-H), 0.43 (s, 2B, B9-Br and B12-Br). ^11B NMR (CD$_3$COCD$_3$) δ -12.71 (m, 12B), -8.13 (d, 4B, J = 151.9 Hz), -4.92 (m, 2B), -0.723 (m, 2B). MS (HR-ESI) C$_{17}$H$_{33}$B$_{10}$BrN$_2$O$_5$ (M+ Na$^+$) calcd 557.2401, found 557.2431.

\[
\text{3-\{5-(9-[^{125}\text{I}]\text{Iodo-\text{o-carboran-1-yl}) penty1\}\text{thymidine and 3-\{5-(12-[^{125}\text{I}]\text{Iodo-\text{o-carboran-1-yl})penty1\}\text{thymidine (3.23 a / 3.23 b - N5-[^{125}\text{I}]\}
}
\]

An alkaline Na$^{125}$I solution (10 µL, pH 8-11) with a specific activity of 17 Ci (629 GBq)/mg was diluted to 100 µL using deionized water. Compound 3.2 a / 3.2 b (10 mg, 0.017 mmol), dissolved in 200 µL DMF, was transferred to a 3 mL conical reaction vial (MINUM-WARE ®). A solution of Hermann’s catalyst in 100 µL of DMF (10 mol %)
was added to the reaction vial. To this reaction mixture, 10 µL of the diluted alkaline Na\(^{125}\)I solution was added dropwise. The reaction mixture was then heated at 110 ºC for 1 h, cooled to rt and filtered using GHP Acrodisc ® 13 mm syringe filter to separate the catalyst from the reaction mixture. The filtrate was evaporated under the flow of argon at 30 ºC and the residue purified by semi preparative HPLC. \(R_f\) 0.28 (dichloromethane: methanol, 10:1). HPLC retention time = 25.50 min (analytical HPLC, Beckman Ultrasphere® column, 1mL flow rate, solvent system: Acetonitrile: water, 50: 50, isocratic elution).

![Chemical Structure](image)

**3-Butyl-1-(2-deoxy-β-D-erythro-pentofuranosyl) thymine (β-N3-butyldthymidine, 3.24)**

To a stirred suspension of thymidine (4 g, 16.51 mmol) and K\(_2\)CO\(_3\) (6.23 g, 45.07 mmol) in 20 mL of DMF: acetone (5:5) was added butyl bromide (2.06 g, 15.03 mmol) dropwise at rt. The reaction mixture was stirred at 50ºC overnight and then concentrated in vacuo. Water was added to the residue followed by extraction with 3 x 50 mL of ethyl acetate. The organic phases were combined, washed with brine, dried over anhydrous MgSO\(_4\), and evaporated. The product was purified by silica gel column chromatography to yield 6 g (81%) of **3.24**. \(R_f\) 0.36 (dichloromethane: methanol, 15:1). \(^1\)H NMR (CD\(_3\)OD) \(\delta\) 0.95 (t, 3H, CH\(_3\), \(J = 7.3\) Hz), 1.35 (m, 2H, CH\(_2\)), 1.57 (m, 2H, CH\(_2\)), 1.90 (s,
3H, CH₃), 2.24 (m, 2H, H-2‘), 3.76 (ddd, 2H, H-5‘ J = 3.4 Hz, J = 12.0 Hz, J = 29.6 Hz), 3.90 (t, 3H, CH₂-N and H-4‘ J = 7.3 Hz), 4.39 (m, 1H, H-3‘), 6.30 (t, 1H, H-1‘, J = 6.7 Hz), 7.83 (d, 1H, H-6 J = 0.7 Hz).

¹³C NMR (CD₃OD) δ 13.36 (C₃H₃), 14.29 (C₃H₃), 21.32 (C₃H₂), 30.88 (C₃H₂), 41.48 (C₂‘), 42.25 (CH₂-N), 62.90 (C-5‘), 72.24 (C-3‘), 87.22 (C-1‘), 88.99 (C-4‘), 110.83 (C-5), 136.55 (C-6), 152.44 (C-2), 165.54 (C-4)

3-Butylthymine (3.25)

To a solution of 25 mg of 3.24 (0.084 mmol) in 5 mL of dichloromethane was added 32 mg (0.252 mmol) of I₂. The reaction mixture was refluxed for 72 h and subsequently washed with a saturated aqueous NH₄Cl solution. The aqueous layer was extracted with 3 x 15 mL of dichloromethane; the combined organic layers were dried over anhydrous Na₂SO₄, and evaporated. The product was purified by silica gel column chromatography (dichloromethane: methanol, 25:1) to give a 3-butylthymine in 80% yield. Rₚ: 0.315 (hexanes: ethyl acetate, 3:7); ¹H NMR (DMSO-d₆) δ 0.86 (t, 3H, CH₃, J = 7.3 Hz), 1.23 (m, 2H, CH₂), 1.43 (m, 2H, CH₂), 1.70 (s, 1H, CH₃), 3.72 (t, 2H, CH₂-N, J = 7.38 Hz), 7.24 (s, 1H, C-6), ¹³C NMR (CD₃OD) δ 12.98 (CH₃), 14.19 (CH₃), 21.21 (CH₂), 30.84 (CH₂), 41.49 (CH₂-N), 110.03 (C-5), 137.39 (C-6), 153.56 (C-2), 166.48 (C-4). MS (HR-ESI) C₉H₁₄N₂O₂Na (M+Na)⁺ calcd 205.0953, found 205.0953.
3-Butyl-1-(2-deoxy-α-D-erythro-pentofuranosyl) thymine (α-N3-butylthymidine, 3.26)

Compound 3.26 was isolated in 20% (5 mg) yield as a side product in the reaction described for 3.25. $R_f$ 0.36 (dichloromethane: methanol, 15:1). $^1$H NMR (CD$_3$OD) $\delta$ 0.95 (t, 3H, CH$_3$, $J = 7.3$Hz), 1.35 (m, 2H, CH$_2$), 1.57 (m, 2H, CH$_2$), 1.90 (s, 3H, CH$_3$), 2.01 (m, 2H, H-2’), 3.74 – 3.87 (m, 5H, H-3’,H-4’and H-5’), 3.91 (m, 2H, CH$_2$-N), 4.16 (m, 1H, H-3’), 6.04 (dd, 2H, H-1’, $J = 2.5$Hz, $J = 10.8$Hz), 7.53 (s, 1H, H-6). $^{13}$C NMR (CD$_3$OD) $\delta$ 13.11 (CH$_3$).14.18 (CH$_3$), 21.21 (CH$_2$), 30.77 (CH$_2$), 37.19 (C-2’), 42.32 (CH$_2$-N), 67.28 (C-5’), 67.76 (C-3’), 68.29 (C-1’), 79.65 (C-4’), 110.99 (C-5), 136.36 (C-6), 152.15 (C-2), 16.31 (C-4). MS (HR-ESI) C$_{14}$H$_{22}$N$_2$O$_5$ (M+Na)$^+$ calcd 321.1426, found 321.1458

3-Butyl-(3’, 5’-di-O-methyl) thymidine (3.27)
To the suspension of NaH (650 mg, 16.25 mmol) in THF (5 mL) was added a solution of 3.24 (1 g, 3.35 mmol) in 5 mL THF at 0 °C. The reaction mixture was sonicated for 20 minutes at rt and methyl iodide (2.38 g, 16.76 mmol) was added dropwise. Subsequently, the reaction mixture was sonicated for 40 minutes at rt and then quenched by dropwise addition of MeOH until the release of the hydrogen gas completely stopped. The reaction mixture was evaporated in vacuo followed by addition of saturated aqueous solution of NH₄Cl and extraction with 3 x 30 mL of dichloromethane. The organic phases were dried over anhydrous MgSO₄ and evaporated in vacuo. The residue was purified by silica gel column chromatography (hexanes: ethyl acetate, 1:1 to yield 870 mg (80 %) of 3.27.

Rf 0.44 (hexanes: ethyl acetate, 1:1). ¹H NMR (CDCl₃) δ 0.91 (t, 3H, CH₃ J = 7.3 Hz), 1.34 (m, 2H, CH₂), 1.57 (m, 2H, CH₂), 1.91 (s, 3H, CH₃), 2.03 (m, 1H, H-2’), 2.40 (ddd, 1H, H-2’, J = 2.6 Hz, J = 6.0 Hz, J = 13.6 Hz), 3.33 (s, 3H, OCH₃), 3.42 (s, 3H, OCH₃), 3.61 (ddd, 2H, H-5’, J = 2.7 Hz, J = 10.5 Hz, J = 38.0 Hz), 3.91 (m, 2H, CH₂-N), 3.98 - 4.11 (m, 2H, H-3’, H-4’), 6.32 (t, 1H, H-1’, J = 6.41 Hz), 7.56 (s, 1H, H-6). ¹³C NMR (CDCl₃) δ 13.86 (CH₃), 14.16 (CH₃), 20.62 (CH₂), 30.10 (CH₂), 37.76 (C-2’), 41.53 (CH₂-N), 57.27 (CH₃-O), 59.50 (CH₃-O), 73.38 (C-5’), 81.52 (C-3’), 84.04 (C-1’), 86.14 (C-4’), 110.30 (C-5), 134.18 (C-6), 151.30 (C-2), 163.88 (C-4). MS (HR-ESI) C₁₆H₁₆N₂O₅ (M+ Na) + calcd 349.1739, found 349.1724.
3-Butyl-(3′, 5′-di-O-carbobenzyloxycarbonyl) thymidine (3.28)

To the solution of 3.24 (1.26 g, 3.86 mmol) and DMAP (1.41 g, 11.54 mmol) in 10 mL of dichloromethane was added benzyl chloroformate (1.98 g, 11.58 mmol) dropwise at 0 °C. The resulting suspension was stirred at rt for 17 h and the progress of the reaction was monitored by TLC analysis. Following completion of the reaction, the solution was diluted with dichloromethane and the organic layer was extracted with 2 M HCl. The organic phase was separated, dried with MgSO₄, and evaporated in vacuo. The product was purified by silica gel chromatography (hexanes: ethyl acetate, 4:1) to furnish 3.28 in 1.64 gm (75%) yield. *R*ₐ 0.18 (hexanes: ethyl acetate, 4:1). ¹H NMR (CDCl₃) δ 0.92 (t, 3H, CH₃, J = 7.3 Hz), 1.35 (m, 2H, CH₂), 1.57 (m, 2H, CH₂), 1.78 (s, 3H, CH₃), 2.23 (m, 1H, H-2′), 2.49 (ddd, 1H, H-2′, J = 1.5 Hz, J = 5.6 Hz, J = 14.3 Hz), 3.90 (m, 2H, H-5′), 4.30 (d, 1H, H-4′, J = 2.3 Hz), 4.43 (m, 2H, CH₂-N), 5.20 (m, 5H, CH₂-Ph, H-3′), 6.43 (dd, 1H, H-1′, J = 5.8 Hz, J = 8.5 Hz), 7.30 (s, 1H, C-6), 7.37 (s, 10H). ¹³C NMR (CDCl₃) δ 13.14, 13.71, 20.14, 29.58, 37.31, 41.19, 67.23, 70.20, 70.27, 77.75, 81.62, 85.09, 110.79, 128.39, 128.57, 128.67, 128.79, 128.89, 132.69, 134.50, 150.77, 154.30, 154.48, 163.14. MS (HR-ESI) C₃₀H₃₄N₂O₉ (M+ Na)⁺ calcd 589.2162, found 589.2172.
3-Butyluridine (3.29)

Compound 3.29 was synthesized from uridine (2 g, 8.19 mmol) by adopting the procedure described for the synthesis of 3.24 (2.1 g, 86 %). Rf: 0.26 (dichloromethane:methanol, 15:1).\(^1\)H NMR (CD\(_3\)OD) \(\delta\) 0.95 (t, 3H, CH\(_3\), \(J = 7.3\) Hz), 1.36 (m, 2H, CH\(_2\)), 1.59 (m, 2H, CH\(_2\)), 3.80 (ddd, 1H, H-5', \(J = 2.8\) Hz, \(J = 12.3\) Hz, \(J = 47.3\) Hz), 3.91 (m, 2H, CH\(_2\)-N), 4.00 -4.18 (m, 3H, H-2', H-3', H-4'), 5.76 (d, 1H, H-5, \(J = 8.1\) Hz), 5.92 (d, 1H, H-1', \(J = 3.8\) Hz), 8.02 (d, 1H, H-6, \(J = 8.1\) Hz). \(^1\)C NMR (CD\(_3\)OD) \(\delta\) 14.12 (CH\(_3\)), 21.15 (CH\(_2\)), 30.67 (CH\(_2\)), 41.91 (CH\(_2\)-N), 62.11 (C-5'), 71.09 (C-3'), 75.87 (C-2'), 86.23(C-4'), 91.56 (C-1'), 102.00 (C-5), 140.79 (C-6), 152.54 (C-2), 165.04 (C-4). MS (HR-ESI) C\(_{13}\)H\(_{20}\)N\(_2\)O\(_6\) (M+ Na)\(^+\) calcd 323.1219, found 323.1214.

3-Butyl-(2', 3', 5'-tri-O-methyl) uridine (3.30)
Compound \textbf{3.30} was synthesized from \textbf{3.29} (650 mg, 2.16 mmol) by adopting the procedure described for the synthesis of \textbf{3.27} (600 mg, 81 %). \( R_f \) 0.31 (hexanes: ethyl acetate, 2:1). \(^1\)H NMR (CDCl\(_3\)) \( \delta \) 0.92 (t, 3H, CH\(_3\), \( J = 7.3 \) Hz), 1.35 (m, 2H, CH\(_2\)), 1.58 (m, 2H, CH\(_2\)), 3.42 (s, 6H, CH\(_2\)-O), 3.55-3.81 (m, 2H, H-5’), 3.61 (s, 3H, CH\(_2\)-O), 3.90 (m, 4H, H-2’, H-4’, CH\(_2\)-N), 4.16 (dd, 1H, H-3’, \( J = 2.2 \) Hz, \( J = 5.1 \) Hz), 5.71 (d, 1H, H-5, \( J = 8.1 \) Hz), 5.92 (s, 1H, H-1’), 7.89 (d, 1H, H-6, \( J = 8.1 \) Hz). \(^{13}\)C NMR (CDCl\(_3\)) \( \delta \) 13.74 (C\(\text{H}_3\)), 20.15 (C\(\text{H}_2\)), 29.60 (C\(\text{H}_2\)), 40.81(C\(\text{H}_2\)), 58.16 (C\(\text{H}_3\)-O), 58.35 (C\(\text{H}_3\)-O), 59.17 (C\(\text{H}_3\)-O), 70.23 (C-5’), 76.37 (C-3’), 80.69 (C-2’), 81.88 (C-4’), 88.29 (C-1’), 101.41 (C-5), 137.62 (C-6), 150.69 (C-2), 162.67 (C-4) MS (HR-ESI) C\(_{16}\)H\(_{26}\)N\(_2\)O\(_6\) (M+ Na) \(^+\) calcd 365.1689, found 365.1674

\begin{figure}
\centering
\includegraphics[width=0.2\textwidth]{3-Butyl-5-iodouridine.png}
\caption{3-Butyl-5-iodouridine (3.31)}
\end{figure}

To a solution of \textbf{3.29} (250 mg, 0.83 mmol) in 5 mL acetonitrile was added dropwise a 1M solution of ICl in dichloromethane (0.8 mL, 0.8 mmol). The reaction mixture was refluxed for 48 h, concentrated in vacuo, and the residue was purified by silica gel column chromatography (dichloromethane: methanol, 15:1) to give a \textbf{3.31} (280 gm, 80 %). \( R_f \) 0.32 (dichloromethane: methanol, 15:1). \(^1\)H NMR (CDCl\(_3\)) \( \delta \) 0.96 (t, 3H, CH\(_3\), \( J = 7.4 \) Hz), 1.36 (m, 2H, CH\(_2\)), 1.59 (m, 2H, CH\(_2\)), 3.83 (ddd, 2H, H-5’, \( J = 2.4 \) Hz, \( J = 4.8 \) Hz, \( J = 11 \) Hz), 4.15 (dd, 1H, H-3’, \( J = 2.2 \) Hz, \( J = 5.1 \) Hz), 5.71 (d, 1H, H-5, \( J = 8.1 \) Hz), 5.92 (s, 1H, H-1’), 7.89 (d, 1H, H-6, \( J = 8.1 \) Hz).
12.2 Hz, $J = 60.8$ Hz), 3.96 (m, 2H, CH$_2$-N), 4.02-4.18 (m, 3H, H-2, H-3, H-4), 5.88 (d, 1H, H-1, $J = 3.3$Hz), 8.65 (s, 1H, H-6). $^{13}$C NMR (CDCl$_3$) $\delta$ 14.15, 21.16, 30.62, 43.57, 61.52, 67.63, 70.68, 76.27, 86.20, 91.93, 145.65, 152.16, 161.87. MS (HR-ESI) C$_{13}$H$_{19}$IN$_2$O$_6$ (M+ Na)$^+$ calcd 449.0186, found 449.0170.

![Image of the compound](image.png)

**3-Butyl-(2′, 3′, 5′-tri-O-methyl)-5-iodouridine (3.32)**

Compound 3.32 was synthesized from compound 3.30 (310 mg, 0.90 mmol) by adopting the procedure described for the synthesis of 3.31 (320 mg, 89%). $R_f$ 0.58 (hexanes: ethyl acetate, 2:1). $^1$H NMR (CDCl$_3$) $\delta$ 0.93 (t, 3H, CH$_3$, $J = 7.3$ Hz), 1.35 (m, 2H, CH$_2$), 1.59 (m, 2H, CH$_2$), 3.43 (s, 3H, O-CH$_3$), 3.52 (s, 3H, O-CH$_3$), 3.62 (m, 4H, O-CH$_3$, H-5′), 3.81-4.20 (m, 6H, H-5′, H-4′, H-3′, H-2′, CH$_2$-N), 5.90 (s, 1H, H-1′), 8.66 (s, 1H, H-6). $^{13}$C NMR (CDCl$_3$) $\delta$ 14.15 (CH$_3$), 20.57 (CH$_2$), 29.95 (CH$_2$), 43.00 (CH$_2$-N), 58.67 (CH$_3$-O), 58.83 (CH$_3$-O), 59.72 (CH$_3$-O), 67.83 (C5-I), 70.11 (C-5′), 76.30 (C-3′), 81.53 (C-2′), 82.37 (C-4′), 88.82 (C-1′), 143.73 (C-6), 150.75 (C-2), 160.04 (C-4). MS (HR-ESI) C$_{16}$H$_{25}$IN$_2$O$_6$ (M+ Na)$^+$ calcd 491.0655, found 491.0656.
3-Methylthymine (4.3)

To a solution of 4.2 (3.6 g, 14.04 mmol) in 80 mL of acetonitrile was added 3.59 g (14.14 mmol) of I₂. The reaction mixture was refluxed for 35 h and then concentrated in vacuo. The quantity of 100 mL of methanol was added to the dark residue and the resulting solution was filtered and concentrated. The product was purified by column chromatography over silica gel (dichloromethane: methanol, 10:1) to afford 1.2 g of 4.3 (~61%). Rf 0.57 (dichloromethane: methanol; 10:1). ¹H-NMR (CDCl₃) δ 1.94 (s, 3H, CH₃), 3.34 (s, 3H, CH₃), 7.06 (s, 1H, H-6), 10.12 (s, 1H, NH). ¹³C NMR (CDCl₃) δ 12.99 (CH₃), 27.32 (CH₃), 110.05 (C-5), 134.21 (C-6), 153.41 (C-2), 164.25 (C-4). MS (HR-ESI) C₆H₈N₂O₂ (M + H)⁺ calcd 141.0664, found 141.0652.

3-Ethylthymine (4.10)

To a suspension of 60% NaH (290 mg, 12.08 mmol) in DMF was added compound 4.7 (500 mg, 2.21 mmol) at 0 ºC. The reaction mixture was stirred at 0 ºC for 30 min after which ethyl bromide (290 mg, 2.66 mmol) was added. The reaction mixture was then allowed to reach rt followed by the additional stirring for additional 2 h. The
reaction mixture was then quenched by dropwise addition of cold water until the effervescence ceased. The reaction mixture was then concentrated in vacuo, the white-yellowish semisolid residue was diluted with 20 mL methanol, and K$_2$CO$_3$ (0.35 g, 2.53 mmol) was added. The reaction mixture was then stirred for additional 2 h at room temperature and then concentrated in vacuo. The residue was purified by silica gel column chromatography to obtain 4.10 (240 mg, 83 %). $R_f$ 0.59 (dichloromethane: methanol; 10:1). $^1$H-NMR (DMSO) $\delta$ 1.06 (t, 3H, $J = 7.0$ Hz, CH$_3$), 1.76 (m, 3H, CH$_3$), 3.79 (q, 2H, $J = 7.0$ Hz, CH$_2$), 7.28 (s, 1H, H-5), 10.86 (s, 1H, NH). $^{13}$C NMR (DMSO) $\delta$ 12.38, 12.71, 34.58, 107.13, 136.12, 150.96, 163.49. MS (HR-ESI) C$_7$H$_{10}$N$_2$O$_2$ (M+ Na)$^+$ calcd 177.064, found 177.0631.

![3-Methyluracil](image)

**3-Methyluracil (4.11)**

Compound 4.11 was synthesized in 70 % yield (1.05 g) according to the procedure described for the synthesis of 4.10 using 2.5 g (11.83 mmol) of 4.8 and 2 g (14.09 mmol) of methyl iodide as starting materials. $R_f$ 0.53 (dichloromethane: methanol; 10:1) $^1$H-NMR (DMSO) $\delta$ 3.11 (s, 3H, CH$_3$), 5.58 (d, 1H, $J = 7.6$ Hz, H-5), 7.43 (d, 1H, $J = 7.6$ Hz, H-6), 11.12 (s, 1H, NH). $^{13}$C NMR (DMSO) $\delta$ 26.40, 99.55, 140.38, 151.60, 163.31. MS (HR-ESI) C$_5$H$_6$N$_2$O$_2$ (M+ Na)$^+$ calcd 149.0327, found 149.0319.
5-Iodo-3-methyluracil (4.12)

To a solution of di-tert-butyl-dicarbonate [(tert-BOC)₂O] (2.29 g, 10.50 mmol) and 5-I-uracil (2.5 g, 10.50 mmol) in MeCN (50 mL) was added 4-DMAP (13 mg, 0.106 mmol). The reaction mixture was stirred for 3.5 h at rt and then concentrated in vacuo. The yellowish-white residue (4.9, 3.5 g, 96 % crude yield) was used without further purification methylation reaction, as described in the following. \(^1\)H-NMR (DMSO) \(\delta\) 2.49 (s, 1H, CH₃), 3.16 (s, 9H, (CH₃)₃), 7.92 (s, 1H, H-5), 11.44 (s, 1H, NH). \(^{13}\)C NMR (DMSO) \(\delta\) 28.01, 66.67, 71.40, 145.13, 151.19, 160.68.

To a suspension of 60% NaH (0.29 g, 12.08 mmol) in DMF was added crude 4.9 (3.5 g, 10.11 mmol) at 0 ºC. The reaction mixture was stirred at 0 ºC for 30 min followed by addition of methyl iodide (1.72 g, 12.13 mmol). The reaction mixture was allowed to warm up to rt over a 30 min period and then concentrated in vacuo. Approximately 30 mL of methanol was added to the resulting oily white suspension until the effervescence ceased. Apparently, the Boc-protective was removed in this step. The reaction mixture was then neutralized (pH 7) using 0.5 N HCl. The thick yellowish precipitate was extracted with EtOAc (3 x 50 mL), the organic extracts were washed with brine (100 mL), dried over anhydrous MgSO₄, and concentrated in vacuo. The product was co-evaporated with 2 x 30 mL of ethanol to obtain 2.5 g (75 %) of pure 4.12. \(R_f\) 0.6 (dichloromethane: methanol; 10:1) \(^1\)H-NMR (DMSO) \(\delta\) 3.17 (s, 1H, CH₃), 7.94 (s, 1H,
H-5), 11.45 (s, 1H, NH). $^{13}$C NMR (DMSO) δ 27.94, 66.53, 145.05, 151.11, 160.60. MS (HR-ESI) $\text{C}_5\text{H}_5\text{IN}_2\text{O}_2$ (M+ Na)$^+$ calcd 274.9293, found 274.9277.

(2R, 4S/4R)-4-Acetoxy-2-[( tert-butyldiphenylsilyl) oxy] methyl)dioxolane (4.23)

Compound 4.23 was synthesized in 10 steps from D-mannose according to the procedure described in detail in Section 4.6.2. $^1$H-NMR (CDCl$_3$) δ 1.06, 1.07 (s, 9H), 1.96, 2.10 (s, 3H, CH$_3$), 3.72-4.19 (m, 4H, H-5', H-6'), 5.25, 5.38 (t, 1H, $J = 4.1$ and 3.1 Hz each, H-2'), 6.32, 6.39 (m, 1H, H-4'), 7.41-7.70 (m, 10H, -Ph$_2$). MS (HR-ESI) $\text{C}_{22}\text{H}_{28}\text{O}_5\text{Si}$ (M+ Na)$^+$ calcd 423.1604, found 423.1608.

3-Methyl-1-[(2R, 4S)-1-[(2- (tert-butyldiphenylsilyloxy) methyl]-1, 3-dioxolan-4-yl] thymine (4.24 A) and 3-methyl-1-[(2R, 4R)-1-[(2-(tert-butyldiphenylsilyloxy) methyl]-1, 3-dioxolan-4-yl] thymine (4.24 B)

A mixture of compound 4.3 (30 mg, 0.43 mmol), a catalytic amount of (NH$_4$)$_2$SO$_4$ (~5-10 mg), and HMDS (5 mL) was refluxed overnight. The resulting clear solution was
concentrated in vacuo under anhydrous conditions to afford colorless oil. To this oily residue was added a solution of 4.23 (40 mg, 0.10 mmol) in anhydrous DCE (5 mL). To this reaction mixture was added dropwise TMSOTf (0.036 mL, 0.20 mmol). The mixture was stirred for 3 h at rt and then quenched by the addition of saturated aqueous sodium bicarbonate (10 mL). The mixture was stirred for an additional 30 min, followed by dilution with water (20 mL) and extraction with dichloromethane (3 x 30 mL). The combined organic layers were washed with saturated aqueous sodium bicarbonate solution and brine, dried with magnesium sulfate, and concentrated in vacuo. The residue was purified by column chromatography over silica gel to obtain 4.24 A in ~30 % yield (10 mg) $R_f$ 0.47 (hexanes: ethyl acetate, 2:1) $^1$H-NMR (CDCl$_3$) δ 1.07 (s, 9H, -C(CH$_3$)$_3$), 1.96 (s, 3H, CH$_3$), 3.34 (s, 3H, CH$_3$), 3.71 (m, 2H, H-6'), 4.01 (dd, 1H, $J = 2.2$ Hz, $J = 9.5$ Hz, H-5'), 4.39 (dd, 1H, $J = 5.4$ Hz, $J = 9.5$Hz, H-5'), 5.55 (t, 1H, $J = 3.1$ Hz, H-2'), 6.34 (dd, 1H, $J = 2.2$ Hz, $J = 5.3$ Hz, H-4'), 7.17 (s, 1H, H-6), 7.43 (m, 6H, -Si(Ph)$_2$), 7.69 (m, 4H, -Si(Ph)$_2$). $^{13}$C NMR (CDCl$_3$) δ 13.41, 19.12, 26.72, 27.78, 64.79, 71.06, 83.05, 106.05, 110.49, 127.82, 129.92, 132.54, 135.53, 135.60, 151.26, 163.50 MS (HR-ESI) C$_{26}$H$_{32}$N$_2$O$_5$Si (M+ Na)$^+$ calcd 503.1978, found 503.1969. Target compound 4.24 B was isolated in ~30 % yield (16.0 mg). $R_f$ 0.34 (hexanes: ethyl acetate, 2:1) $^1$H-NMR (CDCl$_3$) δ 1.06 (s, 9H, -C(CH$_3$)$_3$), 1.70 (s, 3H, CH$_3$), 3.32 (s, 3H, CH$_3$), 3.90 (m, 2H, H-6'), 4.15 (m, 2H, H-5'), 5.07 (t, 1H, $J = 2.9$ Hz, H-2'), 6.36 (dd, 1H, $J = 2.7$ Hz, $J = 4.2$ Hz, H-4'), 7.25 (d, 1H, $J = 5.8$ Hz, H-6), 7.38 (m, 6H, -Si(Ph)$_2$), 7.67 (d, 4H, $J = 7.2$ Hz, -Si(Ph)$_2$). $^{13}$C NMR (CDCl$_3$) δ 13.17, 19.30, 26.77, 27.84, 63.82, 71.17, 81.56, 105.30, 110.65, 127.81, 129.94, 132.86, 135.51, 135.55, 151.40, 163.44. MS (HR-ESI) C$_{26}$H$_{32}$N$_2$O$_5$Si (M+ Na)$^+$ calcd 503.1978, found 503.1969.
3-Ethyl-1-[(2R, 4S)-1-[(2-tert-butyldiphenylsilyloxy)methyl]-1, 3-dioxolan-4-yl] thymine (4.27A) and 3-ethyl-1-[(2R, 4R)-1-[(2-tert-butyldiphenylsilyloxy)methyl]-1, 3-dioxolan-4-yl] thymine (4.27 B)

Compounds 4.27 A and 4.27 B were synthesized according to the procedure described for the synthesis of compounds 4.24 A and 4.24 B using 4.23 (40 mg, 0.10 mmol) and 4.10 (33 mg, 0.21 mmol) as starting materials Target compound 4.26 A (12 mg, 25 %): Rf 0.65 (hexanes: ethyl acetate, 2:1) 1H-NMR (CDCl3) δ 1.07 (s, 9H, -Si(Ph)) 1.22 (t, 3H, J = 7.1 Hz), 1.95 (s, 3H, CH3), 3.71 (m, 2H, H-6'), 4.00 (m, 3H, CH2, H-5'), 4.39 (dd, 1H, J = 5.4 Hz, J = 9.5 Hz, H-5'), 5.55 (t, 1H, J = 3.1 Hz, H-2'), 6.35 (dd, 1H, J = 2.2 Hz, J = 5.4 Hz, H-4'), 7.15 (d, 1H, J = 1.1 Hz, H-6), 7.43 (m, 6H, -Si(Ph)2), 7.69 (m, 4H, -Si(Ph)2). MS (HR-ESI) C27H34N2O5Si (M+ Na)+ calcd 517.2135, found 517.2097. Target compound 4.26 B (16 mg, 33 %): Rf 0.52 (hexanes: ethyl acetate, 2:1) 1H-NMR (CDCl3) δ 1.07 (s, 9H, -Si(Ph)), 1.21 (t, 3H, J = 7.0 Hz, CH3), 1.71 (s, 3H, CH3), 3.91 (m, 2H, CH2), 3.99 (td, 2H, J = 6.3 Hz, J = 7.2 Hz), 4.15 (t, 2H, J = 6.9 Hz), 5.08 (t, 1H, J = 3.3 Hz, H-2'), 6.38 (dd, 1H, J = 2.8 Hz, J = 4.9 Hz, H-4'), 7.25 (s, 1H, H-6), 7.40 (m, 6H, -Si(Ph)2), 7.69 (d, 4H, J = 6.7 Hz, -Si(Ph)2). MS (HR-ESI) C27H34N2O5Si (M+ Na)+ calcd 517.2135, found 517.2096.
3-Methyl-[(2R, 4R)-1-[2-[(tert-Butyldiphenylsilyloxy) methyl]-1, 3-dioxolan-4-yl]] uracil (4.28 A) and 3-methyl [(2R, 4S)-1-[2-[(tert-butyldiphenylsilyloxy) methyl]-1, 3-dioxolan-4-yl]] uracil (4.28 B)

Compounds 4.28 A and 4.28 B were synthesized and isolated as a mixture (27 mg, ~60%) according to the procedure described for the synthesis of compounds 4.24 A and 4.24 B using 4.23 (40 mg, 0.10 mmol) and 4.11 (27 mg, 0.21 mmol) as starting materials. 

Rf: 0.56 (hexanes: ethyl acetate, 1:1) \(^1\)H NMR (CDCl\(_3\)) \(\delta\) 1.07 (s, 9H, -C(CH\(_3\))\(_3\)), 3.32 (s, 6H, CH\(_3\)), 3.71 (m, 4H, H-6’), 3.94 (m, 1H, J = 1.5 Hz, J = 5.0 Hz, H-5’), 4.02 (m, 2H), 4.19 (m, 2H), 4.41 (dd, 2H, J = 5.1 Hz, J = 9.5 Hz), 5.03 (t, 1H, J = 2.5 Hz), 5.54 (m, 2H), 5.81 (d, 2H, J = 8.1 Hz, H-5), 6.28 (dd, 1H, J = 2.1 Hz, J = 5.2 Hz, H-4’), 6.34 (dd, 1H, J = 1.8 Hz, J = 5.0 Hz, H-4’), 7.34 (d, 2H, J = 8.1 Hz, H-6), 7.46 (m, 12H, -Si(Ph)\(_2\)), 7.68 (m, 8H, -Si(Ph)\(_2\)). \(^{13}\)C NMR (CDCl\(_3\)) \(\delta\) 19.13, 26.72, 26.77, 27.55, 64.77, 71.32, 71.85, 83.44, 102.09, 106.15, 127.83, 127.88, 129.95, 130.03, 132.63, 132.79, 135.54, 135.60, 136.63, 151.26, 162.72. MS (HR-ESI) C\(_{25}\)H\(_{30}\)N\(_2\)O\(_5\)Si (M+ Na)\(^+\) calcd 489.1822, found 489.1787.
3-Methyl-1-[(2R, 4S)-2-(hydroxymethyl)-1, 3-dioxolan-4-yl] thymine (4.31)

A mixture of 4.24 A (10 mg, 0.021 mmol) and 1.0 M tetra-n-butylammonium fluoride (TBAF) in THF (0.024 mL, 0.024 mmol) in THF (3 mL) was stirred at room temperature for 1 h. The reaction mixture was concentrated in vacuo and the residue was purified by silica gel column chromatography (dichloromethane: methanol, 20:1) to yield compound 4.31 (4.5 mg, 90 %) as a colorless oil. R:\(f\): 0.47 (dichloromethane: methanol; 15:1). \(^1\)H-NMR (DMSO) \(\delta\) 1.85 (s, 3H, CH\(_3\)), 3.16 (s, 3H, CH\(_3\)), 3.44 (m, 2H, H-6\(\text{'}\)), 4.05 (dd, 1H, \(J = 3.0\) Hz, \(J = 9.5\) Hz, H-5\(\text{'}\)), 4.30 (dd, 1H, \(J = 5.6\) Hz, \(J = 9.4\) Hz, H-5\(\text{'}\)), 5.03 (t, 1H, \(J = 6.0\) Hz, OH), 5.51 (t, 1H, \(J = 3.6\) Hz, H-2\(\text{'}\)), 6.20 (dd, 1H, \(J = 3.1\) Hz, \(J = 5.5\) Hz, H-4\(\text{'}\)), 7.49 (s, 1H, H-6). \(^{13}\)C NMR (DMSO) \(\delta\) 12.80, 27.38, 61.94, 69.26, 82.75, 105.73, 108.64, 134.24, 150.68, 162.94. MS (HR-ESI) C\(_{10}\)H\(_{14}\)N\(_2\)O\(_5\) (M+ Na)\(^+\) calcd 265.0800, found 265.078.
3-Methyl-1-[(2R, 4R)-2-hydroxymethyl-1, 3-dioxolan-4-yl] thymine (4.32)

Compounds 4.32 (6.5 mg, 81%) was synthesized from compound 4.24 B (16 mg, 0.033 mmol) according to the procedure described for the synthesis of compounds 4.31. Rf: 0.48 (dichloromethane: methanol; 15:1). $^1$H-NMR (DMSO) $\delta$ 1.81 (s, 3H, CH$_3$), 3.16 (s, 3H, CH$_3$), 3.66 (dd, 2H, $J = 2.1$ Hz, $J = 4.1$ Hz, H-6′), 4.08 (dd, 1H, $J = 5.7$ Hz, $J = 9.9$ Hz, H-5′), 4.24 (m, 1H, H-5′), 4.92 (t, 1H, $J = 2.4$ Hz, H-2′), 5.24 (t, 1H, $J = 6.1$ Hz, OH), 6.27 (m, 1H, H-4′), 7.78 (s, 1H, H-6). $^{13}$C NMR (DMSO) $\delta$ 12.94, 27.41, 60.16, 70.37, 81.03, 105.15, 108.32, 134.64, 150.82, 162.90. MS (HR-ESI) C$_{10}$H$_{14}$N$_2$O$_5$ (M+ Na)$^+$ calcd 265.0800, found 265.0783.

3-Ethyl-1-[(2R, 4S)-2-hydroxymethyl-1, 3-dioxolan-4-yl] thymine (4.37)
Compounds 4.37 (4.5 mg, 75%) was synthesized from compound 4.27 A (12 mg, 0.024 mmol) according to the procedure described for the synthesis of compounds 4.31. Rf: 0.50 (dichloromethane: methanol; 15:1). $^1$H-NMR (DMSO) δ 1.08 (t, 3H, $J = 7.0$ Hz, CH$_3$), 1.85 (s, 3H, CH$_3$), 3.44 (ddd, 2H, $J = 2.3$ Hz, $J = 7.4$ Hz, $J = 9.3$ Hz, H-6’), 3.83 (q, 2H, $J = 6.9$ Hz, CH$_2$), 4.07 (dd, 2H, $J = 3.0$ Hz, $J = 9.5$ Hz, H-5’), 4.30 (dd, 1H, $J = 5.6$ Hz, $J = 9.4$ Hz, H-5’), 5.03 (t, 1H, $J = 6.0$ Hz, OH), 5.50 (t, 1H, $J = 3.6$ Hz, H-2’), 6.21 (dd, 1H, $J = 3.0$ Hz, $J = 5.5$ Hz, H-4’), 7.49 (s, 1H, H-6). $^{13}$C NMR (DMSO) δ 12.67, 12.79, 35.57, 61.95, 69.18, 82.59, 105.69, 108.92, 134.42, 150.31, 162.49. MS (HR-ESI) C$_{11}$H$_{16}$N$_{2}$O$_{5}$ (M+ Na)$^+$ calcd 279.0952, found 279.0959.

3-Ethyl-1-[(2R, 4R)-2-hydroxymethyl-1, 3-dioxolan-4-yl] thymine (4.38)

Compounds 4.38 (6.0 mg, 75%) was synthesized from compound 4.27 B (16 mg, 0.032 mmol) according to the procedure described for the synthesis of compounds 4.31. Rf: 0.52 (dichloromethane: methanol; 15:1). $^1$H-NMR (DMSO) δ 1.08 (t, 3H, $J = 7.0$ Hz, CH$_3$), 1.81 (s, 3H, CH$_3$), 3.65 (d, 2H, $J = 6.0$ Hz, H-6’), 3.83 (dd, 2H, $J = 6.8$ Hz, $J = 13.8$ Hz, CH$_2$), 4.09 (dd, 1H, $J = 5.8$ Hz, $J = 10.0$ Hz, H-5’), 4.25 (d, 1H, $J = 10.0$ Hz, H-5’), 4.93 (t, 1H, $J = 2.2$ Hz, H-2’), 5.22 (t, 1H, $J = 6.1$ Hz, OH), 6.28 (d, 1H, $J = 5.4$ Hz, H-4’), 7.76 (s, 1H, H-6). $^{13}$C NMR (DMSO) δ 12.68, 12.90, 35.56, 60.16, 70.27, 80.86.
105.10, 108.58, 134.81, 150.41, 162.41. MS (HR-ESI) $\text{C}_{11}\text{H}_{16}\text{N}_{2}\text{O}_{5}$ (M+ Na)$^+$ calcd 279.0952, found 279.0963.

3-Methyl [(2R, 4R)-2-hydroxymethyl-1, 3-dioxolan-4-yl] uracil (4.39 A) and 3-methyl [(2R, 4S)-2-hydroxymethyl-1, 3-dioxolan-4-yl] uracil (4.39 B)

Compounds 4.39 A and 4.39 B were synthesized (8 mg, 82%) as a mixture from compound 4.28 A and 4.28 B (20 mg, 0.043 mmol) according to the procedure described for the synthesis of compounds 4.31 $R_f$: 0.46 (dichloromethane: methanol; 15:1). $^1$H NMR (DMSO) $\delta$ 3.13 (s, 3H, CH$_3$), 3.43 (m, 2H, H-6'), 3.64 (dd, 2H, $J = 2.2$ Hz, $J = 5.9$ Hz, H-6'), 4.08 (m, 2H, H-5'), 4.28 (m, 2H, H-5'), 4.93 (t, 1H, $J = 2.5$ Hz, H-2'), 5.03 (t, 1H, $J = 6.1$ Hz, OH), 5.21 (t, 1H, $J = 5.9$ Hz, OH), 5.46 (t, 1H, $J = 3.6$ Hz, H-2'), 5.76 (d, 2H, $J = 8.0$ Hz, H-5), 6.16 (dd, 1H, $J = 3.0$ Hz, $J = 5.5$ Hz, H-4'), 6.23 (d, 1H, $J = 5.0$ Hz, H-4'), 7.62 (d, 2H, $J = 8.1$ Hz, H-6), 7.88 (d, 2H, $J = 8.1$ Hz, H-6) $^{13}$C NMR (DMSO) $\delta$ 27.11, 27.14, 60.22, 61.88, 69.46, 70.55, 81.33, 83.14, 100.8, 100.86, 105.30, 105.84, 138.72, 138.86, 150.80, 150.96, 162.17. MS (HR-ESI) $\text{C}_{9}\text{H}_{12}\text{N}_{2}\text{O}_{5}$ (M+ Na)$^+$ calcd 251.0644, found 251.0655.
L-(2S, 4S) - and L-(2S, 4R)-4-Acetoxy-2-[(benzoyloxy)-methyl] dioxolane (4.54)

Compound 4.54 was synthesized from L-gulonic acid-\(\gamma\)-lactone according to the procedure described in detail in Section 4.6.4.\(^{352}\) \(^1\)H NMR (CDCl\(_3\)) \(\delta\) 2.00, 2.11 (2s, 3H, OAc), 3.99-4.28 (m, 2H, H-5), 4.43-4.53 (m, CH\(_2\)OBz), 5.45, 5.56 (2 t, \(J = 3.8\) Hz, 3.6 Hz, 1 H, H-2), 6.38-6.45 (m, 1H, H-4), 7.42-7.59, 8.04-8.10 (m, 3H, 2H, OBz). MS (HR-ESI) C\(_{13}\)H\(_{14}\)O\(_6\) (M+ Na)\(^+\) calcd 289.0688, found 289.0693.

3-Methyl-1-[(2S, 4R)-1-[2-(benzoyloxy) methyl]-1, 3-dioxolan-4-yl] thymine (4.64 A) and 3-methyl-1-[(2S, 4S)-1-[2-[(benzoyloxy) methyl]-1, 3-dioxolan-4-yl] thymine (4.64 B)

A mixture of compound 4.3 (50 mg, 0.355 mmol), a catalytic amount of (NH\(_4\))\(_2\)SO\(_4\) (~5-10 mg), and HMDS (5 mL) was refluxed overnight. The resulting clear solution was concentrated in vacuo under anhydrous conditions to afford a colorless oil. To this oily residue was added dropwise a solution of 4.54 (0.035 g, 0.13 mmol) in anhydrous dichloroethane (6 mL) followed by TMSOTf (0.095 mL, 0.52 mmol). The mixture was stirred for 3 h at room temperature and then quenched with saturated aqueous sodium
bicarbonate (10 mL). The mixture was stirred for an additional 30 min followed by dilution with water (20 mL) and extraction with dichloromethane (3 x 30 mL). The combined organic layer was washed with saturated aqueous sodium bicarbonate solution and brine, dried with magnesium sulfate, and concentrated in vacuo. The residue was purified and both anomers were separated by silica gel column chromatography to obtain **4.64 A** in 26 % yield (12 mg). $R_f$ 0.37 (hexanes: ethyl acetate, 1:1) $^1$H-NMR($\text{CDCl}_3$) $\delta$

1.97 (s, 3H, CH$_3$), 3.34 (s, 3H, CH$_3$), 4.12 (dd, 1H, $J = 2.6$ Hz, $J = 9.8$ Hz, H-6’), 4.44 (m, 3H, H-5’, H-6’), 5.79 (t, 1H, $J = 3.7$ Hz, H-2’), 6.38 (dd, 1H, $J = 2.6$ Hz, $J = 5.4$ Hz, H-4’), 7.18 (s, 1H, H-6), 7.47 (t, 2H, $J = 7.7$ Hz, -OBz), 7.60 (dt, 1H, $J = 1.1$ Hz, $J = 7.8$ Hz, -OBz), 8.06 (d, 2H, $J = 8.3$ Hz, -OBz). $^{13}$C NMR ($\text{CDCl}_3$) $\delta$ 13.24, 27.84, 63.46, 70.55, 83.15, 103.45, 110.79, 128.55, 129.30, 129.76, 132.34, 133.47, 151.23, 163.38, 166.97. MS (HR-ESI) C$_{17}$H$_{18}$N$_2$O$_6$ (M+ Na)$^+$ calcd 369.1063, found 369.1038. Target compound **4.64 B** was isolated in 43 % yield (20 mg). $R_f$ 0.29 (hexanes: ethyl acetate, 2:1) $^1$H-NMR($\text{CDCl}_3$) $\delta$

1.69 (s, 3H, CH$_3$), 3.33 (s, 3H, CH$_3$), 4.22 (d, 2H, $J = 3.5$ Hz, H-6’), 4.66 (ddd, 2H, $J = 3.0$ Hz, $J = 12.5$ Hz, $J = 35.2$ Hz, H-5’), 5.29 (t, 1H, $J = 3.0$ Hz, H-2’), 6.43 (dd, 1H, $J = 3.5$ Hz, $J = 4.4$ Hz, H-4’), 7.29 (s, 1H), 7.45 (t, 2H, $J = 7.7$ Hz, -OBz), 7.59 (t, 1H, $J = 7.4$ Hz, -OBz), 8.05 (m, 2H, -OBz). $^{13}$C NMR ($\text{CDCl}_3$) $\delta$ 13.12, 27.88, 62.27, 71.85, 81.58, 102.88, 110.02, 128.60, 129.78, 132.49, 133.56, 151.40, 163.31, 165.97. MS (HR-ESI) C$_{17}$H$_{18}$N$_2$O$_6$ (M+ Na)$^+$ calcd 369.1063, found 369.1049.
3-Ethyl-1-[(2S, 4R)-1-2-[benzoyloxy] methyl]-1, 3-dioxolan-4-yl] thymine (4.67A)
and 3-ethyl-1-[(2S, 4S)-1-2-[benzoyloxy] methyl]-1, 3-dioxolan-4-yl] thymine (4.67 B)

Compounds 4.67 A and 4.67 B were synthesized in 21% yield (10 mg) according to
the procedure described for the synthesis of compounds 4.64 A using 4.54 (35 mg, 0.13
mmol) and 4.10 (43 mg, 0.28 mmol) as starting materials. Rf 0.39 (hexanes: ethyl acetate,
1:1) 1H NMR (CDCl3) δ 1.25 (m, 3H, CH3), 1.96 (s, 3H, CH3), 4.00 (m, 2H, CH2H-6’),
4.12 (m, 2H, CH2), 4.45 (m, 2H, H-5’), 5.79 (t, 1H, J = 3.8 Hz, H-2’), 6.39 (dd, 1H, J =
2.6 Hz, J=5.3 Hz, H-4’), 7.16 (s, 1H, H-6), 7.47 (t, 2H, J = 7.7 Hz, -OBz), 7.60 (t, 1H,
J=7.4 Hz, -OBz), 8.06 (d, 2H, J = 7.3Hz, -OBz). Target compound 4.64 B was isolated in
32% yield (15 mg). Rf 0.28 (hexanes: ethyl acetate, 1:1) 1H NMR (CDCl3) δ 1.17 (t, 3H,
J = 7.2Hz, CH3), 1.77 (s, 3H, CH3), 3.94 (m, 2H, H-6’), 4.23 (m, 2H,CH2), 4.66 (m, 2H,
H-5’), 5.20 (t, 1H, J = 2.7 Hz, H-2’), 6.44 (m, 1H, H-4’), 7.26 (s, 1H, H-6), 7.46 (dd, 2H,
J = 7.3Hz, J = 13.3Hz, -OBz), 7.60 (t, 1H, J = 7.6 Hz), 8.07 (m, 2H, -OBz).
3-Methyl-(2S, 4R)-1-[2-[(benzoyloxy) methyl]-1, 3-dioxolan-4-yl] uracil (4.68 A) and 3-methyl (2S, 4S)-1-[2-[(benzoyloxy) methyl]-1, 3-dioxolan-4-yl] uracil (4.68 B)

The mixture of compounds 4.68 A and 4.68 B (25 mg, ~60%) were synthesized according to the procedure described for the synthesis of compounds 4.64 A and 4.64 B using 4.54 (35 mg, 0.13 mmol) and 4.11 (35 mg, 0.28 mmol) as starting materials.

Rf 0.17 (hexanes: ethyl acetate, 1:1) $^1$H NMR (CDCl$_3$) 3.31 (s, 3H, CH$_3$), 4.18 (m, 2H), 4.46 (m, 2H), 4.68 (ddd, 4H, $\times$ = 2.5 Hz, $\times$ = 12.7 Hz, $\times$ = 40.9 Hz), 5.28 (t, 2H, $\times$ = 2.5 Hz), 5.54 (d, 2H, $\times$ = 8.1 Hz), 5.76 (t, 2H, $\times$ = 3.8 Hz), 5.83 (d, 1H, $\times$ = 8.2 Hz), 6.31 (dd, 1H, $\times$ = 2.4 Hz, $\times$ = 5.1 Hz, H-4'), 6.39 (dd, 1H, $\times$ = 2.4 Hz, $\times$ = 4.6 Hz, H-4'), 7.36 (d, 1H, $\times$ = 8.1 Hz, H-6), 7.47 (t, 4H, $\times$ = 7.7 Hz, H-6), 7.60 (m, 2H), 8.05 (m, 4H). $^{13}$C NMR (CDCl$_3$) $\delta$ 27.59, 27.63, 61.85, 63.57, 70.87, 71.65, 81.80, 83.53, 102.42, 103.14, 128.56, 128.60, 129.76, 133.49, 133.69, 136.43, 136.95, 151.21, 151.32, 162.48, 162.56, 165.85, 165.93. MS (HR-ESI) C$_{16}$H$_{16}$N$_2$O$_6$ (M+ Na)$^+$ calcd 355.0906, found 355.0880.
3-Methyl-1-[(2S, 4R)-2-(hydroxymethyl)-1, 3-dioxol-4-yl] thymine (4.71)

A mixture of 4.64 A (12 mg, 0.035 mmol) in NH₃ in methanol (70 %, 5 mL) was stirred at room temperature for 2 days. Following evaporation, the residue was purified by silica gel column chromatography to yield 4.71 (5.59 mg, 66 %) as a colorless oil.

Rf. 0.41 (dichloromethane: methanol; 15:1) ¹H-NMR (CD₃OD) δ 1.92 (s, 3H, CH₃), 3.28 (s, 3H, CH₃), 3.60 (m, 2H, H-6’), 4.07 (dd, 1H, J = 3.1 Hz, J = 9.5 Hz, H-5’), 4.41 (dd, 1H, J = 5.5 Hz, J = 9.5 Hz, H-5’), 5.49 (t, 1H, J = 3.3 Hz, H-2’), 6.23 (dd, 1H, J = 3.1 Hz, J = 5.5 Hz, H-4’), 7.47 (s, 1H). ¹³C NMR (CD₃OD) δ 13.23, 28.11, 63.61, 71.65, 85.28, 107.35, 110.91, 135.55, 152.70, 165.68. MS (HR-ESI) C₁₀H₁₄N₂O₅ (M+ Na)⁺ calcd 265.0800, found 265.0787.

3-Methyl-1-[(2S, 4S)-2-(hydroxymethyl)-1, 3-dioxol-4-yl] thymine (4.72)
Compounds 4.72 (10 mg, 71 %) was synthesized from the compound 4.64 B (20 mg, 0.08 mmol) according to the procedure described for the synthesis of compounds 4.71.

\[ R_f = 0.45 \, \text{(dichloromethane: methanol; 15:1).} \]

\[ ^1\text{H-NMR (CD}_3\text{COCD}_3) \delta 1.84 \, (s, \, 3H, \, \text{CH}_3), \]

3.21 (s, 3H, CH$_3$), 3.82 (m, 2H, H-6'), 4.19 (dd, 1H, \( J = 5.7 \) Hz, \( J = 10.0 \) Hz, H-5'), 4.27 (dd, 1H, \( J = 1.4 \) Hz, \( J = 10.0 \) Hz, H-5'), 5.02 (t, 1H, \( J = 2.3 \) Hz, H-2'), 6.35 (dd, 1H, \( J = 1.5 \) Hz, \( J = 5.6 \) Hz, H-4'), 7.85 (s, 1H, H-6). ¹³C NMR (CD$_3$COCD$_3$) \( \delta \) 13.32, 27.74, 61.81, 71.88, 82.46, 106.45, 109.84, 135.35, 152.18, 163.96. MS (HR-ESI) C$_{10}$H$_{14}$N$_2$O$_5$ (M+ Na)$^+$ calcd 265.0800, found 265.0782.

\[
\begin{align*}
\text{3-Ethyl-1-[(2S, 4R)-2-(hydroxymethyl)-1, 3-dioxolan-4-yl] thymine (4.77)}
\end{align*}
\]

Compound 4.77 (4.9 mg, 70%) was synthesized from the compound 4.67 A (10 mg, 0.028 mmol) according to the procedure described for the synthesis of compounds 4.71.

\[ R_f = 0.48 \, \text{(dichloromethane: methanol, 15:1).} \]

\[ ^1\text{H-NMR (DMSO)} \delta 1.08 \, (t, \, 3H, \, J = 7.0 \) Hz, CH$_3$), 1.85 (s, 3H, CH$_3$), 3.43 (td, 2H, \( J = 4.7 \)Hz, \( J = 9.0 \) Hz, H-6'), 3.83 (q, 2H, \( J = 6.8 \) Hz, CH$_2$), 4.07 (dd, 1H, \( J = 3.0 \) Hz, \( J = 9.5 \) Hz, H-5'), 4.30 (dd, 1H, \( J = 5.6 \) Hz, \( J = 9.5 \) Hz, H-5'), 5.02 (t, 1H, \( J = 6.0 \) Hz, OH), 5.50 (t, 1H, \( J = 3.6 \) Hz, H-2'), 6.21 (dd, 1H, \( J = 3.0 \) Hz, \( J = 5.6 \) Hz, H-4'), 7.49 (s, 1H, H-6). ¹³C NMR (DMSO) \( \delta \) 12.66, 12.79, 35.56,
3-Ethyl-1-[(2S, 4S)-2-(hydroxymethyl)-1, 3-dioxolan-4-yl] thymine (4.78)

Compounds 4.78 (7.5 mg, 75%) was synthesized from the compound 4.67 B (15 mg, 0.042 mmol) according to the procedure described for the synthesis of compounds 4.71. 

\[ R_f \ 0.50 \text{ (dichloromethane: methanol, 15:1)} \]

\[^1H-NMR\ (DMSO) \ \delta \ 1.08 \text{ (t, 3H, } J = 7.0 \text{ Hz, CH}_3\), 1.81 \text{ (s, 3H, CH}_3\), 3.65 \text{ (d, 2H, } J = 5.6 \text{ Hz, H-6}’\), 3.83 \text{ (dd, 2H, } J = 6.3 \text{ Hz, } J = 13.5 \text{ Hz, CH}_2\), 4.09 \text{ (dd, 1H, } J = 5.8 \text{ Hz, } J = 9.9 \text{ Hz, H-5}’\), 4.25 \text{ (d, 1H, } J = 10.0 \text{ Hz, H-5}’\), 4.93 \text{ (s, 1H, H-2}’\), 5.22 \text{ (t, 1H, } J = 6.0 \text{ Hz, OH)}, 6.28 \text{ (d, 1H, } J = 5.6 \text{ Hz, H-4}’\), 7.75 \text{ (s, 1H, H-6)}. \]

\[^{13}C-NMR\ (DMSO) \ \delta \ 12.68, 12.90, 35.56, 60.16, 70.27, 80.86, 105.10, 108.58, 134.81, 150.41, 162.42. \]

\[ \text{MS (HR-ESI) } C_{11}H_{16}N_2O_5 \ \text{(M+Na)}^+ \ \text{calcd 279.0957, found 279.0951}. \]
3-Methyl-[(2S, 4R)-2-(hydroxymethyl)-1, 3-dioxolan-4-yl] uracil (4.79 A) and 3-
methyl-[(2S, 4S)-2-(hydroxymethyl)-1, 3-dioxolan-4-yl] uracil (4.79 B)

The mixture of compounds 4.79 A and 4.79 B (12 mg, ~70%) was synthesized from
the mixture of compounds 4.68 A and 4.68 B (25 mg, 0.075 mmol) according to the
procedure described for the synthesis of compound 4.71. Rf, 0.46 (dichloromethane:
methanol, 15:1)1H-NMR (DMSO) δ 3.14 (s, 6H, CH3), 3.44 (m, 2H, H-6’), 3.65 (m, 2H,
H-6’), 4.09 (m, 2H, H-5’), 4.29 (m, 2H, H-5’), 4.94 (t, 1H, J = 2.6 Hz, H-2’), 5.03 (t, 1H,
J = 5.9 Hz, OH), 5.22 (t, 1H, J = 5.9 Hz, OH), 5.47 (t, 1H, J = 3.7 Hz, H-2’), 5.76 (dd,
2H, J = 1.3 Hz, J = 8.1 Hz, H-5), 6.17 (dd, 1H, J = 2.8 Hz, J = 5.4 Hz, H-4’), 6.24 (d, 1H,
J = 5.4 Hz, H-4’), 7.63 (d, 1H, J = 8.1 Hz), 7.89 (d, 1H, J = 8.1 Hz, H-6). 13C NMR
(DMSO) δ 27.10, 27.13, 60.21, 61.87, 69.45, 70.55, 81.32, 83.13, 100.79, 100.85, 105.30,
105.83, 138.71, 138.85, 150.79, 150.96, 162.16, 162.18. MS (HR-ESI) C9H12N2O5
(M+Na) + calc 251.0638, found 251.0636.
5.4. X-ray crystallography

The data collection crystal, composed of 3.7 a, was a clear, colorless rectangular plate. Examination of the diffraction pattern on a Nonius Kappa CCD diffractometer indicated a monoclinic crystal system. All work was done at 150 K using an Oxford cryosystems cooler. The data collection strategy was set up to measure a quadrant of reciprocal space with a redundancy factor of 4.7, which means that 90% of these reflections were measured at least 4.7 times. Phi and omega scans with a frame width of 1.0° were used. Data integration was done with Denzo,\textsuperscript{364} and an absorption correction and merging of the data was done with Sortav.\textsuperscript{365,366} Merging the data and averaging the symmetry equivalent reflections (for the Laue group m) resulted in an Rint value of 0.040.

The structure was solved in Cc by the direct methods procedure in SHELXS-97.\textsuperscript{367} Full-matrix least-squares refinements based on F\textsuperscript{2} were performed in SHELXL-97, as incorporated in the WinGX package.\textsuperscript{368} For each methyl group, the hydrogen atoms were added at calculated positions using a riding model with U (H) = 1.5*Ueq (bonded carbon atom). The torsion angle, which defines the orientation of the methyl group about the C-C or Si-C bond, was refined. The hydrogen atoms in the cluster were refined isotropically.

The final refinement cycle was based on 4013 intensities and 226 variables and resulted in agreement factors of R1 (F) = 0.032 and wR2 (F\textsuperscript{2}) = 0.044. For the subset of data with I > 2*sigma (I), the R1 (F) value is 0.024 for 3573 reflections. The final difference electron density map contains maximum and minimum peak heights of 0.62 and -0.38 e/Å\textsuperscript{3}. Neutral atom scattering factors were used and include terms for anomalous dispersion.\textsuperscript{369}
5.5. Biological studies

5.5.1. Activity measurements with recombinant human thymidine kinase (hTK1)

The assay cocktail for the measurement of the initial velocity of the phosphorylation of $[^{3}H]$ dThd contained 50 mM Tris-HCl pH 7.6, 2 mM MgCl$_2$, 2 mM ATP, 0.5 mg/mL BSA, 5 mM DTT, 15 mM NaF, 2 ng of enzyme and 1 μM tritiated dThd (25 Ci/mmol) in a total volume of 50 μL. The reaction was performed at 37 °C and the aliquots from the assay mixtures were removed at 0, 5, 10, and 15 min and applied to DE-81 filter paper. The filters were washed 3 x 5 min with 5 mM ammonium formate and 1 x 5 min in water. Reaction products were then eluted with 0.5 mL 0.1 M HCl and 0.2 M KCl and the quantitation of tritiated thymidine monophosphate was performed using liquid scintillation counting.

5.5.2. Phosphoryl transfer assays with recombinant hTK1 and recombinant Bacillus Anthracis thymidine kinase (BaTK)

Phosphoryl transfer assays were performed with 100 μM $[^{32}P]$ of ATP in the same buffer as described in Section 5.5.1 using 100 μM substrate and 50 ng of hTK1 or 10 ng for BaTK in a total volume of 20 μL. The assay was conducted at 37 °C over a period of 20 min after which the phosphoryl transfer was terminated by heating the reaction mixture at 100°C for 2 min. After centrifugation, 2 μL of the reaction cocktail were separated by thin layer chromatography on PEI-cellulose plates using isobutyric acid: NH$_3$OH: H$_2$O (v/v, 66:1:33) as the mobile phase, and then quantified by phosphoimaging analysis.
BIBLIOGRAPHY


32. Gustafson, E. A.; Chillemi, A. C.; Sage, D. R.; Fingeroth, J. D. The Epstein-Barr virus thymidine kinase does not phosphorylate ganciclovir or acyclovir and demonstrates a narrow substrate specificity compared to the herpes


67. Bandyopadhyaya, A. K.; Tiwari, R.; Tjarks, W. Comparative molecular field analysis and comparative molecular similarity indices analysis of boron-


108. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21 + G


122. Kassis, A. I. 5-\[^{123}I/^{125}I\]-iodo-2'-deoxyuridine for cancer diagnosis and therapy. 

123. Kassis, A. I.; Adelstein, S. J. Preclinical animal studies with 

124. Kassis, A. I.; Adelstein, S. J. 5-[\(^{125}I\)]iodo-2'-deoxyuridine in the radiotherapy 
   935-939.

125. Kassis, A. I.; Adelstein, S. J.; Mariani, G. Radiolabeled nucleoside analogs in 

126. Kassis, A. I.; Kirichian, A. M.; Wang, K.; Semnani, E. S.; Adelstein, S. J. 
   Therapeutic potential of 5-[\(^{125}I\)]iodo-2'-deoxyuridine and methotrexate in the 
   941-946.

127. Kassis, A. I.; Tumeh, S. S.; Wen, P. Y. C.; Baranowska-Kortylewicz, J.; Van 
   den Abbeele, A. D.; Zimmerman, R. E.; Carvalho, P. A.; Garada, B. M.; 
   DeSisto, W. C.; *et al.* Intratumoral administration of 5-[\(^{123}I\)]iodo-2'-deoxyuridine 

128. Kassis, A. I.; Van den Abbeele, A. D.; Wen, P. Y. C.; Baranowska- 
   Kortylewicz, J.; Aaronson, R. A.; DeSisto, W. C.; Lampson, L. A.; Black, P. 
   M.; Adelstein, S. J. Specific uptake of the auger electron-emitting thymidine 
   analog 5-[\(^{123}I/^{125}I\)]iodo-2'-deoxyuridine in rat brain tumors: Diagnostic and 


272


357. Baiocchi, R. A.; Caligiuri, M. A. Low-dose interleukin 2 prevents the development of Epstein-Barr virus (EBV)-associated lymphoproliferative disease in scid/scid mice reconstituted IP With EBV-seropositive human


APPENDIX

Figure A1: $^1$H NMR spectrum of 3.24. The inset shows the typical splitting pattern (apparent triplet) of C1’-H
Figure A2: $^{13}$C NMR spectra of 3.24. The red box highlights the chemical shifts associated with C4’, C1’, C3’ and C5’ of the deoxyribofuranose ring.
Figure A3: $^{13}$C DEPT 135 NMR spectra of 3.24. The red box highlights the chemical shifts for C4’, C1’, C3’ and C5’. Based on multiplicity pattern observed in this spectrum, the probability that the signal at 62.80 ppm relates to C5’ becomes higher.
Figure A4: HSQC spectra of 3.24. The correlation between $^1$H NMR and $^{13}$C NMR for the assignments of both $^1$H NMR and $^{13}$C NMR signals.
Figure A5: $^1$H NMR spectra of 3.26. The inset shows the typical splitting pattern (doublete of a doublet) of the C1’-H proton of an $\alpha$-nucleoside.
Figure A6: $^{13}$C NMR spectra of 3.26. The inset shows key signals for the carbohydrate portion.
Figure A7: $^{13}$C DEPT135 NMR spectra of 3.26. The red box highlights the chemical shifts for C4’, C1’, C3’ and C5’. Based on multiplicity pattern observed in this spectrum, the probability that the signal at 67.00 ppm relates to C5’ becomes higher.
**Figure A8**: HSQC NMR spectra of 3.26. The correlation between $^1$H NMR and $^{13}$C NMR for the assignments of both $^1$H NMR and $^{13}$C NMR signals.
Figure A9: HMBC spectra of N1-Boc-thymine. The spectra did not show coupling of the N3 hydrogen with C5 and C6.
Figure A10: The HMBC spectra of N3-butyl thymine (see Chapter 3, compound 3.24) showing coupling of the α-methylene hydrogens with C-2 and C-4.