A KINETIC AND BIOCHEMICAL APPROACH TO UNDERSTANDING THE MECHANISMS OF NOVEL DNA POLYMERASES.

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

By

Kevin Andrew Fiala, B.S.

*****

The Ohio State University

2007

Dissertation Committee:

Approved by

Professor Zucai Suo, Adviser

Professor Michael Ibba

Professor Richard P. Swenson

Professor Dongping Zhong

Ohio State Biochemistry Program
ABSTRACT

DNA polymerases are the enzymes responsible for the vital task of faithfully duplicating genomes in order to pass on these genetically encoded instructions to their offspring. However, the process of faithfully propagating this information is hindered in all organisms due to endogenous and exogenous agents that damage DNA. While DNA repair mechanisms correct the vast majority of the resulting DNA lesions, unrepaired lesions do persist in the presence of fully functional repair mechanisms. Fortunately cells have evolved a class of promiscuous enzymes known as lesion bypass polymerases that have been shown to bypass DNA lesions that stall the high fidelity replicative DNA polymerases. Here, we have studied two DNA polymerases, human DNA polymerase λ and *Sulfolobus solfataricus* DNA polymerase IV (Dpo4), which are thought to be involved in the previously mentioned cellular processes of DNA repair and DNA lesion bypass respectively.

In the process of establishing a minimal kinetic mechanism for the incorporation of a single nucleotide into undamaged DNA catalyzed by human DNA polymerase λ, we discovered a novel mechanism in which one of its non-enzymatic N-terminal domains, the Proline-rich domain, dramatically increases the fidelity of the C-terminal DNA polymerase β-like domain by 10- to 100-fold to the level equivalent to that observed with DNA polymerase β, with which it shares 33% sequence identity. Moreover, we have also explored the effects of various structurally distinct DNA substrates on the catalytic
efficiency of nucleotide incorporation where we determined the downstream strand and its 5’-phosphate increase the incorporation efficiency by 15- and 11-fold respectively.

We have used *S. solfataricus* Dpo4 as a model Y-family DNA polymerase to elucidate the kinetic mechanism for nucleotide incorporation at both 37 °C and 56 °C, demonstrating that Dpo4 uses an induced-fit mechanism to select and incorporate a correct nucleotide into undamaged DNA independent of reaction temperature. We have also demonstrated using a variety of techniques that Dpo4 predominantly uses two distinct pathways (A-rule and lesion loop-out mechanism) to bypass an abasic site lesion. Taken together, these observations provide compelling evidence for the observation made by Joyce and Benkovic that DNA polymerases defy a unified description.
Dedicated to my mother and father
ACKNOWLEDGMENTS

I would like to thank my adviser Dr. Zucai Suo, who has fostered an environment privy to the pursuit of answers to interesting and challenging scientific questions and who has given me the opportunity to mature scientifically. Dr. Suo has assiduously provided the support to allow me to excel as a graduate student and has equipped me with the tools necessary to succeed post-graduate school.

I would like to thank my committee members Dr. Michael Ibba, Dr. Richard Swenson, and Dr. Dongping Zhong for their support and suggestions during my tenure as a graduate student. These committee members have provided excellent suggestions for many of my manuscripts and have been instrumental in helping me successfully compete for numerous awards and fellowships, in addition to helping me complete this dissertation.

I would also like to thank Professor Edward Behrman and Professor George Serif for initially getting me interested in the field of Biochemistry as an undergraduate student. I would like to acknowledge and thank Dr. Caroline Breitenberger and Ryan Periera for their support and encouragement as an undergraduate researcher in the laboratory of Dr. Breitenberger. Many thanks to Dr. Breitenberger who has also helped me with several fellowship applications.

I want to acknowledge and thank the current and past members of the Suo laboratory Wissam Abdel-Gawad, Nikunj Bhatt, Jessica Brown, Yuxia Dong, Wade Duym, Jason Fowler, Cameron Hypes, Sean Newmister, John Pryor, Michelle Roettger,
Sudarshan Seshadri, Shanen Sherrer, Susmitha Sompali, and Jun Zhang, all of whom I have collaborated with over the past five years. I would also like to acknowledge in particular the help of two undergraduate students, Cameron Hypes and John Pryor, who played important roles in the Dpo4 lesion bypass project and the human Y-family polymerase bypass project respectively, and who were both very deservedly awarded prizes at the Biological Sciences Undergraduate Research Colloquium (Cameron Hypes, 2005 and John Pryor, 2007) and the university-wide Denman Undergraduate Research Competition (John Pryor, 2007). I also want to thank Jason Fowler for all of his help and support. Besides providing great suggestions in regard to my research, he has helped me tremendously with a myriad of technical issues. His contributions cannot be overemphasized. In addition, I wanted to thank J. Marcela Hernandez for her helpful conversations, advice, and suggestions throughout my career at Ohio State University.

Thanks to Roger Woodgate (National Institutes of Health) for providing us with the initial clone of Dpo4. I would also like to thank Professor Zhigang Wang (University of Kentucky) for supplying us with the initial clone of the full-length human polymerase lambda gene. I also would like to thank Professor Hong Ling (University of Western Ontario), John-Stephen Taylor (Washington University), and Wei Yang (NIDDK, National Institutes of Health) for wonderful collaborations and for the materials and ideas they shared with our laboratory.

I am very grateful for the support of my parents and family who have always encouraged me and have shown incredible patience over the past five years. I am also very grateful to Marissa Vargas-Tokuda who has shouldered the vast majority of my frustration over the past two years. I want to thank them for their unwavering support.

Finally I would like to thank the current director of OSBP Dr. Ross Dalbey, the Biochemistry Department, the respective staff, faculty, and students for providing me
with an excellent interactive atmosphere that complemented my scientific and professional development.

This research was supported by grants awarded to my adviser from the National Science Foundation Career Award, the American Chemical Society Petroleum Fund, TriLink Biotechnologies in addition to the start-up fund from the Biochemistry Department. I would like to thank the NIH-sponsored Chemistry-Biology Interface Training Program at Ohio State University, the American Heart Association, the Herta Camerer Gross Graduate Research Fellowship, and the Presidential Fellowship at Ohio State University for providing me with funding for the entirety of my Ph.D. career.
VITA

January 6, 1978 ............... Born - Cleveland, Ohio

2001 ......................... B.S. Biochemistry, The Ohio State University

2002-Present.................. Ph.D. Biochemistry, The Ohio State University

PUBLICATIONS


viii


FIELDS OF STUDY

Major Field: Ohio State Biochemistry Program
# TABLE OF CONTENTS

Abstract ........................................................................................................... ii  
Dedication ........................................................................................................ iv  
Acknowledgments ............................................................................................ v  
Vita ................................................................................................................... viii  
Table of Contents ............................................................................................. x  
List of Tables .................................................................................................. xvii  
List of Figures .................................................................................................. xix  
List of Schemes ................................................................................................. xxiv  

Chapters:  

1 Introduction .................................................................................................. 1  

1.1 Introduction ............................................................................................ 1  

1.2 Structural and Mechanistic Analysis of DNA Polymerases .............. 4  

1.3 Focus of Dissertation ............................................................................ 9  

1.4 Figures and Schemes .......................................................................... 11  

1.5 References ............................................................................................ 14  

x
Pre-steady state kinetic studies of the fidelity of *Sulfolobus solfataricus* DNA polymerase IV at 37 °C

2.1 Introduction

2.2 Materials and Methods

2.3 Results

2.4 Discussion

2.5 Figures, Tables, Schemes

2.6 References

Mechanism of DNA polymerization catalyzed by *Sulfolobus solfataricus*

3.1 Introduction

3.2 Materials and Methods

3.3 Results

3.4 Discussion
3.5 Figures, Tables, Schemes.......................................................79

3.6 References..............................................................................93

4 Mechanism of abasic lesion bypass catalyzed by *Sulfolobus solfataricus*
DNA polymerase IV at 37 °C.......................................................96

4.1 Introduction.................................................................96

4.2 Materials and Methods................................................98

4.3 Results................................................................................99

4.4 Discussion..............................................................108

4.5 Figures, Tables, Schemes........................................118

4.6 References.............................................................140

5 Sloppy bypass of abasic lesion bypass catalyzed by *Sulfolobus solfataricus*
DNA polymerase IV.............................................................144

5.1 Introduction.............................................................144

5.2 Materials and Methods.............................................146

xii
5.3 Results and Discussion……………………………………………………………148

5.4 Figures, Tables, Schemes……………………………………………………….160

5.5 References…………………………………………………………………….177

6  Mechanistic consequences of temperature on DNA polymerization
catalyzed by *Sulfolobus solfataricus* DNA polymerase IV………………………180

6.1 Introduction……………………………………………………………………180

6.2 Materials and Methods……………………………………………………….182

6.3 Results…………………………………………………………………………186

6.4 Discussion……………………………………………………………………195

6.5 Figures, Tables, Schemes……………………………………………………204

6.6 References……………………………………………………………………224

7  Mechanism of template-independent nucleotide incorporation catalyzed by
a template-dependent DNA polymerase at 37 °C………………………………..228

7.1 Introduction…………………………………………………………………228
7.2 Materials and Methods...............................................................230

7.3 Results.....................................................................................233

7.4 Discussion...............................................................................238

7.5 Figures, Tables, Schemes..........................................................247

7.6 References...............................................................................266

8 Pre-steady state kinetic studies of the fidelity and mechanism of polymerization catalyzed by truncated DNA polymerase lambda...........269

8.1 Introduction...............................................................................269

8.2 Materials and Methods.............................................................272

8.3 Results.....................................................................................277

8.4 Discussion...............................................................................287

8.5 Figures, Tables, Schemes..........................................................297

8.6 References...............................................................................318
9 Up-regulation of the fidelity of human polymerase lambda by its non-enzymatic proline-rich domain at 37 °C…………………………………………………………………………322

9.1 Introduction……………………………………………………………………….322

9.2 Materials and Methods…………………………………………………………324

9.3 Results………………………………………………………………………………326

9.4 Discussion……………………………………………………………………………..329

9.5 Figures, Tables, Schemes…………………………………………………………………………335

9.6 References………………………………………………………………………………345

10 Kinetic effect of a downstream strand and its 5'-terminal moieties on single-nucleotide gap-filling synthesis catalyzed by human DNA polymerase lambda at 37 °C…………………………………………………………………………………………………………………………348

10.1 Introduction………………………………………………………………………………….348

10.2 Materials and Methods………………………………………………………………….350

10.3 Results……………………………………………………………………………………353

10.4 Discussion…………………………………………………………………………………..356
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.5</td>
<td>Figures, Tables, Schemes</td>
<td>361</td>
</tr>
<tr>
<td>10.6</td>
<td>References</td>
<td>373</td>
</tr>
<tr>
<td>11</td>
<td>Additional Results, Conclusions, and Future Directions</td>
<td>376</td>
</tr>
<tr>
<td>11.1</td>
<td>Introduction</td>
<td>376</td>
</tr>
<tr>
<td>11.2</td>
<td>Results and Future Directions for the Y-family Polymerases</td>
<td>377</td>
</tr>
<tr>
<td>11.3</td>
<td>Results and Future Directions for the Y-family Polymerases</td>
<td>384</td>
</tr>
<tr>
<td>11.4</td>
<td>Conclusions</td>
<td>392</td>
</tr>
<tr>
<td>11.5</td>
<td>Figures, Tables, Schemes</td>
<td>396</td>
</tr>
<tr>
<td>11.6</td>
<td>References</td>
<td>419</td>
</tr>
</tbody>
</table>

List of References | 421
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 DNA Substrates</td>
<td>48</td>
</tr>
<tr>
<td>2.2 Pre-steady State Kinetic Parameters of Dpo4</td>
<td>49</td>
</tr>
<tr>
<td>2.3 Comparison of Fidelity of Dpo4 with other DNA Polymerases</td>
<td>51</td>
</tr>
<tr>
<td>3.1 Estimated Kinetic Constants of Dpo4 at 37 °C</td>
<td>90</td>
</tr>
<tr>
<td>4.1 DNA Primer and Templates</td>
<td>129</td>
</tr>
<tr>
<td>4.2 Binding Affinity of Dpo4 to DNA Substrates at 23 °C</td>
<td>130</td>
</tr>
<tr>
<td>4.3 Kinetic Parameters for Incorporation at the Two Pause Sites</td>
<td>131</td>
</tr>
<tr>
<td>4.4 Comparison of the Substrate Specificity into 21/22AP</td>
<td>132</td>
</tr>
<tr>
<td>4.5 Kinetic Parameters for Incorporation Downstream of Damaged DNA</td>
<td>133</td>
</tr>
<tr>
<td>4.6 Kinetic Parameters for Incorporation Downstream of Damaged DNA</td>
<td>134</td>
</tr>
<tr>
<td>4.7 Realignment Possibilities of the Looped-out DNA Intermediates</td>
<td>135</td>
</tr>
<tr>
<td>4.8 Kinetic Parameters for DNA Substrates AP-1 and AP-8</td>
<td>136</td>
</tr>
<tr>
<td>4.9 Preference Between the A-rule and the Lesion Loop-out Mechanism</td>
<td>137</td>
</tr>
<tr>
<td>4.10 Kinetic Parameters for Damaged DNA Substrate 20/41AP</td>
<td>138</td>
</tr>
<tr>
<td>5.1 DNA Primers and Templates</td>
<td>163</td>
</tr>
<tr>
<td>5.2 Mutation Spectrum of Short Oligonucleotide Sequencing Assay</td>
<td>164</td>
</tr>
<tr>
<td>5.3 Comparison of Nucleotide Incorporation Opposite the AP Site</td>
<td>165</td>
</tr>
<tr>
<td>5.4 Comparison of Methods for Incorporation Opposite the AP Site</td>
<td>166</td>
</tr>
</tbody>
</table>
6.1 Kinetic Parameters for Incorporation into D-1.................................220
6.2 Kinetic Constants of Dpo4 at 37 °C and 56 °C.................................222
7.1 DNA Substrates..............................................................................262
7.2 Summary of X-ray Crystallographic Data........................................263
7.3 Comparison of Efficiencies for Blunt-end and Recessed DNA.........264
8.1 DNA 21-19/41-mer Substrates..........................................................314
8.2 Pre-steady State Kinetic Parameters of tPolλ..................................315
8.3 Comparison of Kinetic Parameters of tPolλ and Polβ......................316
9.1 Kinetic Parameters for Incorporation Catalyzed by fPolλ..................342
9.2 Kinetic Parameters for Incorporation Catalyzed by dPolλ................343
10.1 Kinetic Parameters for Incorporation by fPolλ..............................368
10.2 Kinetic Parameters with D-DNA, D-1OH, and D-1N....................369
10.3 Pre-steady State Kinetic Parameters of Correct dTTP Incorporation..370
11.1 Kinetic Parameters for Incorporation by Dpo4 and the Y12A mutant..414
11.2 DNA Substrates for the Dpo4 Incorporation Assays.......................415
11.3 Kinetic Parameters for Incorporation of matched rNTPs by fPolλ.....416
11.4 Kinetic Parameters for Incorporation of mismatched rNTPs by fPolλ..417
LIST OF FIGURES

1.1 Ternary Structure of Dpo4 with Undamaged DNA Substrate..........11
2.1 Purification of *S. solfataricus* Dpo4........................................40
2.2 Effects of Mg$^{2+}$ Concentration on Dpo4 Activity.........................41
2.3 Effects of NaCl Concentration on Dpo4 Activity..........................42
2.4 Effects of pH on Dpo4 Activity..............................................43
2.5 Concentration Dependence of the Rate of Correct Incorporation........44
2.6 Concentration Dependence of the Rate of Correct Incorporation........45
2.7 Concentration Dependence of the Rate of Incorrect Incorporation......46
2.8 Concentration Dependence of the Rate of Incorrect Incorporation.....47
3.1 Sequence of the D-1 DNA Substrate........................................79
3.2 Pre-steady State and Steady-state Kinetics for dTTP Incorporation ....80
3.3 Steady-state Kinetics of dTTP Incorporation into D-1 by Dpo4.........81
3.4 Active Site Titration of Dpo4 with D-1.......................................82
3.5 Elemental Effect on the Rate of Correct Nucleotide Incorporation.....83
3.6 Elemental Effect on the Rate of Incorrect Nucleotide Incorporation...84
3.7 Processive Polymerization by Dpo4...........................................85
3.8 Image of the Processive Polymerization by Dpo4..........................86
3.9 Measurement of the Dissociation Rate of E-DNA-dNTP Complex.......87
3.10 Measurement of the Dissociation Rate of E’-DNA-dNTP Complex.....88
3.11 Pulse-chase and Pulse-quench Experiment.................................89
4.1 Running Start Nucleotide Incorporation Assay for Dpo4...............118
4.2 Running Start Nucleotide Incorporation Assay for SsoPol B1........119
4.3 Measurement of the Binding of Dpo4 to 21/41AP.........................120
4.4 Measurement of the Binding of Dpo4 to 21/41AP Secondary Plot....121
4.5 Concentration Dependence of dATP on Nucleotide Incorporation...122
4.6 Concentration Dependence of dATP on Nucleotide Incorporation...123
4.7 Quantitative Effect of the AP site on Incorporation via A-rule.......124
4.8 Quantitative Effect of the AP site on Incorporation via Loop-out...125
4.9 The Chemical Structure of dPTP............................................126
4.10 Concentration Dependence of Observed dPTP Rate Constant......127
4.11 Incorporation of dATP into 21/41AP Catalyzed by Dpo4............128
5.1 Mutation Spectrum of AP Site Bypass Catalyzed by Dpo4 at 37 °C.....160
5.2 Mutation Spectrum of AP Site Bypass Catalyzed by Dpo4 at 75 °C....161
5.3 BmgB I Digestion Assay.......................................................162
6.1 The D-1 DNA Substrate......................................................204
6.2 Concentration Dependence of the Rate Constant of dTTP into D-1...205
6.3 Concentration Dependence of the Rate Constant of dTTP into D-1...206
6.4 Pre-steady State and Steady-state Kinetics of dTTP Incorporation ....207
6.5 Steady-state Kinetics of dTTP Incorporation into D-1 at 56 °C........208
6.6 Active Site Titration of Dpo4 at 56 °C.................................209
6.7 Elemental Effect on the Rate for Correct Incorporation at 56 °C.....210
6.8 Elemental Effect on the Rate for Incorrect Incorporation at 56 °C...211
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.9</td>
<td>Measurement of the Dissociation Rate Constant for E'-DNA-dNTP...212</td>
</tr>
<tr>
<td>6.10</td>
<td>Pulse-chase and Pulse-quench Experiment at 56 °C ......................213</td>
</tr>
<tr>
<td>6.11</td>
<td>DNA Dissociation from the E-DNA-dNTP Complex at 56 °C ..........214</td>
</tr>
<tr>
<td>6.12</td>
<td>Thermal Denaturation Curve of Dpo4 ........................................215</td>
</tr>
<tr>
<td>6.13</td>
<td>Dpo4 Stability Assay ..........................................................216</td>
</tr>
<tr>
<td>6.14</td>
<td>Temperature Dependence of the Correct Rate and Fidelity ..........217</td>
</tr>
<tr>
<td>6.15</td>
<td>Temperature Dependence of the Incorrect Rate Constant ..........218</td>
</tr>
<tr>
<td>6.16</td>
<td>Activation Energy for Correct Nucleotide Incorporation ............219</td>
</tr>
<tr>
<td>7.1</td>
<td>Gel Mobility Shift Assay for Binding of Dpo4 to B-1 DNA ..........247</td>
</tr>
<tr>
<td>7.2</td>
<td>Equilibrium Dissociation Constant for the Dpo4-B-1 DNA ............248</td>
</tr>
<tr>
<td>7.3</td>
<td>Time Course of Incorporation of 200 µM dATP onto B-1 .................249</td>
</tr>
<tr>
<td>7.4</td>
<td>Time Course Incorporation of 1300 µM dCTP onto B-1 ..................250</td>
</tr>
<tr>
<td>7.5</td>
<td>Time Course of Incorporation of 1300 µM dGTP onto B-1 ...............251</td>
</tr>
<tr>
<td>7.6</td>
<td>Time Course of incorporation of 400 µM dTTP onto B-1 .................252</td>
</tr>
<tr>
<td>7.7</td>
<td>Time Course of Incorporation of 2 µM dPTP onto B-1 ...................253</td>
</tr>
<tr>
<td>7.8</td>
<td>Time Course of Incorporation of 1300 µM dATP onto B-1A .............254</td>
</tr>
<tr>
<td>7.9</td>
<td>Concentration Dependence on the Rate of dATP Incorporation ..........255</td>
</tr>
<tr>
<td>7.10</td>
<td>Concentration Dependence on the Rate of dPTP Incorporation ..........256</td>
</tr>
<tr>
<td>7.11</td>
<td>Crystal Structure of Dpo4-Blunt-end DNA with ddATP ..................257</td>
</tr>
<tr>
<td>7.12</td>
<td>Crystal Structure of Dpo4-Blunt-end DNA with ddATP ..................258</td>
</tr>
<tr>
<td>7.13</td>
<td>Proposed Mechanisms for Blunt-end Addition of dPTP ..................259</td>
</tr>
<tr>
<td>7.14</td>
<td>Proposed Mechanisms for Blunt-end Addition of dATP ..................260</td>
</tr>
<tr>
<td>7.15</td>
<td>Comparison of the Structures of Dpo4 Ternary Complexes ..........261</td>
</tr>
<tr>
<td>8.1</td>
<td>Schematic Representations of DNA Polymerases λ and β ...............297</td>
</tr>
<tr>
<td>8.2</td>
<td>Purification of the N-terminal Truncated Human Polymerase λ .......298</td>
</tr>
</tbody>
</table>
8.3 Optimization of Mg2+ Concentration on the activity of tPolλ........299
8.4 Optimization of NaCl Concentration on the activity of tPolλ........300
8.5 Optimization of pH on the activity of tPolλ............................301
8.6 Measurement of the dNTP koff Rate Constant and kp..............302
8.7 Measurement of the DNA Dissociation Rate Constant and kp........303
8.8 Concentration Dependence of the Rate of Correct Incorporation.....304
8.9 Concentration Dependence of the Rate of Correct Incorporation....305
8.10 Concentration Dependence of the Rate of Incorrect Incorporation...306
8.11 Concentration Dependence of the Rate of Incorrect Incorporation...307
8.12 Progression of Incorporation for 100 µM dCTP into D-6............308
8.13 Progression of Incorporation for 100 µM dCTP into D-1.............309
8.14 Progression of Incorporation for 100 µM dCTP into D-7.............310
8.15 Progression of Incorporation for 100 µM dCTP into D-8.............311
8.16 Progression of Incorporation for 100 µM dCTP into D-12...........312
8.17 Progression of Incorporation for 100 µM dTTP into D-12...........313
9.1 Schematic Representation of fPolλ, dPolλ, tPolλ, and Polβ...........335
9.2 Single-nucleotide Incorporation into Gapped DNA Substrates........336
9.3 dTTP Incorporation into 5'-[32P]-labeled D-1.........................337
9.4 dTTP Incorporation into 5'-[32P]-labeled D-1.........................338
9.5 Comparison of the Substrate Specificity by fPolλ and tPolλ...........339
9.6 Comparison of the Substrate Specificity by fPolλ and tPolλ...........340
9.7 Comparison of Incorporation Efficiencies of fPolλ and dPolλ..........341
10.1 Schematic Representation of the Domains of Polλ and Polβ.........361
10.2 dRP-DNA Substrate...........................................................362
10.3 H-DNA Substrate.............................................................363
10.4 D-DNA Substrate.............................................................364
10.5 D1-OH Substrate........................................................................365
10.6 D-1N Substrate........................................................................365
10.7 Concentration Dependence of a Correct Incorporation Rate.......366
10.8 Concentration Dependence of a Correct Incorporation Rate.......367
11.1 Mutation Spectrum of AP Bypass Catalyzed by Polymerase η......396
11.2 Mutation Spectrum of AP Bypass Catalyzed by Polymerase κ......397
11.3 Comparison of the Linker Sequences of Dpo4 and Dbh.................398
11.4 DNA Binding Affinity for the Dpo4 Linker Mutants.......................399
11.5 dTTP Incorporation into Gapped D-1 Substrate by Polymerase λ.....400
11.6 Single-nucleotide Gapped DNA Substrates.................................401
11.7 Pulse-chase and Pulse-quench Experiments.................................402
11.8 Control and AP Lesion Containing DNA Substrates.......................403
11.9 Elongation of 17/41-mer by Human fPolλ..................................404
11.10 Elongation of 17/41-mer by Human tPolλ..................................405
11.11 Elongation of 17/41-mer by Human Polymerase μ........................406
11.12 Elongation of 17/41-mer by Human TdT..................................407
11.13 Elongation of 17/41-mer by Human Polymerase β.......................408
11.14 Mutation Spectrum of AP Bypass Catalyzed by Polymerase β........409
11.15 Control and Cisplatin-modified DNA Substrates.........................410
11.16 Elongation of 17/44-mer by Human Polymerase β.......................411
11.17 Elongation of 17/44-mer by Human Polymerase μ.......................412
11.18 Elongation of 17/44-mer by Human fPolλ..................................413
# LIST OF SCHEMES

<table>
<thead>
<tr>
<th>Schemes</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 DNA Polymerase Catalyzed Reaction</td>
<td>12</td>
</tr>
<tr>
<td>1.2 Minimal Kinetic Mechanism for a DNA Polymerase</td>
<td>13</td>
</tr>
<tr>
<td>2.1 Minimal Kinetic Mechanism for a DNA Polymerase</td>
<td>52</td>
</tr>
<tr>
<td>2.2 Detailed Minimal Kinetic Mechanism for dNTP Incorporation</td>
<td>53</td>
</tr>
<tr>
<td>2.3 The “dNTP”-stabilized Misalignment Model</td>
<td>54</td>
</tr>
<tr>
<td>3.1 Structures of the Sp and Rp Isomers of dTTP(\alpha)</td>
<td>91</td>
</tr>
<tr>
<td>3.2 Detailed Kinetic Mechanism for dNTP incorporation by Dpo4</td>
<td>92</td>
</tr>
<tr>
<td>4.1 Proposed Mechanisms to Bypass an AP Lesion</td>
<td>139</td>
</tr>
<tr>
<td>4.2 Proposed Minimal Kinetic Mechanism of AP Site Bypass by Dpo4</td>
<td>140</td>
</tr>
<tr>
<td>5.1 Two Primary Pathways Used by Dpo4 to Bypass and AP Site Lesion</td>
<td>167</td>
</tr>
<tr>
<td>5.2 Diagram of the Short Oligonucleotide Sequencing Assay</td>
<td>168</td>
</tr>
<tr>
<td>5.3 Proposed Mechanisms for ‘AP Bypass Products’</td>
<td>169</td>
</tr>
<tr>
<td>5.4 Proposed Mechanisms for ‘AP Bypass Products’</td>
<td>171</td>
</tr>
<tr>
<td>5.5 Proposed Mechanisms for ‘AP Bypass Products’</td>
<td>173</td>
</tr>
<tr>
<td>5.6 Proposed Mechanisms for ‘AP Bypass Products’</td>
<td>174</td>
</tr>
<tr>
<td>5.7 Possible Sequence Contexts for the BmgB I Digestion Assay</td>
<td>176</td>
</tr>
<tr>
<td>6.1 Proposed Kinetic Mechanisms for dNTP Incorporation by Dpo4</td>
<td>223</td>
</tr>
<tr>
<td>7.1 Structure of dPTP</td>
<td>265</td>
</tr>
<tr>
<td>8.1 Minimal Kinetic Mechanism for dNTP Incorporation into DNA</td>
<td>317</td>
</tr>
<tr>
<td>9.1 Minimal Mechanism for dNTP Incorporation into DNA</td>
<td>344</td>
</tr>
<tr>
<td>10.1 Reaction Sequence of Uracil-Mediated Short-Patch BER</td>
<td>371</td>
</tr>
</tbody>
</table>
10.2 Minimal Kinetic Mechanism for dNTP Incorporation onto D-DNA.....372
11.1 Detailed Kinetic Mechanism for dNTP Incorporation by fPol...........418
CHAPTER 1

INTRODUCTION

1.1 Introduction

DNA polymerases are the enzymes responsible for the replication of an organism’s genetic information. After initially being discovered in the 1950’s by Arthur Kornberg, scientists have devoted decades of research into characterizing these fascinating enzymes. These scientists have broadly classified DNA polymerases into six families based on phylogenetic relationships and include the A-, B-, C-, D-, X-, and Y-families. While all DNA polymerases share the same overall organization of their respective structural domains, the architecture of the enzyme active site, the requirement of accessory proteins, in addition to rate constants for performing catalysis, polymerase fidelity, and shear size differ greatly among this class of enzymes. Moreover, besides the three conserved catalytic carboxylic residues found in the active sites of all DNA polymerases, there is little sequence similarity shared between members of different polymerase families, let alone, within the same family. Yet, these enzymes are capable of performing the same fundamental reaction involving the addition of a nucleotide triphosphate (dNTP) to a DNA substrate (DNA_n) via an electrophilic attack of the terminal 3’-hydroxyl group of the DNA substrate on the weak bond between the alpha- and beta-phosphates of a dNTP, performed in the confines of the polymerase active site (Scheme 1.1 and 1.2) resulting in a DNA molecule elongated by one nucleotide (DNA_{n+1}).
Interestingly, DNA polymerases have evolved to perform a variety of diverse biological functions within the cell, including, but not limited to high-fidelity processes such as DNA duplication and the repair of DNA damage in addition to the often error-prone, low-fidelity processes of somatic hypermutation and DNA lesion bypass.

Polymerase families. The A-family DNA polymerases are represented by *Escherichia coli* (*E. coli*) DNA polymerases I. These enzymes are characterized as high-fidelity DNA polymerases that typically possess both 3’ → 5’ and 5’ → 3’ exonuclease activity. Although found in bacteria, plants, and the mitochondria of higher Eukaryotes, this family also has viral members (1). Eukaryotic DNA polymerase gamma (γ) is a well characterized polymerase of this family that is solely responsible for the replication and repair of the mitochondrial genome (2).

In an analogous manner as described above, *E. coli* DNA polymerase II defines the B-family DNA polymerases with members in Archaea, Eukaryotes, and viruses (1). Most of these family members are found in Eukaryotes and include DNA polymerases alpha (α), DNA polymerases delta (δ), DNA polymerases epsilon (ε), DNA polymerases zeta (ξ). DNA polymerase α functions as a primase (3) and possesses only 3’ → 5’ exonuclease activity, while DNA polymerases δ and ε are responsible for the synthesis of the leading and lagging strands and possess both 3’ → 5’ and 5’ → 3’ exonuclease activities. DNA polymerase ξ is unusual in that it has a significantly lower fidelity in comparison to the high-fidelity DNA polymerases α, δ, and ε. This is due to its lack of 3’ → 5’ exonuclease activity (4). Intriguingly, DNA polymerase ξ has been shown to fairly efficiently extend from mispaired primer termini (5-7). Due to this observation, coupled with the fact that it is unable to incorporate nucleotides opposite DNA lesions like other low-fidelity members (5-7), DNA polymerase ξ is thought to function in tandem with other Y-family
DNA polymerases (see below) to bypass DNA lesions in Eukaryotes by performing the extension step after the Y-family member incorporates a nucleotide opposite the lesion (reviewed in (8)).

The C-family DNA polymerases are exclusively bacterial and are homologous to *E. coli* DNA polymerase III. These DNA polymerases possess 3' → 5' exonuclease activity and typically function in genome replication. *E. coli* DNA polymerase III, the most well-studied member of this family, is the major replicative polymerase in this organism (9).

The D-family DNA polymerases are high-fidelity enzymes that have only been identified in Archaea. These replicative DNA polymerases possess proofreading 3' → 5' exonuclease activity and replicate DNA in a template-dependent fashion. One of the most well-known D-family polymerases is *Pyrococcus furiosus* (*Pfu*) DNA polymerase for its use in the technique of PCR (10).

The X-family DNA polymerases are a group of enzymes from a larger family of nucleotidyl transferases. Some of these enzymes are not DNA polymerases in the traditional sense, in that they can catalyze template-independent nucleotidyl transferase activity (11). There are several notable Eukaryotic family members including DNA polymerase beta (β), DNA polymerase lambda (λ), DNA polymerase mu (μ), terminal deoxynucleotidyl transferase (TdT), DNA polymerases sigma 1 (σ1) and sigma 2 (σ2) and yeast DNA polymerase IV. These family members can also be found in Bacteria, Archaea, and viruses.

Finally, the Y-family DNA polymerases are a class of low-fidelity enzymes that are all devoid of 3' → 5' exonuclease activity and possess very low processivity with DNA (12-
18). These enzymes, which are conserved in all three life domains yet are not found in viruses, have been shown to have more open and solvent accessible active sites (19-24) that can accommodate and facilitate the bypass of often bulky DNA lesions that halt the synthesis of the high-fidelity replicative DNA polymerases of the other polymerase families. Typically higher eukaryotes have four Y-family members, DNA polymerase eta (η), DNA polymerase iota (ι), DNA polymerase kappa (κ), and REV1. The most well-known of these is DNA polymerase η which, when inactivated in humans, leads to a genetic condition known as Xeroderma Pigmentosum Variant which is characterized by a higher sensitivity to sunlight-induced skin cancer (12,14,25). This enzyme has been shown to bypass sunlight-induced DNA lesions in an error-free fashion (14,25).

1.2 Structural and Mechanistic Analysis of DNA Polymerases

Structure and mechanism. Crystal structure analysis of numerous DNA polymerases illustrate that all DNA polymerases have a similar global ternary structural organization (23,26-31) although they lack significant sequence homology. This similarity in global structure between the DNA polymerase families in all three domains of life and in viral systems, suggests that these enzymes are derived from a common ancestor (32). With the exception of the Y-family members that contain a fourth structural domain (known as the ‘little finger’ domain in Sulfolobus solfataricus DNA polymerase IV (Dpo4), ‘wrist’ domain in Sulfolobus solfataricus DinB homolog DNA polymerase (Dbh), and the ‘polymerase-associated domain’ (PAD) in yeast DNA polymerase η), all structure-known DNA polymerases consist of three distinct domains (palm, thumb, and finger) that are organized and resemble a human right-hand with the finger and thumb domains oriented
so as to “grip” the DNA while forming an active site cleft with the palm domain where the reaction chemistry occurs (Figure 1.1). The active site formed by these three conserved domains is similar in all DNA polymerases at the structural level (33). In the high-fidelity DNA polymerases the finger domain is involved in the proper positioning of the incoming nucleotide while the thumb domain facilitates DNA binding and processivity (34). The palm domain in all DNA polymerases contains at least two, but typically three conserved carboxylic acid moieties that are requisite to catalyze nucleotide incorporation. These active site residues, in coordination with two divalent metal ions, have been shown to assist in the positioning of the nucleotide in the ground-state and facilitate the overall catalysis through electrostatic stabilization. The catalytic contribution fostered by DNA polymerases involves the incorporation of a nucleoside monophosphate (dNMP) into a polymer of nucleic acid via an in-line nucleophilic attack of the 3’-hydroxyl of the DNA substrate on the electron-deficient α-phosphate of the incoming nucleotide via a semi-associative mechanism (35). This reaction is catalyzed by DNA polymerases utilizing the two-metal-ion mechanism as proposed by Steitz (34). These two divalent cations, commonly referred to as metal A and metal B in the literature, have been shown to perform important functions during the catalytic cycle. Metal A facilitates the incorporation of a dNTP by stabilizing the negative charge of the α-phosphate of the nucleotide and activates the 3’-hydroxyl for the attack of the α-phosphate by lowering the affinity of the hydroxyl for its hydrogen atom, while metal B functions to orient the β- and γ-phosphates of the nucleotide in a position that is competent for catalysis (27,28,36) and stabilizes the pyrophosphate release upon product formation. These two cations, typically thought to be Mg$^{2+}$ ions in the physiological setting, reduce the net negative charge of the phosphate moieties of the incoming nucleotide by forming a penta-coordinated transition state with these phosphate residues and through interaction with the active site carboxylic acid moieties. This mechanism
effectively stabilizes the highly charged environment within the DNA polymerase active site during catalysis resulting in a decrease in the activation energy barrier. Amino acid residues in the DNA polymerase active site are required for proper positioning of these metal ions. Upon “matched” nucleotide binding, a rate-limiting conformational change occurs which results in a “closed” conformation capable of orienting the 3’-hydroxyl group of the DNA substrate and the α-phosphate of the dNTP for catalysis. A “mismatched” basepair prevents the necessary conformational change for efficient nucleotide incorporation, resulting in a significantly slower rate constant for nucleotide incorporation. DNA polymerases in the A-, B-, and X-families have been demonstrated to have significant structural changes especially in their finger domains, from the binary crystal structures to the ternary crystal structure (28,29,31,37,38) upon binding the “matched” dNTP. However, while the precise nature of the aforementioned rate-limiting conformational change remains definitively unresolved, recent kinetic (39), structural (40), modeling (41), and fluorescence analyses (42) have provided fairly compelling evidence that this rate-limiting conformational change may be due to a local structural adjustment in the DNA polymerase active site as opposed to the traditional model involving a large rotation of the polymerase finger domain.

**Proofreading exonuclease activity.** To increase the overall incorporation efficiency, some DNA polymerases have evolved catalytic domains that function to proofread each nucleotide incorporated into the DNA duplex, removing misincorporated nucleotides. This proofreading activity, known as 3’ → 5’ exonuclease activity, was first discovered in *E. coli* DNA polymerase I (43) but is known to be a characteristic of the high-fidelity replicative DNA polymerases that are responsible for genome replication. The *in vivo* implications of this exonuclease activity has been illuminated from analysis of transgenic mice lacking the exonuclease alleles for the replicative DNA polymerase δ which reveal a
dramatic increase in susceptibility to cancer within the first year of life (44). Similar results indicating an increase in the rate of mutagenesis have been observed when the exonuclease activities of *E. coli* DNA polymerase III (45) and T7 DNA polymerase (46) have been functionally reduced. In each turnover of the catalytic cycle during continuous synthesis, the DNA polymerase partitions between forward synthesis, pyrophosphorolysis, its exonuclease activity, and DNA dissociation, with the probability of each comprised as a function of the relative rate constant for each individual reaction (47). During processive synthesis, the kinetic partitioning of high-fidelity replicative polymerases favors forward synthesis, while only rarely dissociating or partitioning the DNA into the exonuclease domain. However, once a “mismatched” nucleotide is incorporated, the polymerization rate decreases by several (four to five) orders of magnitude with a concomitant increase in the rate of exonuclease activity by one to two orders of magnitude. Thus the basis of enhanced selectivity offered by the exonuclease domain is due to an inhibition of continued synthesis on a DNA substrate containing a terminal “mismatch” which increases the likelihood of favorable partitioning to the exonuclease domain. From a structural perspective, a “mismatched” basepair prevents the finger domain of high-fidelity DNA polymerases from forming the “closed” conformation in order to bind to the next incoming dNTP. The 3’-end of the DNA is therefore able to partition into the exonuclease active site, typically located at least 30 Å from the polymerase active site by a process thought to be facilitated by the thumb domain.

*The basis of nucleotide binding.* Nucleotide binding is a complex set of interactions facilitated to differing degrees in these enzymes by hydrogen bonding between the template base in the DNA polymerase active site and the incoming nucleotide, as well as the base-stacking between the incoming nucleotide and the terminal basepair of the
elongating stand, stabilization via active site amino acid residues, and even through electrostatic interactions from active sites divalent ions. Arguably, the most influential amongst these interactions in nucleotide binding is hydrogen bonding and base-stacking. The “matched” basepair will hydrogen bond with the template base so as to be positioned in the enzyme active site with the proper Watson-Crick geometry for nucleotide incorporation. These “matched” Watson-Crick basepairs are characterized by a pseudo-2-fold axis between the nitrogenous bases, a symmetrical orientation of the glycosidic bond between the bases and sugar relative to the vector between the C1’ angles (λ angle), and possess the same C1’ distances (48). This base pairing scheme allows the “matched” basepair to be selected for incorporation at least partially due to the geometric restraints imparted from the DNA polymerase active site (49). Nucleotide incorporation efficiency is enhanced via geometric selection processes resulting from steric interactions from the polymerase domains within the nucleotide binding pocket. These interactions check for the proper basepairing complementarity for “matched” basepairs and select against “mismatched” basepairs via a conformational closing of the finger domain of replicative DNA polymerases upon the nucleotide binding pocket. The proper closing and subsequent stabilization of this conformational change in the context of a “matched” basepair facilitates the proper alignment of the α-phosphate and the 3’-hydroxyl of the DNA for catalysis within the active site. Structural analysis of “mismatched” DNA basepairs reveals an asymmetric orientation of the λ angle in addition to different C1’ distances than that observed with a “matched” Watson-Crick basepair (50). The “mismatched” basepair is not stabilized by the conformational change in the finger domain due to the lack of proper geometrical alignment and thus will be less likely to obtain the proper alignment in the active site for efficient catalysis (28,36). From an entropic perspective, the process of base pairing is energetically favorable due to the release of water molecules (increased entropy) coupled with the low energetic cost of
formation of subsequent hydrogen bonds (51). Alternatively, base-stacking interactions, which are thought to account for the overall stability of the DNA helix (52,53), have been suggested to contribute to moderate DNA polymerase fidelity in the absence of hydrogen bonding interactions between the incoming nucleotide and the template base (51). However, not all DNA polymerases require hydrogen bond formation for efficient catalysis, although it has been shown to increase the overall fidelity of these enzymes (see review by E. Kool, (51)).

1.3 Focus of Dissertation

This dissertation will focus on two DNA polymerases, one from the X-family (human DNA polymerase λ) and one from the Y-family (Sulfolobus solfataricus DNA polymerase IV, Dpo4). Both enzymes have been recently identified in their respective organisms and, as such, are just beginning to be characterized. While human DNA polymerase λ shares 33% sequence identity with its X-family homolog human DNA polymerase β, an enzyme known to function in base-excision repair (BER) pathways in vivo, many lines of evidence have shown that human DNA polymerase λ is not a surrogate enzyme for human DNA polymerase β. Besides a drastically different domain organization, recent biochemical, cellular and animal model studies have suggested that human DNA polymerase λ likely functions in immunoglobulin V(D)J recombination and non-homologous end-joining in addition to base-excision repair (BER) in vivo. On the other hand, Dpo4 is a lesion bypass polymerase that is hypothesized to function in vivo in the global bypass of DNA lesions that stall the progression of high-fidelity DNA polymerases which have been shown to accomplish the majority of the genome
replication. These Y-family DNA polymerases are characterized by relatively open and solvent-accessible enzyme active sites which foster comparatively few contacts between the protein and the replicating basepair and as such, are 100 to 100,000 times more error prone when compared to the high-fidelity replicative DNA polymerases. From a mechanistic standpoint, these Y-family DNA polymerases temporarily replace the replicative DNA polymerase at the site of DNA damage and bypass the lesion in either an error-free or error-prone fashion prior to dissociation from DNA in order to allow the replicative DNA polymerase to continue genome duplication. While Dpo4 and DNA polymerase λ ultimately possess the same catalytic capability, detailed characterization of each enzyme reveals evidence for the intriguing caveat that individual DNA polymerases are mechanistically and biologically distinct.
1.4 Figures and Schemes

Figure 1.1. Ternary structure of Dpo4 with an undamaged DNA substrate and incoming nucleotide (dNTP). Conserved structural domains of Dpo4 are shown above and include the palm domain (red), the finger domain (blue), and the thumb domain (green). In addition Dpo4 and the Y-family DNA polymerases have a fourth structural domain known as the little finger (purple). This figure is taken from (23).
Scheme 1.1. DNA polymerase catalyzed reaction.
Scheme 1.2. Minimal kinetic mechanism for a DNA polymerase.

\[ E \cdot \text{DNA}_n + \text{dNTP} \xrightleftharpoons[K_d]{K_p} E \cdot \text{DNA}_n \cdot \text{dNTP} \xrightarrow{k_p} E' \cdot \text{DNA}_{n+1} \cdot \text{PPi} \]
1.5 References


14. Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M.,

15. Levine, R. L., Miller, H., Grollman, A., Ohashi, E., Ohmori, H., Masutani, C.,

16. Bebenek, K., Tissier, A., Frank, E. G., McDonald, J. P., Prasad, R., Wilson, S. H.,

Genet 25, 458-461


Nature 430, 377-380

Science 309, 2219-2222

21. Trincao, J., Johnson, R. E., Escalante, C. R., Prakash, S., Prakash, L., and

22. Uljon, S. N., Johnson, R. E., Edwards, T. A., Prakash, S., Prakash, L., and


Nat Struct Biol 8, 984-989

278, 18767-18775

Science 264, 1891-1903

Biochemistry 36, 11205-11215

Nature 391, 251-258


CHAPTER 2

PRE-STEADY STATE KINETIC STUDIES OF THE FIDELITY OF
SULFOLOBUS SOLFATARICUS DNA POLYMERASE IV AT 37 ºC

2.1 Introduction

Cellular DNA is continually under assault from a wide array of DNA-damaging agents such as reactive oxygen species, UV light, ionizing radiation, and radio-mimetic chemicals (1). Various forms of DNA lesions stall the replication machinery by impeding the progression of replicative polymerases. Over the past few years, a novel superfamily of DNA polymerases, now known as the Y-family, has been identified and collectively demonstrate an ability to bypass DNA lesions, either in an error-free or error-prone manner (2-4). The Y-family distributes amongst the three kingdoms of life and includes *Escherichia coli* DNA polymerases IV (also known as DinB) and V (also known as UmuC), yeast DNA polymerase η (yPolη), human DNA polymerases η (hPolη), κ (hPolκ), and ι (hPolι), *Sulfolobus acidocaldarius* Dbh, and *Sulfolobus solfataricus* DNA polymerase Dpo4 (5). Dpo4, the focus of this paper, is shown to preferentially traverse abasic sites, followed by cisplatin-DNA adducts, *cis-syn* thymine-thymine (TT) dimers, 6-4 TT dimers, and acetyl aminofluorene-DNA adducts (6).

Dpo4 is a thermostable DNA polymerase which can catalyze synthesis at 37 ºC, exhibiting a misincorporation fidelity in the range of $10^{-3}$ to $10^{-4}$ with undamaged DNA.
substrates as revealed by steady-state kinetic analysis (6,7) and by the implementation of a forward mutation assay (7). The latter assay scores a variety of substitution and frameshift errors generated during the copying of a lacZ template in a single-stranded gap of M13mp2 DNA. Surprisingly, the forward mutation assays also reveal that Dpo4 generates C:C mismatches at an unusually high rate (4 x 10^{-2}) and preferentially at cytosine flanked by a 5’-template guanine (7). Interestingly, Dbh, a Dpo4 homolog from *S. acidocaldarius*, makes C:C mismatches most frequently (8). Except for the unusually high C:C rate, the error rates of the other eleven mismatches for Dpo4 are remarkably similar to those for hPolκ (7,9). Dpo4 also demonstrates very low frameshift fidelity and frequently generates deletions of even non-iterated nucleotides, especially cytosine flanked by a 5’-template guanine (7). The frameshift error rate of Dpo4 (7) is similar to those of hPolκ (9) and hPolη (10), while being much higher than human DNA polymerases α (11), β (12), and γ (13). Interestingly, *E. coli* DNA polymerase IV (14,15) and *S. acidocaldarius* Dbh (16) are also found to predominantly generate a -1 frameshift at a pyrimidine flanked by a 5’-template guanine. Based on these findings, a “dNTP-stabilized” misalignment mechanism is hypothesized to be responsible for making -1 frameshift mutations for these Y-family polymerases (7). This mechanism suggests the incoming nucleotide, dCTP, forms a correct basepair with a downstream template base G, on a “looped out” template strand instead of forming an incorrect basepair with the first available template base (14). A mutagenic intermediate, important for verifying this mechanism, is observed in the crystal structure of Dpo4 in a ternary complex with DNA and a mismatched incoming nucleotide (17).

In addition to Dpo4, the fidelity of several Y-family polymerases including yPolη (16,18), hPolη (10,19), hPolκ (9,20), hPolτ (21-23), *E. coli* polymerase IV (14), and Dbh (16) have been estimated by steady-state kinetic analyses. Although the steady-state
kinetic methods can be used to estimate the polymerase fidelity and substrate specificity, the kinetic and thermodynamic basis for the fidelity of the Y-family polymerases has not been established due to the inherent limitations of steady-state kinetics (24). Like other polymerases, the mechanistic basis of enzyme fidelity can only be established by employing pre-steady state kinetic methods (24). So far, pre-steady state kinetic studies on the novel Y-family have been used to determine the fidelity of yPolη based on one correct and one incorrect nucleotide incorporation (25), but the other fourteen possible incorporations were not analyzed. In this paper, we chose Dpo4 as a model enzyme from the Y-family and employ pre-steady state kinetic techniques to study, for the first time, the replication fidelity and its kinetic and thermodynamic basis, based on all sixteen possible nucleotide incorporations into undamaged DNA substrates. The “dNTP-stabilized” misalignment mechanism will also be examined. Additionally, we established the kinetic mechanism of DNA polymerization catalyzed by Dpo4 (26).

2.2 Materials and methods

Materials. These chemicals were purchased from the following companies: \( \gamma^{32} \text{P} \)ATP, Perkin Elmer Life Sciences (Boston, MA); dNTPs, Gibco-BRL (Rockville, MD); calf intestine alkaline phosphatase, Fermentas (Hanover, MD); T4 polynucleotide kinase, USB (Cleveland, OH); Biospin columns, Bio-Rad Laboratories (Hercules, CA).

Cloning and purification of Dpo4. The dpo4 gene from \textit{S. solfataricus} was cloned into the Nde I/Xho I sites of pET22b. Dpo4 fused to a C-terminal His\textsubscript{6} tag was expressed in \textit{E. coli} strain BL21(DE3). An overnight culture of \textit{E. coli} expression strain BL21(DE3)
carrying this plasmid was used to inoculate Luria-Bertani media containing 100 μg/ml ampicillin and the cells were grown at 22 °C. After the OD$_{600}$ reached to 0.5, cultures were induced with 0.2 mM IPTG and incubated at 22 °C until OD$_{600}$ reached 1.7. Cells were harvested (4000 rpm, 15 min) and resuspended in buffer A (10 mM KHPO$_4$, pH 7.0, 50 mM NaCl, 10 mM MgAc$_2$, 10 % glycerol, 0.1% 2-mercaptoethanol). Cells were then lysed by passing through a French Press cell at 16,000 psi twice and the resulting lysate was cleared by spinning in an ultracentrifuge (35,000 rpm, 40 min). Cleared lysate was incubated at 78 °C for 12 minutes to precipitate thermolabile E. coli proteins which were then removed by ultracentrifugation. The supernatant containing predominantly Dpo4 was pooled and incubated overnight at 4 °C with nickel-NTA superflow resin (Qiagen). The supernatant was removed by centrifugation in a swing-bucket centrifuge (2500 rpm, 10 min) and the Dpo4-bound nickel resin was packed into a column. Bound proteins were eluted through a linear gradient of 25 to 350 mM imidazole in buffer B (10 mM KHPO$_4$, pH 7.0, 0.35 M NaCl, 5 to 500m mM imidazole, 2.5 mM MgAc$_2$, 10 % glycerol, 0.1% 2-mercaptoethanol). Dpo4-containing fractions were pooled and applied to a MonoS column (Pharmacia Biotech) and eluted via a three step gradient (50 to 240 mM NaCl, 240 mM NaCl, 240 to 1000 mM NaCl) in buffer C (50 mM Tris-Cl, pH 7.0, 10 to 1000 mM NaCl, 2 mM EDTA, 10% glycerol, 0.1% 2-mercaptoethanol). Fractions containing Dpo4 were pooled, dialyzed against buffer D (50 mM Tris-acetate, pH 7.5, 50 mM NaAc, 1 mM DTT, 0.5 mM EDTA, and 10% glycerol) twice, and concentrated using a Centriprep YM-30 (Millipore). The concentrated Dpo4 was finally dialyzed against the storage buffer (50 mM Tris-acetate, pH 7.5, 50 mM NaAc, 1 mM DTT, 0.5 mM EDTA, and 50% glycerol). Dpo4 was purified to >95% purity based on staining SDS-PAGE gels with Coomassie Blue R-250 (Figure 2.1) and resulted in a yield of approximately 50 mg from 9 liters of initial E. coli culture. The concentration of the purified Dpo4 was
measured spectrophotometrically at 280 nm using the calculated extinction coefficient of 24,058 M⁻¹cm⁻¹ before storing small 20 - 50 µl aliquots of the Dpo4 preparation at -80 ºC.

*Synthetic oligonucleotides.* The DNA substrates listed in Table 2.1 were purchased from either Integrated DNA Technologies or TriLink Biotechnologies, purified by denaturing polyacrylamide gel electrophoresis (18% acrylamide, 8 M urea) and the concentration was determined by UV absorbance at 260 nm with the following extinction coefficients (M⁻¹cm⁻¹): 21-mer, ε = 194 100; D-1 41-mer, ε = 396 700; D-6 41-mer, ε = 394 200; D-7 41-mer, ε = 392 200; D-8 41-mer, ε = 389 500; D-12 41-mer, ε = 392 400.

*Labeling and annealing of the DNA substrates.* The primer strand 21-mer was 5’-³²P labeled by incubation with T4 polynucleotide kinase and [γ-³²P]ATP for 1 hour at 37 ºC. After inactivation of the kinase by heating the reaction mixture for 5 minutes at 95 ºC, unreacted [γ-³²P]ATP was subsequently removed by centrifugation via a Biospin-6 column (Biorad). The 5’-³²P labeled 21-mer was then annealed with the corresponding non-radiolabeled DNA 41-mer at a molar ratio of 1.0 to 1.1 respectively, to form the DNA complex of 21/41-mer. Mixtures to be annealed were denatured at 95 ºC for 8 min, and then cooled slowly to room temperature for several hours.

*Buffers.* All experiments using Dpo4, if not specified, were performed in Buffer R containing 50 mM HEPES (pH 7.5 at 37 ºC), 5 mM MgCl₂, 50 mM NaCl, and 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA. All reactions were carried out at 37 ºC.

*Rapid-quench experiments.* Experiments were carried out in a rapid chemical quench flow apparatus (KinTek, PA). The apparatus contained a computer-controlled stepping
motor and was modified for small reaction volumes (15 µl). Invariably, the experiments were carried out by allowing enzyme and DNA to preincubate in buffer R. An aliquot of this solution (15 µl) was loaded into a sample loop and rapidly mixed with an equal volume of solution containing nucleotide in buffer R from a second sample loop. The reactions were quenched with 90 µl of 0.37 M EDTA (final concentration) after time intervals ranging from 5 ms to several minutes. All concentrations reported in this paper refer to concentrations during the reaction following rapid mixing.

*Measurement of substrate specificity.* A preincubated solution of Dpo4 and DNA at fixed concentrations was mixed with varying concentrations of Mg\(^{2+}\)•dNTP (5 to 2400 µM) in buffer R at 37 °C to start the reaction. The reaction at each concentration of Mg\(^{2+}\)•dNTP was terminated with 0.37 M EDTA at varying times ranging from ms to minutes. The reaction products were analyzed by sequencing gel analysis. The time course of product formation was fit to a single exponential equation for each concentration of Mg\(^{2+}\)•dNTP (see Data Analysis) to give the observed rate of nucleotide incorporation. The observed rates abstracted from a series of time courses of product formation were plotted against the concentrations of Mg\(^{2+}\)•dNTP and the data were fit to a hyperbola (see Data Analysis) to give the equilibrium dissociation constant of dNTP, \(K_d\), and the maximum rate for incorporation of dNTP, \(k_p\). The substrate specificity, \(k_p/K_d\) was then calculated.

*Product analysis.* Reaction products were analyzed by sequencing gel electrophoresis (17% acrylamide, 8 M urea, 1x TBE running buffer) and quantitated with a Phosphorimager 445 SI (Molecular Dynamics).

*Data analysis.* Data were fit by nonlinear regression using the program KaleidaGraph (Synergy Software). The single turnover experimental data were fit to equation 1 (single
[Product] = A[1 – exp(-k_{obs}t)] \quad (1)

exponential equation), where A represents the reaction amplitude or initial concentration of the binary complex of enzyme and DNA, and \( k_{obs} \) the observed single turnover rate.

Data from the measurement of \( K_d \) of dNTP were fit to equation 2

\[ k_{obs} = k_p[dNTP]/([dNTP] + K_d) \quad (2) \]

(hyperbola), where \( k_p \) is the maximum rate of dNTP incorporation. The substrate specificity and polymerase fidelity were calculated as \( (k_p/K_d) \) and 

\( (k_p/K_d)_{\text{incorrect}}/(k_p/K_d)_{\text{correct}} \), respectively.

### 2.3 Results

*Protein purification.* Dpo4 (352 amino acid residues and 40.2 kDa) is one of the few Y-family polymerases which can be copiously overexpressed and purified from *E. coli* with high yield (6). We modified the purification protocol published previously (6) by adding a C-terminal hexahistidine tag to wild-type Dpo4 for the convenience of protein purification. The C-terminal hexahistidine-tagged Dpo4 was overexpressed in *E. coli* and purified through a series of three steps: heat denaturation at 78 °C for 12 minutes, Ni-affinity chromatography, and MonoS cation-exchange chromatography (Figure 2.1). The heat denaturation step precipitated most of the *E. coli* proteins, but the thermostable Dpo4 remained soluble. The purified Dpo4 was about 88% active based on the results of an active site titration (26). To evaluate the effect of the C-terminal hexahistidine-tag on the activity of Dpo4, we also purified wild-type Dpo4 lacking the C-terminal hexahistidine-tag following the published purification protocol (6) (data not shown). We subsequently performed kinetic experiments with the purified wild-type Dpo4 and it was found to
incorporate a single nucleotide at a rate constant of $4.7 \pm 0.3 \text{ s}^{-1}$ which is similar to the rate of $3.8 \pm 0.2 \text{ s}^{-1}$ observed with the C-terminal hexahistidine-tagged Dpo4 under the same reaction conditions. These results demonstrated the C-terminal hexahistidine-tag had no effect on the activity of Dpo4. This is not surprising since the crystal structure reveals that the C-terminus of Dpo4 is exposed on the protein surface and is distal from the active site and the enzyme-bound DNA (17). It should be noted that the C-terminal hexahistidine-tagged Dpo4 was used in all experiments reported in this dissertation.

*Optimization of reaction conditions.* Unless otherwise stated, all reactions in this paper were performed at 37 °C, rather than 80 °C, the optimal growth temperature of S. solfataricus. This was done primarily for the following four reasons. First, the rapid chemical quench apparatus used in our studies cannot be operated at 80 °C without causing irreversible damage to the apparatus itself. Second, with regard to DNA melting temperatures, the primer would have to be significantly longer than 21 nucleotides to allow the annealed DNA substrate to remain stable at 80 °C. The products generated by incorporation of single nucleotides into a longer primer would become increasingly difficult to separate from the unreacted substrate by sequencing gel electrophoresis. The difficulty incurred in product analysis would likely hinder data acquisition and thus the overall veracity of results. Third, kinetic data obtained at 37 °C allow us to make direct comparisons with other polymerases assayed at the same temperature. Fourth, we attempted to measure nucleotide incorporation rates at more physiologically relevant temperatures for Dpo4 by performing manual-quench experiments of incorporation of a mismatched nucleotide, due to the aforementioned limitations in instrument operation at elevated temperatures and much slower incorporation of a mismatched nucleotide versus a matched one. Experiments involving the misincorporation of dATP into D-1 were initially carried out at 60 °C (with appropriate adjustment of the pH of the reaction buffer...
to 7.5 at 60 °C), however, even at the shortest time intervals (7 seconds) we could accomplish manually, we found all substrate (21-mer) was converted to product (22-mer). We then decided to carry out the same experiments at a slightly reduced temperature, specifically 50 °C (with appropriate adjustment of the buffer pH to 7.5 at 50 °C). These experiments yielded incorporation of incorrect nucleotides that were on average 24-fold faster than the rates observed at 37 °C (see below). Our observations were not unprecedented. Dbh, a homologous polymerase from *S. acidocaldarius* has been shown to have a 40-fold higher nucleotide incorporation rate at 65 °C than at 22 °C (16). Therefore, if the rate increase for correct nucleotide incorporation is similar as what we observed for incorrect nucleotides, the incorporation of a correct nucleotide catalyzed by Dpo4 at 80 °C should be too fast to be measured by a rapid chemical quench apparatus since the apparatus is limited by its mixing dead time of one to three milliseconds. However, it is not unprecedented to study the kinetics of a polymerase at a non-physiological temperature. For example, the kinetic studies of T7 phage DNA polymerase were performed at 20 °C, rather than 37 °C, and the overall results were not affected except the reaction rates were slower (27-29).

To optimize the reaction conditions, all components were kept constant while MgCl₂ concentration, NaCl concentration, and pH were individually varied. Using a rapid chemical quench apparatus, a preincubated solution of Dpo4 and 5′-[³²P]-D-1 (Table 2.1) was mixed with dTTP at 37 °C to initiate the reaction. Since Dpo4 was in 4-fold molar excess over the DNA substrate, about 90% of D-1 should be complexed to Dpo4 based on the measured affinity of 10.6 nM (26). The reactions were quenched by EDTA at various time intervals. The DNA product 22-mer and unreacted substrate 21-mer were separated by sequencing gel electrophoresis, and quantitated using a Phosphorimager. The data were fit to equation 1 (Materials and Methods). The single turnover rate varied
dramatically with the concentration of Mg$^{2+}$, NaCl, and with the buffer pH. The optimal conditions for Dpo4 polymerization were 5 mM MgCl$_2$ (Figure 2.2), 0 to 75 mM NaCl (Figure 2.3), and pH 7.5 (Figure 2.4). These data demonstrated a significant decrease in the single turnover rates when assayed outside optimum conditions. Thus, the optimized reaction buffer contains 50 mM HEPES, pH 7.5, 5 mM MgCl$_2$, 50 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA (buffer R).

**Substrate specificity of correct nucleotide.** An incoming nucleotide specifically binds at the polymerase active site only after the binding of DNA to the enzyme, to form the E•DNA binary complex (30). The ground-state binding affinity of dTTP ($K_d$) to the binary complex (Scheme 2.1) was measured through the dTTP concentration dependence of the single turnover rate ($k_{obs}$) under conditions where the concentration of enzyme was 4-fold greater than the DNA concentration to ensure almost all of the DNA molecules were bound by Dpo4 (see above). A preincubated solution of Dpo4 (120 nM) and radiolabeled D-1 (30 nM) was mixed with increasing concentrations of dTTP in buffer R. EDTA was then added to the reaction mixtures to quench the polymerization at various time intervals. The DNA products and remaining substrate 21-mer were separated by sequencing gel electrophoresis, and quantitated using a Phosphorimager. The product concentration was plotted against the reaction time and the data were fit to equation 1 (Materials and Methods) to yield a single turnover rate at each concentration of dTTP (Figure 2.5). The single turnover rates were then plotted against dTTP concentrations (Figure 2.6). The data were subsequently fit to equation 2 (Material and Methods) to yield a $k_p$ of 9.4 ± 0.3 s$^{-1}$ for the maximum dTTP incorporation rate, and a $K_d$ of 230 ± 17 μM for the binding of dTTP. Since the value of the $K_d$ was unusually high for a correct nucleotide binding to a polymerase•DNA complex, we repeated the measurement and confirmed that this was not an experimental artifact. We also measured and obtained
similar $K_d$ and $k_p$ when DNA was in molar excess to Dpo4 (data not shown). This suggested the single turnover method used in Figures 2.5 and 2.6 did not affect the measured $K_d$ values. To further discredit a possible experimental anomaly, the affinity of correct nucleotide to human polymerase lambda was assayed independently using the same single turnover method and gave a low $K_d$ of 2.4 $\mu$M (31). These observations are consistent with the finding that free DNA polymerase either does not bind a nucleotide prior to the binding of DNA or binds a nucleotide nonspecifically and weakly (30). Additionally, in the case that Dpo4 was binding nucleotides nonspecifically, the relatively small amount of excess enzyme (90 nM) should not affect the experimental results since the dTTP concentrations were up to four orders of magnitude greater than enzyme excess as seen in Figures 2.5 and 2.6. Alternatively, the $K_d$ value could be affected by the concentration of free Mg$^{2+}$ in the reaction buffer since dTTP also binds Mg$^{2+}$. We repeated the experiments in Figures 2.5 and 2.6 with the concentration of free Mg$^{2+}$ kept constant at 5 mM by adding extra MgCl$_2$ to compensate the binding of Mg$^{2+}$ by dTTP. The results from these experiments showed similar values of $K_d$ and $k_p$ for incorporation of dTTP into D-1 (data not shown). Thus, the affinity of correct nucleotide dTTP to Dpo4, regardless of Mg$^{2+}$ compensation, was more than 10- and 50-fold weaker when compared to replicative polymerases (27-29,32-34). Additionally, the value of the substrate specificity ($k_p/K_d$) was calculated to be 0.0409 $\mu$M$^{-1}$s$^{-1}$ (Table 2.2) for the incorporation of dTTP into D-1.

Similar single turnover measurements were performed with the incorporations of dCTP into D-6, dATP into D-7, and dGTP into D-8 (data not shown) and the measured kinetic parameters $k_p$, $K_d$, and $k_p/K_d$ are listed in Table 2.2. The ground-state binding affinity of dCTP (70 $\mu$M) is 2- to 3-fold higher compared to the binding of the other three correct
nucleotides. This could possibly be due to the downstream template guanine (Table 2.1), leading to two consecutive dCTP incorporations with different polymerization kinetics.

Substrate specificity of incorrect nucleotide. Pre-steady state kinetic analysis of single incorrect dCTP incorporation into D-1 (Table 2.1) was carried out in a manner analogous to those experiments described above for the specificity of correct nucleotide incorporation. Dpo4 (120 nM), preincubated with radiolabeled D-1 (30 nM), was reacted with varying concentrations of dCTP in buffer R. The reactions were quenched with 0.37 M EDTA, analyzed by sequencing gel electrophoresis, and the products were quantitated using a Phosphorimager. The observed single turnover rate at each concentration of dCTP was obtained through the fit of the time course of product formation to equation 1 (Figure 2.7). The observed reaction rates were then plotted against the concentrations of dCTP and the data were fit into equation 2 (Materials and Methods) to yield $k_p$, $K_d$, and substrate specificity ($k_p/K_d$) of $0.013 \pm 0.001$ s$^{-1}$, $1036 \pm 95$ μM, and $1.25 \times 10^{-5}$ μM$^{-1}$s$^{-1}$, respectively (Figure 2.8).

With DNA substrates D-1, D-6, D-7, and D-8, the kinetic parameters for the incorporations of all possible incorrect single nucleotides were determined under single turnover conditions (Table 2.2). Overall, the incorrect nucleotide incorporations have 2- to 10-fold lower ground-state binding affinity ($K_d$) and two to three orders of magnitude lower incorporation rates in comparison to the four correct nucleotide incorporations. Surprisingly, the ground-state binding affinity of the mismatched C:T (D-7) and C:C (D-8) is slightly lower than that of correct A:T and G:C basepairs, although the maximum incorporation rates still differ by several hundred-fold. Interestingly, the downstream template base for these two incorrect incorporations is a guanine, which could form a matched basepair if the incoming nucleotide dCTP skips the first available template base
(Table 2.1). To provide further evidence for such a “dNTP-stabilized” misalignment mechanism (14), we changed the downstream template base G (D-8) to an A (D-12) and measured the kinetic parameters for single dCTP incorporation into D-12 under the same single turnover conditions (data not shown). The sequence of the D-12 template (Table 2.1) prevents the “dNTP”-stabilized dCTP incorporation. The measured $K_d$ ($1228 \pm 265 \mu M$) is about 6-fold larger than the $K_d$ of dCTP into D-8 (192 \mu M), supporting this mutagenic mechanism. Interestingly, the above observations did not occur to the mismatched basepair C:A in D-1 although the next template base is also a guanine (Table 2.1). This observation suggests the “dNTP-stabilized” misalignment mechanism is sequence dependant.

The values of substrate specificity and fidelity for all twelve incorrect incorporations were calculated and listed in Table 2.2. The fidelity of Dpo4 with undamaged DNA is in the range of $10^{-3}$-10^{-4}, similar to what was estimated previously by steady-state kinetic analysis (6).

2.4 Discussion

Although S. solfataricus Dpo4 is an archael polymerase, its structural and functional features are most likely conserved throughout the Y-family and thus, it serves as a model for assessing both the lesion bypass properties and reduced replication fidelity inherent to the family (17). Using the approach of single nucleotide incorporation under single turnover conditions, we measured the substrate specificity of incoming nucleotides and calculated the fidelity of Dpo4 on undamaged DNA substrates. The fidelity was
determined to be in the range of $10^{-3}$-$10^{-4}$ (Table 2.2), similar to what was estimated previously by steady-state kinetic analysis (6,7) and a forward mutation assay (7). The fidelity of a polymerase, defined in equation 3, is inversely proportional to the $K_d$ difference and to the $k_p$ difference between the incorporations of a correct versus an incorrect nucleotide.

$$\text{Fidelity} = \frac{(k_p/K_d)_\text{incorrect}}{(k_p/K_d)_\text{correct}} = \frac{1}{\left[\frac{(K_d)_\text{incorrect}}{(K_d)_\text{correct}}\right]}\left[\frac{(k_p)_\text{correct}}{(k_p)_\text{incorrect}}\right]^{-1}$$

$$= (K_d \text{ difference})^{-1}(k_p \text{ difference})^{-1} \quad (3)$$

Both the ground-state binding affinity $K_d$ and the maximum incorporation rate $k_p$ influence the overall fidelity. The kinetic and thermodynamic basis of the fidelity of Dpo4 is discussed below.

*Contribution of the ground-state binding affinity of an incoming nucleotide to the fidelity of Dpo4.* The ground-state binding affinity of mismatched nucleotides ($K_d$) for Dpo4 is, on average, 9.6-fold (1.1 to 18.1-fold) lower than matched nucleotides (Tables 2.2 and 2.3), corresponding to a $\Delta\Delta G (= -RT\ln K)$ equal to 1.6 kcal/mol (27). The average $K_d$ difference (Table 2.3) and the corresponding $\Delta\Delta G$ value are lower than the average values reported for replicative polymerases such as T7 DNA polymerase (300-fold, 3.9 kcal/mol, 20 °C) (27) and human DNA polymerase γ (hPolγ) (263-fold, 4.0 kcal/mol) (35), but similar to the values for another Y-family member yPolη (5.4-fold, 1.0 kcal/mol) (25), and repair enzymes including rabbit DNA polymerase beta (rPolβ) (24.1-fold, 2.3 kcal/mol) (34) and *E. coli* DNA polymerase I (Klenow) (3.0-fold, 0.75 kcal/mol) (36). Previous analysis of DNA-melting experiments found that the free energy differences between primer terminal correct and incorrect basepairs are less than 1.0 kcal/mol at 37 °C. This suggests the selection of a matched nucleotide by the two Y-family polymerases, Dpo4 and yPolη, and the two repair DNA polymerases, rPolβ and
Klenow, in the ground-state is determined primarily by its intrinsic ability to base pair with the template base, with little contribution from the respective loose polymerase active site as illustrated in the crystal structures (17,37). In contrast, the replicative DNA polymerases significantly stabilize the binding of correct nucleotide through hydrogen bonding and van der Waals interactions in relatively tight enzyme active sites. In these polymerases, the difference in ground-state binding affinity of correct and incorrect nucleotides contributes about 100-fold selection to the fidelity (Table 2.3). In addition, Dpo4 (70-230 μM) binds correct nucleotides with at least 10-fold less affinity than replicative polymerases. For example, the $K_d$ values for correct nucleotide incorporation for hPolγ (35) and T7 DNA polymerase (27) are 0.8 μM and 18 μM, respectively. The low ground-state binding affinity observed for matched nucleotides to Dpo4 further suggests a minimal interaction between Dpo4 and the nascent basepair.

The difference in the ground-state binding affinity and the selection of nucleotides at the initial binding step between the low fidelity and the high fidelity polymerases can be explained by the structural differences at the respective polymerase active sites based on the X-ray crystal structures of Dpo4 (17) and other DNA polymerases (38-42). The ternary structure of Dpo4, DNA, and a matched nucleotide in the “closed” state $E'\bullet$DNA•dNTP (Scheme 2.2) reveals the enzyme active site is relatively open and unusually accessible to solvent in comparison to the active sites of replicative polymerases (17). The weak interaction between Dpo4 and the nascent basepair is expected to be further exacerbated in the “open” state $E\bullet$DNA•dNTP (Scheme 2.2) since the finger and little finger domains of Dpo4 have not undergone the conformational change to form the “closed” state which necessarily increases the processivity and confers relaxed interactions between Dpo4 and the replicating basepair. The energetically favorable entropic contribution from the exclusion of water molecules around the nascent
basepair at the Dpo4 active site is diminished due to its greater solvent accessibility (43). Additionally, the Dpo4 amino acid residues (G41, A42, A44, A57, and G58) in direct contact with the replicating basepair are all small (17), unlike the residues found in high fidelity polymerases, which often involve the aromatic ring of Tyr or Phe, or the guanidinium group of Arg, to form planar stacking interactions with the incoming base (38-42). Except for one water-mediated hydrogen bond, the nascent basepair is unrestrained in either minor or major groove and has no specific polar interactions with Dpo4 (17).

Interestingly, yPolη binds correct nucleotide ($K_d = 2.4 \mu$M) (25) with much higher affinity than Dpo4. yPolη, a Y-family member like Dpo4, is expected to possess a loose active site. However, recently, two structures of Dpo4 in a ternary complex with a DNA substrate containing a cis-syn cyclobutane pyrimidine dimer (CPD) and an incoming nucleotide were crystallized (44). From these structures and the available structure of yPolη alone (37), a theoretical model of a Polη-CPD complex was generated. Structure-based sequence alignment suggests that several amino acid residues in yPolη make its active site less solvent accessible when compared to that of Dpo4 (44). For example, the larger side chains Gln55 and Ile60 residues in yPolη purportedly interact with the replicating basepair in the minor groove while the corresponding residues in Dpo4, specifically Val32 and Ala44, are less capable of such an interaction due to their relatively small side chains. Additionally, Arg73 of yPolη, corresponding to Ala57 of Dpo4, has been purported to help orient the incoming nucleotide through charge interactions conferred via the side chain, while the aliphatic portion of this residue shields the major groove of the replicating basepair and the triphosphate moiety of the incoming nucleotide (44). Overall, these and other subtle structural differences between the Dpo4-CPD complex and the theoretical model of the yPolη-CPD complex may procure
significantly different active site environments. Accordingly, these active site environments may possess characteristics that likely affect nucleotide binding affinities.

Finally, although the reaction temperature in our measurements (37 °C) was much lower than the physiological temperature (80 °C) of *S. solfataricus*, the temperature difference should not affect the equilibrium dissociation constant $K_d = k_{off}/k_{on}$ since both association ($k_{on}$) and dissociation ($k_{off}$) will be equally faster at higher temperatures. Interestingly, the ground-state binding affinity of matched nucleotides to *S. acidocaldarius* Dbh (0.2-1.0 mM) (16) and *Thermus aquaticus* DNA polymerase (35-57 μM) (45) is also low compared to other replicative DNA polymerases. More studies are required to investigate the low affinity of matched nucleotides to these thermostable DNA polymerases.

*Contribution of the rate of nucleotide incorporation to the fidelity of Dpo4.* With Dpo4, the incorporation rates of incorrect nucleotides ($k_p^{\text{incorrect}}$) are approximately two to three orders of magnitude slower than the rates of incorporation of correct nucleotides ($k_p^{\text{correct}}$) (Table 2.2). The contribution to the overall replication fidelity from the $k_p$ difference [$k_p^{\text{correct}}/k_p^{\text{incorrect}}$] is about 100-fold greater than the contribution from the $K_d$ difference [$K_d^{\text{incorrect}}/K_d^{\text{correct}}$] (Table 2.3). The data obtained from similar studies with yPolη reveal a similar predominant contribution to fidelity from the $k_p$ difference (Table 2.3) (25). Likewise, the only other DNA polymerases to be extensively studied using pre-steady state kinetics (Table 2.3), show the same 10- to 100-fold greater contribution to fidelity from the $k_p$ difference. These observations suggest that the steps after the initial binding of nucleotide to form E•DNA•dNTP (Scheme 2.2) provide a more stringent selection for the incorporation of correct versus incorrect nucleotides at the polymerase active site than the initial binding. One of the following two steps has been postulated to
limit the incorporation of a nucleotide: an enzyme conformational change (step 2) or the chemistry step (step 3) (Scheme 2.2). The former deals with the previously proposed “induced-fit” model for replicative DNA polymerases (27,32). In this model, the binding of a correct nucleotide will induce a rate-limiting protein conformational change to form a “closed” conformation followed by a fast chemical incorporation step. The conformational change from the “open” to “closed” state serves as a kinetic barrier to improve selectivity. This step is largely a function of Watson-Crick basepairing geometry required by the polymerase active site (46). The latter model is based on the results gleaned from studies of rPolβ (47), which show that the protein conformational change is actually faster than the slow chemistry step. Our kinetic data for Dpo4 (26) suggest that Step 2 in Scheme 2.2 is rate-limiting for the incorporation of a correct nucleotide while Step 3 limits the incorporation of an incorrect nucleotide. More discussion on the contribution of protein conformational change to the overall fidelity of Dpo4 will be discussed in Chapter 3 (26).

Incidentally, the $k_p$ for a correct nucleotide incorporation by Dpo4 (9.4-16.1 s$^{-1}$) (Table 2.2) is much higher than the rate reported for Dbh (0.06-2.3 min$^{-1}$), a Dpo4 homolog from a $S. acidocaldarius$, measured by Potapova et al. using single nucleotide incorporation under “single turnover” conditions at 22 °C (16). However, this discrepancy cannot be logically reconciled by the different reaction temperatures. Most likely, the slow incorporation rates observed with Dbh are due to the fact that this group did not preincubate the enzyme with DNA prior to mixing with dNTP to initiate reaction time courses (16). The order with which components are mixed to start nucleotide incorporation reactions, significantly affects the pre-steady state kinetic results (30). Moreover, the Dbh reaction buffer (pH 8.5, 10 mM MgCl$_2$, no additional salt) was different from buffer R (pH 7.5, 5 mM MgCl$_2$, 50 mM NaCl) used in our studies.
Differential factor between low and high fidelity polymerases. Surprisingly, all four low fidelity polymerases, including the two Y-family members Dpo4 and yPolη and the two DNA repair polymerases Klenow and rPolβ, when compared to the two high fidelity replicative polymerases T7 DNA polymerase and hPolγ, have a 10- to 100-fold lower $K_d$ selection factor while the values of the $k_p$ difference are similar (Table 2.3). Although the overall contribution to the fidelity from the $k_p$ difference is more substantial, this suggests that the discrimination at the initial nucleotide binding step determines whether a DNA polymerase is either a high fidelity or a low fidelity polymerase (Table 2.3). The selection provided by nucleotide ground-state binding affinity is dictated by the tightness conferred by the respective polymerase active site. Similar proclivities in selection during the subsequent nucleotide incorporation steps after the initial binding step is justifiable since all polymerases use a two-metal-ion based catalytic mechanism, and since all polymerases have structurally similar palm domains containing the three conserved active site carboxylates (one Asp and two Glu residues) (48). However, more DNA polymerases need to be kinetically examined to determine whether or not these conclusions can be globally applied.

The “dNTP”-stabilized misalignment model (Scheme 2.3). The ground-state binding affinity of the mismatched dCTP into D-7 (C:T) or D-8 (C:C) is very similar to that of a correct nucleotide, although the $k_p$ differs by 619- and 276-fold (Table 2.2), respectively. Interestingly, the downstream template base, in both cases, is a guanine. These data suggest the mismatched dCTP somehow base pairs with the downstream guanine to form the canonical C:G basepair by “skipping” the first available template base from the primer 3’-terminus. This type of base pairing necessarily increases the ground-state binding affinity of mismatched dCTP, leading to a subsequent decrease in the polymerase
fidelity (1.1 x 10^-3 - 3.2 x 10^-3) at these specific template sequences. Interestingly, such a mechanism of nucleotide incorporation is not plausible for replicative DNA polymerases, due to their inherent inability to accommodate two template bases at the active site simultaneously. To provide further evidence for this mechanism, we designed a substrate (D-12) that mimicked the D-8 substrate except that the downstream guanine in D-8 was replaced with an adenine. The affinity of mismatched dCTP (1228 ± 266 μM) to the D-12 and Dpo4 complex indeed drops to the level of other mismatched basepairs (Table 2.2). This result not only provided additional evidence for the “dNTP”-stabilized misalignment mechanism shown in Scheme 2.3, but also showed this type of dCTP incorporation dominated direct misincorporations (C:T and C:C). Once dCTP was incorporated, the next correct nucleotide incorporated into D-7 or D-8 could be dGTP leading to -1 frameshift, or dCTP, resulting in substitution after realignment of the primer terminus (7). These pathways are currently being studied by our group to determine which is more favored. Interestingly, a forward mutation assay has found that Dpo4 frequently generates deletions and substitutions of even non-iterated nucleotides, especially cytosines flanked by a 5’-template guanine (7). These results are consistent with our pre-steady state kinetic results. The “dNTP”-stabilized misalignment mechanism is also consistent with the ternary structure of Dpo4, DNA, and a mismatched ddGTP, showing Dpo4 simultaneously accommodates two template bases at its active site and the incoming ddGTP skips the first available template base and base pairs with the template base cytosine (17). In addition, when \textit{E. coli} DNA polymerase IV catalyzes single nucleotide incorporation into a template 5’-CAp sequence where the next available template base is 2-aminopurine (Ap), the incorporation of dGTP increases 2-aminopurine fluorescence intensity due to G:C base pairing which forces 2-aminopurine to become extrahelical or “looped-out” while the incorporation of dTTP quenches 2-aminopurine fluorescence signal due to the canonical T:Ap basepairing (14). The results from these 2-aminopurine
fluorescence intensity change experiments support the unusual basepairing in the “dNTP”-stabilized misalignment mechanism and support our observations for Dpo4.

Intriguingly, the rates of incorrect incorporation ($k_p$) of dCTP into D-7 and D-8 (Table 2.2) were not affected by the “dNTP”-stabilized misalignment. These kinetic results are consistent with the mismatched ternary structure of Dpo4 (17). In this ternary structure, the base at the 3’ end of the primer is tilted up, while the incoming guanine moiety is tilted down, thereby orienting the primer 3’-hydroxyl group and the $\alpha$-phosphate of the incoming nucleotide less than 4 Å apart. Additionally, this crystal structure reveals the sugar-phosphate bond of the incoming ddGTP is mobile. The mobility observed for the incoming nucleotide coupled with the observed unconventional basepairing arrangement may justify the slow incorporation rate observed for dCTP incorporation via the “dNTP”-stabilized misalignment mechanism.

Additionally, the “dNTP”-stabilized misalignment mechanism appears to be sequence-dependant. The binding affinity of mismatched dCTP to a template adenine flanked by a 5’-template guanine (D-1) was not increased. We speculate that the reason we did not observe an increase in affinity in this specific case was due to the difficulty of the bulky base adenine, rather than a less bulky pyrimidine base as in D-7 and D-8, in looping out at the enzyme active site, therefore allowing the incoming dCTP to base pair with the downstream template guanine. Interestingly, 50% of the one or two-base deletion errors made by hPolk (9) and *E. coli* DNA polymerase IV (15) are deletions of a template pyrimidine (Py) in a 5’-GPy sequence. Dbh, a homolog of Dpo4, has also shown a remarkable preference to skip a pyrimidine rather than a purine, when the skipped nucleotide is positioned immediately 3’ to a template G residue, generating a single-base deletion (16). Therefore, the “dNTP”-stabilized misalignment mechanism could apply to
all Y-family polymerases when they replicate through 5’-GPy “hotspots”. It will be interesting to see if these Y-family polymerases also use the “dNTP”-stabilized misalignment mechanism to incorporate dGTP into the 5’-CPy sequences. 2-Aminopurine fluorescence intensity studies with *E. coli* polymerase IV (14) and the mismatched Dpo4 ternary crystal structure (17) have shown this is possible. We are currently studying the incorporation of dGTP into 5’-CPy containing sequences.

The high frameshift rates at 5’-GPy, and possibly other sequence contexts, could be detrimental to genome stability especially for higher organisms, although, alternatively, substitutions resulting from realignment of the DNA substrate after incorporation via the “dNTP”-stabilized misalignment mechanism, could arguably promote genetic diversity. Two factors could decrease such high frameshift frequency *in vivo*. The first, previously mentioned, is the realignment of the primer 3’-terminus after the nucleotide incorporation following the “dNTP”-stabilized misalignment mechanism. This will eliminate frameshift mutations and increase substitution rates. The second factor is the inaccessibility of these frameshift hotspots to Y-family polymerases due to their low processivity coupled with the tight association of DNA replicative polymerases with DNA.
2.5 Figures, Tables, Schemes

Figure 2.1. Purification of *S. solfataricus Dpo4*. SDS-PAGE analysis and subsequent Coomassie Blue staining of proteins at each step of the purification was shown. Lane 1, protein marker; lane 2, crude extracts of non-induced cells; lane 3, crude extracts of IPTG-induced cells; lane 4, soluble fraction after cleared lysate was incubated at 78 °C for 12 minutes; lane 5, eluate from Ni-affinity column; lane 6, eluate from MonoS cation-exchange column.
Figure 2.2. Effects of Mg\(^{2+}\) concentration on the enzymatic activity of Dpo4. A preincubated solution of 5'-\(^{32}\)P-labeled-D-1 (30 nM) and 4-fold Dpo4 (120 nM) was rapidly mixed with correct nucleotide (100 \(\mu\)M dTTP) for various time intervals under single turnover conditions. Concentrations of all components were held constant while Mg\(^{2+}\) concentration was varied.
Figure 2.3. Effects of NaCl concentration on the enzymatic activity of Dpo4. A
preincubated solution of 5'-32P-labeled-D-1 (30 nM) and 4-fold Dpo4 (120 nM) was
rapidly mixed with correct nucleotide (100 μM dTTP) for various time intervals under
single turnover conditions. Concentrations of all components were held constant while
NaCl concentration was varied.
Figure 2.4. Effects of pH on the enzymatic activity of Dpo4. A preincubated solution of $5'$-$^{32}$P-labeled-D-1 (30 nM) and 4-fold Dpo4 (120 nM) was rapidly mixed with correct nucleotide (100 μM dTTP) for various time intervals under single turnover conditions. Concentrations of all components were held constant while the reaction pH was varied. Activity was assayed in 25 mM MES-NaOH buffer between pH 6.0 and 7.0, 25 mM Tris-Cl buffer for pH 8.0 and 8.5, and 25 mM glycine-NaOH buffer for pH 9.0 and 10.0.
Figure 2.5. Concentration dependence on the pre-steady state rate of correct nucleotide incorporation. A preincubated solution of Dpo4 (120 nM) and 5'-32P-labeled-D-1 (30 nM) was mixed with increasing concentrations of Mg$^{2+}$-dTTP (17 μM, ●; 50 μM, ○; 75 μM, ■; 100 μM, □; 150 μM, △; 200 μM, ▲; 250 μM, ▼; 400 μM, ▽; 900 μM, ◆; 1500 μM, ◇) for various time intervals. The solid lines are the best fits to the single exponential equation (equation 1).
Figure 2.6. Concentration dependence on the pre-steady state rate of correct nucleotide incorporation. The single exponential rates obtained from Figure 2.5 were plotted as a function of dTTP concentration. The rate data were then fit to the hyperbolic equation (equation 2) yielding a $k_p$ of $9.4 \pm 0.3 \text{ s}^{-1}$ and a $K_d$ of $230 \pm 17 \mu\text{M}$. 
Figure 2.7. Concentration dependence on the pre-steady state rate of incorrect nucleotide incorporation. A preincubated solution of Dpo4 (120 nM) and 5'-32P-labeled-D-1 (30 nM) was mixed with increasing concentrations of Mg$^{2+}$-dCTP (100μM, ●; 300 μM, ○; 600 μM, ■; 1000 μM, □; 1400 μM, ▲; 1800 μM, △) for various time intervals. The solid lines are the best fits to the single exponential equation.
Figure 2.8. Concentration dependence on the pre-steady state rate of incorrect nucleotide incorporation. The single exponential rates obtained from the above data fitting were plotted as a function of dCTP concentration. The rate data were then fit to the hyperbolic equation yielding a $k_p$ of 0.013 ± 0.001 s$^{-1}$ and a $K_d$ of 1036 ± 95 μM.
<table>
<thead>
<tr>
<th>DNA Substrates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D-1</strong></td>
</tr>
<tr>
<td>5’-CGCAGCCGTCACAACCAAACTCA-3’</td>
</tr>
<tr>
<td>3’-GCGTCGGCAGTTGGTGTGAGTACAGCTAGGCAGCGCAGG-5’</td>
</tr>
<tr>
<td><strong>D-6</strong></td>
</tr>
<tr>
<td>5’-CGCAGCCGTCACAACCAAACTCA-3’</td>
</tr>
<tr>
<td>3’-GCGTCGGCAGTTGGTGTGAGTACAGCTAGGCAGCGCAGG-5’</td>
</tr>
<tr>
<td><strong>D-7</strong></td>
</tr>
<tr>
<td>5’-CGCAGCCGTCACAACCAAACTCA-3’</td>
</tr>
<tr>
<td>3’-GCGTCGGCAGTTGGTGTGAGTACAGCTAGGCAGCGCAGG-5’</td>
</tr>
<tr>
<td><strong>D-8</strong></td>
</tr>
<tr>
<td>5’-CGCAGCCGTCACAACCAAACTCA-3’</td>
</tr>
<tr>
<td>3’-GCGTCGGCAGTTGGTGTGAGTACAGCTAGGCAGCGCAGG-5’</td>
</tr>
<tr>
<td><strong>D-12</strong></td>
</tr>
<tr>
<td>5’-CGCAGCCGTCACAACCAAACTCA-3’</td>
</tr>
<tr>
<td>3’-GCGTCGGCAGTTGGTGTGAGTACAGCTAGGCAGCGCAGG-5’</td>
</tr>
</tbody>
</table>

**Table 2.1. DNA Substrates.**
Table 2.2. Pre-Steady State Kinetic Parameters of Dpo4.
<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (μM$^{-1}$s$^{-1}$)</th>
<th>Fidelity$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dTTP</td>
<td>230 ± 17</td>
<td>9.4 ± 0.3</td>
<td>4.09 x 10$^{-2}$</td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>578 ± 188</td>
<td>0.006 ± 0.001</td>
<td>9.86 x 10$^{-6}$</td>
<td>2.4 x 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>1036 ± 95</td>
<td>0.013 ± 0.001</td>
<td>1.25 x 10$^{-5}$</td>
<td>3.1 x 10$^{-4}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>1150 ± 312</td>
<td>0.007 ± 0.001</td>
<td>6.00 x 10$^{-6}$</td>
<td>1.5 x 10$^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>70 ± 8</td>
<td>7.6 ± 0.2</td>
<td>1.09 x 10$^{-1}$</td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>334 ± 128</td>
<td>0.009 ± 0.001</td>
<td>2.69 x 10$^{-5}$</td>
<td>2.5 x 10$^{-4}$</td>
</tr>
<tr>
<td>dTTP</td>
<td>1283 ± 150</td>
<td>0.077 ± 0.004</td>
<td>6.00 x 10$^{-5}$</td>
<td>5.5 x 10$^{-4}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>131 ± 42</td>
<td>0.008 ± 0.001</td>
<td>6.11 x 10$^{-5}$</td>
<td>5.6 x 10$^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>206 ± 46</td>
<td>16.1 ± 0.9</td>
<td>7.82 x 10$^{-2}$</td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>1941 ± 589</td>
<td>0.034 ± 0.006</td>
<td>1.75 x 10$^{-5}$</td>
<td>2.2 x 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>309 ± 84</td>
<td>0.026 ± 0.002</td>
<td>8.41 x 10$^{-5}$</td>
<td>1.1 x 10$^{-3}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>935 ± 129</td>
<td>0.066 ± 0.004</td>
<td>7.06 x 10$^{-5}$</td>
<td>9.0 x 10$^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>171 ± 15</td>
<td>9.4 ± 0.2</td>
<td>5.50 x 10$^{-2}$</td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>617 ± 113</td>
<td>0.016 ± 0.001</td>
<td>2.59 x 10$^{-5}$</td>
<td>4.7 x 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>192 ± 48</td>
<td>0.034 ± 0.002</td>
<td>1.77 x 10$^{-4}$</td>
<td>3.2 x 10$^{-3}$</td>
</tr>
<tr>
<td>dTTP</td>
<td>913 ± 181</td>
<td>0.011 ± 0.001</td>
<td>1.20 x 10$^{-5}$</td>
<td>2.2 x 10$^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>1228 ± 266</td>
<td>0.005 ± 0.001</td>
<td>4.23 x 10$^{-6}$</td>
<td></td>
</tr>
<tr>
<td>Polymerase</td>
<td>Fidelity</td>
<td>$K_d$ difference</td>
<td>$k_p$ difference</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>-----------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>Dpo4$^a$</td>
<td>$1.5 \times 10^{-4}$ to $3.2 \times 10^{-3}$</td>
<td>1.1 to 18.1</td>
<td>$2.4 \times 10^{2}$ to $1.7 \times 10^{3}$</td>
<td></td>
</tr>
<tr>
<td>yPol$^b$</td>
<td>$1.24 \times 10^{-3}$</td>
<td>5.4</td>
<td>$1.5 \times 10^{2}$</td>
<td></td>
</tr>
<tr>
<td>rPol$^c$</td>
<td>$6.3 \times 10^{-5}$ to $2.0 \times 10^{-4}$</td>
<td>4.8 to 43.4</td>
<td>$83.3$ to $2.9 \times 10^{3}$</td>
<td></td>
</tr>
<tr>
<td>E. coli Pol I$^d$</td>
<td>$1.7 \times 10^{-4}$ to $5.3 \times 10^{-4}$</td>
<td>1.7 to 4.2</td>
<td>$5.0 \times 10^{3}$ to $2.4 \times 10^{4}$</td>
<td></td>
</tr>
<tr>
<td>hPoly$^e$</td>
<td>$1.8 \times 10^{-6}$ to $2.9 \times 10^{-4}$</td>
<td>72 to 454</td>
<td>$38.5$ to $3.4 \times 10^{3}$</td>
<td></td>
</tr>
<tr>
<td>T7 DNA Pol$^f$</td>
<td>$2.6 \times 10^{-7}$ to $6.7 \times 10^{-6}$</td>
<td>200 to 400</td>
<td>$2.0 \times 10^{3}$ to $4.0 \times 10^{3}$</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.3. Comparison of Fidelity of Dpo4 with the Replicative, Repair, and Lesion-Bypass DNA Polymerases.**

$^a$at 37 °C from this chapter. $^b$at 37 °C and Ref. (25), the numbers are derived from the incorporations of a correct and an incorrect nucleotides, not from the 16 possible incorporations. $^c$at 37 °C and Ref. (33). $^d$at 22 °C, not including the fidelity contribution from the 3’-5’ and 5’-3’ exonucleases and Ref. (35). $^e$at 37 °C, not including the fidelity contribution from the 3’-5’ exonuclease activity, and Ref. (34). $^f$at 20 °C, not including the fidelity contribution from the 3’-5’ exonuclease activity, and Ref. (28).
Scheme 2.1. Minimal kinetic mechanism for a DNA polymerase.

\[ E\bullet DNA_n + dNTP \overset{K_d}{\underset{k_p}{\leftrightarrow}} E\bullet DNA_n\bullet dNTP \overset{k_p}{\rightarrow} E'\bullet DNA_{n+1} \bullet PPi \]
Scheme 2.2. A detailed minimal kinetic mechanism for incorporation of a dNTP.
Scheme 2.3. The “dNTP”-stabilized misalignment model.
2.6 References


CHAPTER 3

MECHANISM OF DNA POLYMERIZATION CATALYZED BY *SULFOLOBUS SOLFATARICUS* DNA POLYMERASE IV AT 37 °C

3.1 Introduction

The recently discovered Y-family polymerases are found to bypass a variety of DNA lesions which block DNA synthesis catalyzed by replicative polymerases. All the Y-family polymerases that have thus far been biochemically characterized are devoid of intrinsic proof-reading exonuclease activity and are distributive (1-7). Previous steady-state kinetic studies reveal the Y-family polymerases, when compared to similar studies of replicative polymerases, have two to three orders of magnitude lower fidelity. This low fidelity is due to the loose polymerase active site and lack of fidelity checking mechanisms as revealed by the high resolution crystal structure of *Sulfolobus solfataricus* DNA polymerase IV (Dpo4) in a ternary complex with DNA and an incoming nucleotide (8). In addition to the signature right hand shape with palm, finger, and thumb domains found in all structure-known DNA polymerases (9-14), Dpo4 has an additional fourth domain, designated “little finger,” which is absent in both replicative and repair DNA polymerases. The unique biochemical and structural properties of Y-family polymerases suggest they may utilize different kinetic mechanisms to catalyze DNA synthesis in comparison to replicative and repair polymerases. As established by pre-steady state kinetic analysis, replicative polymerases like T7 DNA polymerase,
utilize an induced-fit mechanism in which polymerase fidelity is a result of conformational coupling, where the energy of nucleotide binding is used to drive a rate-limiting protein conformational change preceding a fast chemistry step (15). In contrast, DNA repair polymerases like rabbit DNA polymerase β (rPolβ) incorporate nucleotides by employing what is known as the rate-limiting transition state mechanism, in which the free energy difference in chemical transition states between correct and incorrect basepairs dictates polymerase fidelity through a rate-limiting chemistry step (16).

Recently, a pre-steady state kinetic study on yeast polymerase η (yPolη), another Y-family member, with an undamaged DNA substrate was reported (17). yPolη, like other replicative polymerases, is shown to utilize the induced-fit mechanism to select and incorporate correct nucleotides during polymerization. However, unlike other DNA polymerases, the protein conformational change, as opposed to the chemistry step, in yPolη is suggested to be the rate-limiting step for the incorporation of mismatched nucleotides.

We have implemented pre-steady state kinetic methods to determine the fidelity of DNA polymerization catalyzed by Dpo4 (18). Here, we employ pre-steady state kinetic techniques to investigate the kinetic mechanism of DNA polymerization by Dpo4 with an undamaged template. Dpo4, a thermostable archaeal polymerase (19), was selected to serve as a model Y-family polymerase due to the ability to obtain milligrams of active enzyme required for pre-steady state kinetic studies, which can be overexpressed in *Escherichia coli* and purified using common chromatographic methods (20). Our kinetic results suggest Dpo4 like yPolη follows the induced-fit mechanism for nucleotide discrimination. However, the rate-limiting step for the incorporation of an incorrect nucleotide by Dpo4, unlike yPolη, was limited by the chemistry step.
3.2 Materials and methods

Materials. These chemicals were purchased from the following companies: [$\alpha$-32P]dTTP and [$\gamma$-32P]ATP, Perkin Elmer Life Sciences (Boston, MA); dNTPs, Gibco-BRL (Rockville, MD); $S_\gamma$-dTTP$\alpha$S and $S_\gamma$-dGTP$\alpha$S, Biolog, Life Science Institute (Bremen, Germany); ddTTP, Trilink Biotechnologies (San Diego, CA). Full-length Dpo4 fused to a C-terminal His$_6$ tag was overexpressed in *E. coli* as described in (18). The protein was stored in 20 µl aliquots and stored at -80 °C (18). The DNA substrate (D-1) listed in Figure 3.1 was prepared as described previously (18).

Pre-steady state kinetic assays. All experiments using Dpo4, if not specified, were performed in an optimized reaction buffer R containing 50 mM HEPES (pH 7.5 at 37 °C), 5 mM MgCl$_2$, 50 mM NaCl, and 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA. All reactions were carried out at 37 °C using a rapid chemical quench flow apparatus (KinTek, PA) as described previously (18).

Pulse-chase and pulse-quench experiments. Pulse-chase and pulse-quench experiments were performed in buffer R using a rapid chemical quench flow instrument. A preincubated solution containing Dpo4 (30 nM) and unlabeled D-1 substrate (30 nM) loaded into one sample loop was rapidly mixed with buffer containing [$\alpha$-32P]dTTP (50 µM) from a second sample loop for reaction times ranging from 50 ms to 4.5 s. In the pulse-quench experiment, reactions were immediately quenched with 1 M HCl. In the pulse-chase experiment, reactions were immediately chased by 2.5 mM unlabeled dTTP for 30 s, followed by quenching with 1 M HCl. In both cases, quenched reactions were
treated with chloroform and neutralized with a solution of NaOH/TRIS. Following neutralization, each individual sample was quantitated via sequencing gel analysis.

*Product analysis.* Reaction products were analyzed by sequencing gel electrophoresis (17% acrylamide, 8 M urea, 1x TBE running buffer) and quantitated with a Phosphorimager 445 SI (Molecular Dynamics).

*Data analysis.* Data were fit by nonlinear regression using the program KaleidaGraph (Synergy Software). Data from burst experiments were fit to equation 1 where $A$ is the fraction of active enzyme, $E_0$ the enzyme concentration

$$[\text{Product}] = AE_0[1 – \exp(-k_1t) + k_2t] \quad (1)$$

measured spectrophotometrically, $k_1$ the observed burst rate, and $k_2$ the observed steady-state rate.

Data from the steady-state kinetic experiments were fit to equation 2 where

$$[\text{Product}] = k_{ss}E_0t + E_0 \quad (2)$$

$k_{ss}$ is the steady-state rate constant of dNTP incorporation at initial active enzyme concentration of $E_0$.

Data from the active site titration were fit to equation 3 where $K_d$ represents

$$[\text{E·DNA}] = 0.5(K_d + E_0 + D_o) - 0.5[(K_d + E_0 + D_o)^2- 4E_0D_o]^{1/2} \quad (3)$$

the equilibrium dissociation constant for DNA substrates, $E_0$ the active enzyme concentration, and $D_o$ the DNA concentration.

Data from the single turnover experiments were fit to equation 4 (single exponential), where $E_0$ represents the enzyme concentration measured spectrophotometrically, $A$ the fraction of active enzyme, and $k_{obs}$ the observed single turnover rate.

$$[\text{Product}] = E_0A[1 – \exp(-k_{obs}t)] \quad (4)$$
The data from the processive elongation of 21/41-mer to 27/41-mer were modeled using an improved pc version of KINSIM provided by Carl Frieden, Washington University (St. Louis, MO) (21). Final fitting of the data was accomplished by nonlinear regression based on kinetic simulation using an improved pc version of FITSIM (22).

3.3 Results

Burst kinetics. To determine whether a single nucleotide incorporation catalyzed by Dpo4 follows the biphasic kinetics as observed in other polymerases (23-28), a pre-steady state kinetic analysis for the correct incorporation of dTTP into a DNA substrate D-1 (Figure 3.1) was performed under conditions where the DNA substrate was in slight molar excess over the enzyme in order to observe the kinetics of the first and subsequent turnovers of the enzyme [23]. A preincubated solution of Dpo4 (30 nM) and 5’-[32P]-D-1 (120 nM) in one syringe was mixed with dTTP from a second syringe at 37 °C in the optimized reaction buffer R (18). The reactions were quenched at time intervals ranging from 5 ms to 15 s. The reaction products were analyzed by sequencing gel electrophoresis and quantitated using a Phosphorimager. The resulting time course of formation of labeled 22/41-mer (Figure 3.2) shows the biphasic kinetics: a rapid burst of dTTP incorporation followed by a slower linear phase. The first turnover occurred at a rate of 3.8 ± 0.2 s⁻¹ (burst phase), while subsequent turnovers (linear phase) occurred at a much slower rate of 0.020 ± 0.002 s⁻¹ (Figure 3.2). To confirm the observed linear phase was the bona fide steady-state phase, we conducted an independent steady-state kinetic experiment under conditions where the DNA substrate concentration was approximately 200-fold greater than the enzyme concentration. The observed steady-state rate was 0.07
s$^{-1}$ (Figure 3.3) which was similar to the rate observed in the linear phase of the burst experiment (Figure 3.2). Thus, the linear phase was indeed the steady-state phase. Like other polymerases (23-28), the pre-steady state burst observed in Figure 3.2 suggested that the linear phase was limited by the slow DNA dissociation from the binary complex of enzyme and DNA binary complex (E•DNA). The former could be specifically limited by the protein conformational change step, the chemistry step, or both. The latter was confirmed by the assessment of DNA dissociation from the E•DNA binary complex. A preincubated solution of Dpo4 (50 nM) and 5'-radiolabeled D-1 (100 nM) was mixed with a large excess of activated calf thymus DNA (1mg/mL) for time intervals varying from 10 s to 20 min. The large excess of calf thymus DNA trapped Dpo4 as it dissociated from the labeled D-1 substrate. The subsequent reaction with dTTP was held constant for 15 s to afford ample time for the extension of the available labeled substrate. Reactions were quenched upon addition of EDTA and mixtures were analyzed by gel analysis. The resulting data were fit to a single-exponential equation to yield a rate constant of 0.022 s$^{-1}$ ($k^{-1}$) which is close to the value of 0.020 s$^{-1}$ observed in the steady-state phase in Figure 3.2.

Active site titration. Since the catalysis in the first turnover was much faster than the equilibration of the enzyme and DNA ($E + DNA = E•DNA$), a titration of the enzyme active site with DNA can be used to measure the equilibrium dissociation constant ($K_d$) of the E•DNA binary complex by examining the DNA concentration dependence of the burst amplitude (23). A preincubated solution of Dpo4 (22 nM, as determined by UV absorbance measurements), Mg$^{2+}$, and increasing concentration of 5’-$^{32}$P-labeled D-1 was rapidly mixed with dTTP and Mg$^{2+}$. Reactions were quenched with EDTA after 1.2 sec which allowed adequate time to reach the maximum burst amplitude with negligible contribution of multiple turnovers. The products were analyzed by sequencing gel
electrophoresis. The experiments were performed in triplicate and the mean product concentration was plotted against the DNA concentration. The solid line (Figure 3.4) was a fit of the data to a quadratic equation (Materials and Methods) which gave a $K_d$ value for the active complex between Dpo4 and DNA 21/41-mer of $10.6 \pm 1.3$ nM, and an enzyme amplitude of $19.4 \pm 0.5$ nM or 88% of the protein concentration measured spectrophotometrically at 280 nm (18). Thus, the enzyme is not 100% active and the enzyme concentrations for the following experiments were corrected for the amount of active enzyme concentrations. On the basis of the DNA dissociation rate of 0.02 s$^{-1}$ (see above), the apparent second-order binding rate constant of the Dpo4•21/41-mer complex was thus calculated ($k_{on} = k_{off}/K_d = 1.9 \times 10^6$ M$^{-1}$s$^{-1}$). This indicated the binding rate of DNA to Dpo4 was below diffusion limit. A similar phenomenon has been obtained previously for the binding of DNA to T7 DNA polymerase ($1.12 \times 10^7$ M$^{-1}$s$^{-1}$) (23).

**Association and dissociation rates of incoming nucleotide.** The ground-state binding affinity of dTTP ($K_d$) was measured through the dTTP concentration dependence of the single turnover rate, yielding the maximum dTTP incorporation rate ($k_p$) of $9.4 \pm 0.3$ s$^{-1}$, and a $K_d$ of $230 \pm 17$ μM for the binding of dTTP in Chapter 2 (18). Since the active site of Dpo4 is loose and relatively accessible to solvent as revealed by the crystal structure of Dpo4 (8), we assume the upper limit of the association rate constant of dTTP ($k_{on}$) is close to the diffusion limit of $1.0 \times 10^8$ M$^{-1}$s$^{-1}$. The upper limit of the dissociation rate constant of dTTP from the E•DNA•dNTP complex $k_{off} = k_{on}K_d = 23,000$ s$^{-1}$ was thereby estimated. This dNTP dissociation rate is too fast to be measured by current rapid chemical quench and stopped-flow techniques.

**Elemental effect of the incorporation of a matched incoming nucleotide.** The maximum incorporation rate, $k_p$, can be a direct measure of the chemistry step of phosphodiester
bond formation, or the protein conformational change step preceding the chemistry step, or both. To distinguish between these three possibilities, the effect of an incoming nucleotide analog, characterized by a substitution of the \( \alpha \)-phosphate with a phosphothioate group (Scheme 3.1), on \( k_{obs} \), the observed incorporation rate, was assayed. This approach was based on the observations that a rate-limiting chemical step involving the making or breaking of a phosphate bond shows a phosphothioate elemental effect (29,30). A preincubated solution of Dpo4 (120 nM) and 5'-\( ^{32} \)P-labeled D-1 (30 nM) was reacted with either dTTP (100 \( \mu \)M) or \( S_p \)-dTTP\( \alpha \)S (100 \( \mu \)M, >95% purity) in buffer R. The \( S_p \) isomer (Scheme 3.1), rather than the \( R_p \) isomer of dTTP\( \alpha \)S, was used due to the stereoselectivity of polymerases in the presence of Mg\(^{2+} \) as observed with rPol\( \beta \) (31). The reactions were quenched at various times. The data were fit into a single exponential equation (equation 4) with \( k_{obs} \) values of \( 6.27 \pm 0.38 \) s\(^{-1} \) and \( 4.49 \pm 0.30 \) s\(^{-1} \) for dTTP and \( S_p \)-dTTP\( \alpha \)S, respectively (Figure 3.5). Therefore, the elemental effect of the matched dTTP incorporation was 1.4. On the basis of modeling studies of the hydrolysis of phosphate diesters (29), an elemental effect of 4 to 11 has generally been taken as evidence for a rate-limiting chemistry step during DNA synthesis catalyzed by a polymerase (17). This suggested the chemistry step was not rate-limiting for the incorporation of correct nucleotides by Dpo4 in the burst phase. This conclusion was consistent with the experimental results shown below.

**Elemental effect of the incorporation of a mismatched incoming nucleotide.** To probe whether the chemistry step, the protein conformational change step, or both were rate-limiting during incorrect nucleotide incorporation, an analogous elemental effect study was performed with the incorporation of mismatched incoming dGTP into D-1. A preincubated solution of Dpo4 (120 nM) and 5'-\( ^{32} \)P-labeled D-1 (30 nM) was reacted with either dGTP (100 \( \mu \)M) or \( S_p \)-dGTP\( \alpha \)S (100 \( \mu \)M, >95% purity) in buffer R and the
products were analyzed and quantitated as above. The observed single turnover rates ($k_{obs}$) were $0.0021 \pm 0.0001 \text{s}^{-1}$ and $0.00037 \pm 0.00002 \text{s}^{-1}$ for dGTP and Sp-dGTP$\alpha$S, respectively (Figure 3.6). The elemental effect of dGTP incorporation into D-1 was calculated to be 5.8, which suggested the chemistry step limits the rate of incorrect nucleotide incorporation (29). However, additional evidence is required to confirm this observation.

Notably, we performed experiments assessing a possible elemental effect for both correct and incorrect nucleotide incorporations at subsaturating nucleotide concentrations based on our results from nucleotide binding (18). However, since the difference between the maximum incorporation rate $k_p$ and the observed $k_{obs}$ at a subsaturating dNTP concentration equals $[\text{dNTP}]/(K_d + [\text{dNTP}])$, the value of the elemental effect derived from the $k_p$ ratio should be identical or similar to the elemental effect derived from the $k_{obs}$ ratio if the $K_d$ values of dNTP and Sp-dNTP$\alpha$S are equal or close. Due to the lack of interactions between Dpo4 and the nucleotide at the polymerase active site revealed by ternary crystal structures of Dpo4 (8), there will be essentially no difference in dNTP and its $\alpha$-thio analog and, as such, the $K_d$ difference will be minimal at best. Structural analyses of other polymerase active sites, like DNA polymerase $\beta$, also suggest a minimal difference between the ground-state binding affinity of dNTP and its analogue dNTP$\alpha$S (31). Therefore, the subsaturating concentrations of nucleotides dNTP and dNTP$\alpha$S used in our studies should not affect the value of the elemental effect and the overall conclusion.

**Processive polymerization.** The processivity of Dpo4 was investigated by including three dNTPs in the reaction mixture, thus permitting the elongation of D-1, a 21/41-mer DNA substrate, to a 27/41-mer. The experiment was performed under conditions where D-1
was in molar excess over Dpo4 to ensure the measured kinetics of sequential elongation steps was a function of single enzyme binding events. A preincubated solution of Dpo4 (35 nM) and 5'-32P radiolabeled D-1 (100 nM) was mixed with dTTP, dCTP, and dGTP (1.2 mM each) in the reaction buffer R with Mg2+ compensation (see above) for reaction times ranging from several milliseconds to 2.5 seconds prior to being quenched with EDTA and subsequently analyzed by gel assay. The time courses of product formation and substrate disappearance were fit by nonlinear regression using computer simulation programs of Kinsim (21) and Fitsim (22) and a mechanism consisting of a series of six single nucleotide incorporations and DNA dissociations (Figures 3.7 and 3.8). The rate constants of formation of intermediates were as follows: 22-mer at 2.3 ± 0.1 s⁻¹, 23-mer at 2.9 ± 0.2 s⁻¹, 24-mer at 3.5 ± 0.4 s⁻¹, 25-mer at 4.3 ± 0.6 s⁻¹, 26-mer at 3.9 ± 0.6 s⁻¹, and 27-mer at 3.4 ± 0.4 s⁻¹. This simulation also provided DNA dissociation rates: 21/41-mer at 0.26 ± 0.05 s⁻¹; 22/41-mer at 0.18 ± 0.11 s⁻¹; 23/41-mer at 0.19 ± 0.12 s⁻¹; 24/41-mer at 0.18 ± 0.16 s⁻¹; 25/41-mer at 0.18 ± 0.18 s⁻¹; 26/41-mer at 0.15 ± 0.11 s⁻¹; 27/41-mer at 0.20 ± 0.09 s⁻¹. These dissociation rates in the range of 0.15 to 0.26 s⁻¹ were much higher than the E•DNA dissociation rate of 0.02 s⁻¹ (Figure 3.2). If DNA dissociated at rate of 0.02 s⁻¹, each of the intermediates shown in Figure 3.7 should have returned to the baseline as in the case of T7 DNA polymerase (23). However, the intermediates accumulated in the range of 5 to 10 nM which suggested that DNA dissociated at a faster rate from a yet unidentified intermediate state formed after the binding of dNTP to the E•DNA complex. This intermediate could be the initial binding complex E•DNA•dNTP, the complex E’•DNA•dNTP after a putative protein conformational change, or the complex E’•DNA•PPi after the chemistry step.

Measurement of the DNA dissociation rate of the E•DNA•dNTP complex. DNA can either dissociate from the ternary complex, E•DNA•dNTP, or be converted to product

67
after the putative protein conformational change and chemistry steps. To measure the DNA dissociation rate \((k_{\text{off}})\) of the E\(\bullet\)DNA\(\bullet\)dNTP complex, the nucleotide analog \(S_p\)-dTTP\(\alpha\)S, rather than dTTP, was used in this assay to procure a slightly larger kinetic partitioning due to the slower rate of incorporation of \(S_p\)-dTTP\(\alpha\)S relative to dTTP (Figure 3.5). A large molar excess of unlabeled D-1 was used to trap any dissociated enzyme molecules which were bound by radiolabeled DNA substrate molecules. The experiment was conducted by reacting a preincubated Dpo4 (55 nM) and 5’-radiolabeled D-1 (60 nM) with \(S_p\)-dTTP\(\alpha\)S (1.2 mM) alone or with \(S_p\)-dTTP\(\alpha\)S (1.2 mM) and unlabeled D-1 (2.5 \(\mu\)M). Incorporation of \(S_p\)-dTTP\(\alpha\)S occurred at a rate of 11.0 \(s^{-1}\) \((k_{\text{obs}})\) in both time courses, but the amplitudes of the burst were 53.1 \(\pm\) 0.6 nM and 51.2 \(\pm\) 0.7 nM in the absence and presence of the DNA trap respectively (Figure 3.9). Since the relative amplitude (51.2/53.1) is equal to \(k_{\text{obs}}/(k_{\text{obs}} + k_{\text{off}})\), the dissociation rate \(k_{\text{off}}\) of the E\(\bullet\)DNA\(\bullet\)dNTP complex was estimated to be 0.41 \(s^{-1}\). This rate was similar to the rates observed in Figures 3.7 and 3.8, indicating the low processivity of Dpo4 was due to fast DNA dissociation from the E\(\bullet\)DNA\(\bullet\)dNTP ternary complex. Similar results have been observed with HIV-1 reverse transcriptase (26).

**Measurement of the dissociation rates of the E’\(\bullet\)DNA\(\bullet\)dNTP complex.** Steady-state kinetic experiments were used to measure the dissociation rate of DNA from a E’\(\bullet\)DNA\(\bullet\)dNTP complex that could not undergo the chemical step due to the presence of ddTTP that, once incorporated, does not have the required hydroxyl group at the 3’-end of the primer strand for a subsequent extension reaction. The experiments were conducted by reacting a preincubated solution of Dpo4 (1.2 nM) and 5’-radiolabeled D-1 (200 nM) with ddTTP (1.2 mM) alone or with ddTTP (1.2 mM) and the next correct nucleotide dCTP (1.2 mM). The steady-state rate in the presence of ddTTP alone (0.028 \(\pm\) 0.001 \(s^{-1}\)) (Figure 3.10) was a direct measure of the rate of DNA dissociation from the
E•DNA product complex. This rate was close to the steady-state rate measured in Figures 3.2 and 3.3. In the presence of ddTTP and dCTP, an extremely slow off rate of $0.0040 \pm 0.0003 \text{s}^{-1}$ was observed from the dissociation of the E’•DNA•dNTP complex (Figure 3.10). Similar results were observed previously in HIV-1 reverse transcriptase (26).

These results suggested the binding of the next correct nucleotide inhibits the dissociation of DNA from the E’•DNA•dNTP complex. More importantly, the slow dissociation of DNA from the E’•DNA•dNTP provided kinetic evidence for the existence of the protein conformational change prior to the chemistry step.

**Pulse-chase and pulse-quench experiments.** Pulse-chase and pulse-quench experiments were conducted to seek additional, but independent evidence for the existence of a slow conformational change step that would precede the chemistry step in the incorporation of a matched nucleotide by Dpo4. These experiments were carried out by comparing two separate time courses. In each case, a preincubated solution of Dpo4 (30 nM) and unlabeled D-1 (30 nM) was mixed with $[\alpha^{\text{32P}}]$dTTP (50 μM) for various time intervals in a rapid chemical quench-flow instrument. In the first experiment, reactions were quenched by the addition of 1 M HCl, while in the second experiment, each reaction was chased with unlabeled dTTP (2.5 mM) for an additional 30 seconds, followed by acid-quench (1 M HCl). The reaction mixtures were denatured by the addition of chloroform and neutralized with 1 M NaOH prior to analysis by sequencing gel electrophoresis. In the pulse-quench reactions, 1 M HCl quenched all the enzyme-bound species. In the chased reactions, the enzyme-bound complex was allowed to partition between both the reverse and forward directions, and a stable bound complex of Dpo4, DNA, and $[\alpha^{\text{32P}}]$dTTP, if any, would be chased by the large amount of cold dTTP in the forward direction to form an excess of radiolabeled product. Comparison of the kinetics for these experiments provided direct evidence for the existence of an enzyme-bound complex.
preceding the chemical step (23,32). The pulse-quench data were fit to a single exponential equation to yield an amplitude of $20.0 \pm 0.6$ nM and a rate constant of $0.62 \pm 0.17$ s$^{-1}$. The chased data were also fit to a single exponential equation with an amplitude of $25.5 \pm 1.0$ nM and a corresponding rate of $0.62 \pm 0.24$ s$^{-1}$ (Figure 3.11). The incorporation rates of [$\alpha$-$^{32}$P]-dTTP (50 μM) for the pulse-chase and pulse-quench experiments were lower than expected ($\sim 2$ s$^{-1}$) and it was due to experimental error. The obvious difference in amplitudes (5.5 nM) between the two sets of experiments was reproducible and demonstrated the formation of an intermediate complex prior to the chemical step (23,32). Theoretically, this intermediate complex could be either E•DNA, E•DNA•dNTP, or E’•DNA•dNTP. However, we know that this complex cannot be the E•DNA binary complex since such an intermediate would bind cold dTTP in the pulse-chase conditions and be removed from observation. This would result in identical curves in each of the two sets of experiments. Likewise, this intermediate species was not the E•DNA•dNTP ternary complex since the difference in amplitudes (20%) would require that the complex partitioned between formation of product (amplitude = 80%, 2 s$^{-1}$) and release of dTTP to form the E•DNA complex ($\sim 0.09$ s$^{-1}$) and form E, DNA, dTTP (0.41 s$^{-1}$) (32). The off rate of dTTP from the E•DNA•dNTP complex is much slower than the estimated dissociation rate of dTTP 23,000 s$^{-1}$ based on its ground-state binding affinity (see above). The high off rate of dNTP from the E•DNA•dNTP complex has also been estimated with T7 DNA polymerase ($\geq 1000$ s$^{-1}$) (23) and the Klenow fragment of E. coli DNA polymerase I ($\geq 1000$ s$^{-1}$) (33). Therefore, this intermediate species was most likely the E’•DNA•dNTP ternary complex.
3.4 Discussion

In this chapter we have established the elementary steps of nucleotide incorporation into undamaged DNA catalyzed by Dpo4. On the basis of our pre-steady state kinetic results, we propose a kinetic mechanism for DNA polymerization by Dpo4 shown in Scheme 3.2. In Scheme 3.2, E, DNA_n, dNTP, and PP_i represent Dpo4, D-1, dTTP, and pyrophosphate, respectively. The estimated kinetic constants are listed in Table 3.1. On the basis of the previous findings that free DNA polymerases, such as *E. coli* polymerase I (34) and T7 DNA polymerase (23), do not specifically bind a nucleotide prior to the binding of DNA, we propose Dpo4 first binds DNA followed by dNTP, to form the E•DNA•dNTP ternary complex. The pre-steady state time course of single dTTP incorporation revealed biphasic kinetics (Figure 3.2). The burst of product formation in the first turnover was limited by nucleotide incorporation, while the slow steady-state phase of the subsequent turnovers was limited by the dissociation of DNA product from the enzyme. The steady-state turnover rate was further measured and confirmed by two independent steady-state kinetic experiments with dTTP (Figure 3.3) and ddTTP (Figure 3.10), and a third experiment involving the direct measurement of the dissociation rate of the E•DNA binary complex (see above). The rate-limiting step in the first turnover was determined to be the first protein conformational change after the initial collision of dNTP with the E•DNA binary complex: E•DNA•dNTP ↔ E’•DNA•dNTP. We have observed three lines of evidence supporting such a protein conformational change: (i) the small elemental effect of 1.4 for the incorporations of dTTP versus S_p-dTTPαS suggested the chemistry step was not rate-limiting based on the previous studies of the elemental effect (4-11) observed in the hydrolysis of pyrophosphate (29). This approach has been used to establish the rate-limiting step in the nucleotide incorporation catalyzed by other
polymerases (17,23,35,36). Thus, the potential rate-limiting step must be a step preceding the chemistry step (Step 4 in Scheme 3.2). Since the initial binding of dNTP (Step 2) is a fast equilibrium step, the only candidate was the intermediate step between steps 2 and 4 (Scheme 3.2) which was therefore suggested to be the protein conformational change; (ii) in the presence of ddTTP and next correct nucleotide dCTP, we observed an extremely slow dissociation rate (0.004 s\(^{-1}\)) of DNA from a ternary complex which could not undergo the chemistry step for the incorporation of dCTP due to the lack of a 3'-hydroxyl group on the primer strand after the incorporation of ddTTP (Figure 3.10). Since the dissociation of the E•DNA (0.02-0.07 s\(^{-1}\)) and E•DNA•dNTP (0.41 s\(^{-1}\)) complexes were much faster, the tight binding ternary complex observed in the presence of ddTTP and dCTP was thus deduced to be E’•DNA•dNTP. After the initial binding of dNTP to form the loose E•DNA•dNTP complex, the enzyme undergoes an isomerization to tighten the binding of both DNA and dNTP at the active site. The fast dissociation of the E•DNA•dNTP complex was observed in the processive polymerization experiment shown in Figures 3.7 and 3.8 and the rate was measured directly by following kinetic partitioning in the incorporation of Sp-dTTP\(\alpha\)S in the presence of an unlabeled DNA trap (Figure 3.9); (iii) the reaction amplitude difference between the pulse-chase (25.5 nM) and pulse-quench (20.0 nM) clearly indicated the existence of the E’•DNA• [\(\alpha-^{32}\)P]-dTTp complex which accumulated but was chased forward by a large excess of cold dTTP (Figure 3.11). Such an amplitude difference has been observed in \(E.\ coli\) DNA polymerase I (32), T7 DNA polymerase (23), and yPol\(\eta\) (17). In addition, about 5.5 nM of the pool of Dpo4 was in the E’•DNA• [\(\alpha-^{32}\)P]-dTTP complex which did not yield products under pulse-quench conditions, but was chased to products by large excess of cold dTTP under the pulse-chase conditions. For the E’•DNA• [\(\alpha-^{32}\)P]-dTTP complex to accumulate, it is necessary to have a step after the chemistry step to serve as a kinetic “roadblock” (32). If PP\(_i\) dissociated rapidly after the chemistry step, the E’•DNA• [\(\alpha-\)
$^{32}$P]-dTTP complex would not accumulate. A slow conformational change
$(E'\cdot DNA_{n+1}\cdot PP_i \leftrightarrow E\cdot DNA_{n+1}\cdot PP_i)$ is thus included in Scheme 3.2. However, further
evidence is required to confirm the presence of this step. Additional experiments,
particularly those that investigate the reverse elementary steps, are needed to measure the
remaining microscopic rate constants in Scheme 3.2.

*Induced-fit mechanism of nucleotide incorporation.* Scheme 3.2 shows Dpo4, like yPolη
(17) and other polymerases (23-28), uses the induced-fit mechanism to select correct
incoming nucleotides. The fidelity of Dpo4 was contributed primarily from the protein
conformational change preceding the chemistry step (18). Presently, the contributions of
individual amino acid residues and domains of Dpo4 to the protein conformational
change have not been explicitly defined due to the lack of the structural information
regarding the binary complex of a Y-family polymerase and DNA. However, a
superimposition of the conserved palm domains of Dpo4 in its ternary complex (8) and
yPolη alone (37) reveals a protein conformational change: the inward rotation of the
finger and little finger domains of Dpo4 towards DNA by $\sim 48^\circ$ (38). The superimposition
of the structure of yPolη alone (37) with the recently solved structure of Dpo4 in the
presence of DNA (containing a cis-syn T-T dimer) and a correct incoming nucleotide also
suggests a large conformational change for Dpo4: $16^\circ$, $16^\circ$, and $45^\circ$ for the finger, thumb,
and little finger domains, respectively (39). The large swing of the finger domain
between the “open” conformation of $E\cdot DNA$ and the “closed” conformation of
$E'\cdot DNA\cdot dNTP$ has been observed in other polymerases such as T7 DNA polymerase
(13), HIV-1 reverse transcriptase (12), *Thermus aquaticus* DNA polymerase I (9), and
DNA polymerase $\beta$ (10). In contrast, the structure of Dbh, a homolog of Dpo4 from
*Sulfolobus acidocaldarius*, in the absence of both DNA and an incoming dNTP (40,41),
shows this Y-family polymerase appears to be in the “closed” conformation. Results from
the overlay of this structure with the Dpo4•DNA•dNTP ternary structure show no obvious structural change which appears to support the conclusion reached by these authors that the “induced-fit” mechanism does not apply to Y-family polymerases (41). However, this conclusion is contradictory to our kinetic observations for Dpo4 and the previously published kinetic results of yPolη which suggest a rate-limiting protein conformational change during correct nucleotide incorporation (17). We speculate that the “closed” conformation observed in the Dbh crystal structure (40,41) is probably a result of crystal packing and does not reflect the inherent protein conformation in solution. This speculation is somewhat supported by the unusual fact that DNA and a correct incoming nucleotide are required in the crystallization buffer in order to grow the crystals of the full-length Dbh alone (40). It would be extremely interesting to see the NMR structures of Dbh and Dpo4 in solution. Alternatively, the protein conformational change for a Y-family polymerase could occur exclusively in the unique little finger domain, rather than the finger domain as observed in replicative and repair DNA polymerases. This possibility is unlikely considering that the little finger domain of Dpo4 is not an essential component of the active site (42). However, it is possible that a large swing of the little finger domain leads to minor adjustment of the positions of DNA, nucleotide, and the active site residues of Dpo4. Since the E•DNA•dNTP ternary complex of Dpo4 is solved (8), the structural insight into the first conformational change in Scheme 3.2 will be clear once the structure of the Dpo4•DNA binary complex is available. Based on the kinetic results of the two Y-family polymerases Dpo4 and yPolη (17), we further speculate the “induced-fit” mechanism may apply to all Y-family DNA polymerases.

processivity of Dpo4. Based on the rate constants derived from the processive DNA synthesis experiment (Figures 3.7 and 3.8), the average DNA dissociation and product
formation rates during processive polymerization were calculated to be 0.21 s\(^{-1}\) and 3.3 s\(^{-1}\), respectively. Processivity, defined as the ratio of the polymerization rate divided by the DNA dissociation rate, was calculated to be 16 at 37 °C for Dpo4. This was slightly higher than the processivity of yPol\(\eta\) (5.2) (17), and was similar to the previously estimated processivity of Dpo4 with the same DNA/enzyme ratio (20). The processivity of Dpo4 and yPol\(\eta\) is much lower than the processivity of replicative polymerases like T7 DNA polymerase (1500) and human mitochondrial DNA polymerase \(\gamma\) (2250) (43). The low processivity of the Y-family polymerases combined with their low fidelity is consistent with their role in the bypass of DNA lesions. In theory, the Y-family polymerases are envisioned to dissociate from DNA shortly after traversing a damaged site, allowing a high-fidelity replicative polymerase to “re-associate” with the DNA and continue processive synthesis. Since \(S.\ solfataricus\) is an aerobic crenarchaeon that metabolizes sulfur and grows optimally at 80 °C and pH 2-4 (19), its genome is more likely to undergo depurination and depyrimidination reactions. The resulting abasic lesions are probably bypassed by Dpo4, thus suggesting an adaptive feature provided by this enzyme for the survival of \(S.\ solfataricus\) under such harsh conditions (20). We are currently studying the fidelity of lesion bypass by Dpo4 using pre-steady state kinetics.

The high processivity of replicative polymerases is due to the function of the processivity factors associated with these polymerases, e.g. \(E.\ coli\) thioredoxin for T7 DNA polymerase, the small subunit for human mitochondrial DNA polymerase \(\gamma\), eukaryotic proliferating cell nuclear antigen (PCNA) for polymerases \(\delta\) and \(\varepsilon\) (44). The processivity of several Y-family polymerases including Dbh (45) and \(E.\ coli\) DNA polymerase IV (46) is greatly increased in the presence of cofactors, but the processivity of human DNA polymerase \(\kappa\) is not affected (47). Recently, a heterotrimer of three PCNA homologs, PCNA1, 2, and 3, from \(S.\ solfataricus\) was found to dramatically increase the
processivity of *S. solfataricus* DNA polymerase B1 (48,49). It will be interesting to test the effect of the *S. solfataricus* PCNA heterotrimer on the processivity of Dpo4. If the processivity of Dpo4 is increased greatly, the protein complex could be used in the error-prone polymerase chain reactions (PCR) to engineer random mutations since the fidelity of Dpo4 is low (18,20,50). The biological implication of the potential interactions between Dpo4 and the PCNA heterotrimer requires additional studies.

**DNA binding affinity.** The affinity of Dpo4 for a synthetic DNA substrate, D-1, was measured to be 10.6 nM using the active site titration assay (Figure 3.4). This indicates Dpo4 binds DNA with the same affinity as replicative polymerases such as T7 DNA polymerase (18 nM) (23), human mitochondrial DNA polymerase γ (9.9 nM) (43), HIV-1 reverse transcriptase (4.7 nM) (26), and the repair enzyme *E. coli* DNA polymerase I (5 nM) (36). This is not surprising since, overall, the DNA binding pocket in polymerases is quite large and the interactions between DNA and these polymerases are quite intense. The little finger domain and the tether linking the thumb and little finger domain of Dpo4, which are absent in replicative polymerases, wrap around DNA acting like a small clamp (8). These unusual structural features compensate for the abnormally small thumb and finger domains of Dpo4, increasing the DNA binding affinity. This is confirmed by the significantly lower processivity of a truncated Dpo4 mutant lacking the little finger domain when compared to the processivity of wild-type Dpo4 (8).

The faster dissociation of the ternary complex, E•DNA•dNTP (0.41 s⁻¹), compared to the binary complex, E•DNA (0.02-0.07 s⁻¹), is very intriguing. Similar results have been observed for HIV-1 RT (26). This phenomenon suggests that correct dNTP binding causes an increase in the rate of DNA dissociation from the enzyme. The structural implication of faster DNA dissociation is not clear. We speculate the binding of a correct
nucleotide may disrupt various specific interactions between DNA and the enzyme such as hydrogen bonding and van der Waals contacts. However, the faster protein conformational change, \( E \cdot \text{DNA} \cdot \text{dNTP} \leftrightarrow E' \cdot \text{DNA} \cdot \text{dNTP} \) (9.4 s\(^{-1}\)), in the presence of a correct incoming nucleotide, will limit the effect that the dissociation of DNA from \( E \cdot \text{DNA} \cdot \text{dNTP} \) has on substrate elongation due to favorable kinetic partitioning. More importantly, this protein conformational change step leads to the formation of a tighter binding state \( E' \cdot \text{DNA} \cdot \text{dNTP} \) which dissociates at 0.004 s\(^{-1}\) (Figure 3.10). Thus, the protein conformational change results in a tighter binding of DNA to Dpo4. Similar results have been observed in HIV-1 reverse transcriptase (26).

*Rate-limiting step in the incorporation of a mismatched nucleotide.* The elemental effect for the incorporations of a mismatched dGTP and S\(\_\)dGTP\(\alpha\)S was measured to be 5.8 (Figure 3.6), suggesting the chemistry step is rate-limiting for the incorporation of incorrect nucleotides. This conclusion is solely based on the elemental effect of 4 to 11 observed with the hydrolysis of phosphate diesters, which is limited by the chemistry step (29). Interestingly, the elementary effect observed with Dpo4 is similar to the value of 5.0 observed with correct nucleotide incorporation by rPol\(\beta\) (31). Recent structural and stopped-flow fluorescence studies have determined the rate-limiting step of DNA polymerization by rPol\(\beta\) is in fact the chemistry step (51). The combined results from these studies of rPol\(\beta\), support our results with Dpo4 for a rate-limiting chemistry step for incorrect nucleotide incorporation. Moreover, large elemental effects associated with mismatched nucleotides incorporation has been observed with other polymerases including T7 DNA polymerase (17-34) (24) and *E. coli* DNA polymerase I (65) (35). Thus, the assignment of the rate-limiting chemistry step for the incorporation of mismatched nucleotides based on the elemental effect seems to be a general trend for all DNA polymerases. However, a smaller elemental effect (1.9) is observed in the
mismatched nucleotide incorporation catalyzed by ypO1 (17). This exception is not clearly understood. Additional mechanistic evidence other than the elemental effect is needed to evaluate whether the chemistry step limits the incorrect nucleotide incorporation by Dpo4 and ypO1.
3.5 Figures, Tables, Schemes

D-1

5’-CGCAGCGGTCCCAACCACTCA-3’
3’-GCGTCGGCAGGTTGGTTGAGTAGCAGCTAGGTTACGGCAGG-5’

Figure 3.1. Sequence of the D-1 DNA substrate.
Figure 3.2. Pre-steady state and steady-state kinetics of dTTP incorporation into D-1 by Dpo4. A preincubated solution of Dpo4 (30 nM) and 5'-32P-labeled-D-1 (120 nM) was rapidly mixed with dTTP (0.10 mM). The reactions were quenched at various times with 0.37 M EDTA and products were quantitated by sequencing gel analysis. The data were fit by nonlinear regression to a biphasic curve (equation 1) with rate constants equal to 3.8 ± 0.2 s$^{-1}$ and 0.020 ± 0.002 s$^{-1}$ for the exponential and linear phases respectively.
Figure 3.3. Steady-state kinetics of dTTP incorporation into D-1 by Dpo4. dTTP incorporation into D-1 was independently measured under steady-state conditions by preincubating Dpo4 (1.2 nM) and 5'32P-labeled-D-1 (250 nM) and then starting the reactions with the addition of Mg2+-dTTP (0.10 mM). Reactions were terminated at various times by the addition of 0.37 M EDTA and quantitated. The data were fit to a straight line (equation 2) and the steady-state rate constant was calculated from the slope of this line divided by the enzyme concentration to give a rate constant equal to 0.07 s−1.
Figure 3.4. Active site titration of Dpo4 with D-1. A solution of Dpo4 (22 nM as determined by UV absorbance measurements) was preincubated with increasing concentrations of 5'-32P-labeled-D-1 and subsequently mixed with a solution containing Mg2+-dTTP (0.10 mM) in a rapid chemical quench flow apparatus. The reactions were quenched after 1.2 sec and the products were analyzed by sequencing gel electrophoresis. The burst amplitude was then plotted as a function of substrate concentration and the data were fit to the quadratic equation (equation 3) which gave a $K_d$ for the Dpo4-D-1 complex of $10.6 \pm 1.3$ nM and an enzyme amplitude of $19.4 \pm 0.5$ nM.
Figure 3.5. $S_p$-dTTP($\alpha$S) elemental effect on the rate of correct nucleotide incorporation. A preincubated solution of Dpo4 (120 nM) and $5'$-$^{32}$P-labeled-D-1 (30 nM) was mixed with either 0.10 mM dTTP (●) or $S_p$-dTTPαS (○) in parallel time courses. The products were quantitated and the data were fit by nonlinear regression to the single exponential equation (equation 4) yielding $k_{obs}$ values of 6.27 ± 0.38 s$^{-1}$ and 4.49 ± 0.30 s$^{-1}$ for dTTP and $S_p$-dTTPαS, respectively. The elemental effect of 1.4 was calculated from the ratio of the $k_{obs}$ values for correct nucleotide incorporation into D-1.
Figure 3.6. Sp-dGTP(αS) elemental effect on the rate of incorrect nucleotide incorporation. A preincubated solution of Dpo4 (120 nM) and 5'-32P-labeled-D-1 (30 nM) was mixed with either 0.10 mM dGTP (●) or Sp-dGTPαS (○) in parallel time courses. The products were quantitated and the data were fit to the single exponential equation (equation 4) yielding $k_{obs}$ values of 0.0021 ± 0.0001 s$^{-1}$ and 0.00037 ± 0.00002 s$^{-1}$ for dGTP and Sp-dGTPαS, respectively. The elemental effect of 5.8 was calculated from the ratio of the $k_{obs}$ values for incorrect nucleotide incorporation into D-1.
Figure 3.7. Processive polymerization by Dpo4. A preincubated solution of Dpo4 (35 nM) and 5'-32P-labeled-D-1 (100 nM) was rapidly mixed with dTTP, dCTP, and dGTP (1.2 mM each) in the reaction buffer containing Mg2+. The reactions were quenched at various times with 0.37 M EDTA and analyzed by sequencing gel analysis. The amount of remaining substrate 21-mer (■) and each product formed (22-mer, ●; 23-mer, △; 24-mer, ◆; 25-mer, □; 26-mer, ○; 27-mer, ◊) was plotted as a function of the reaction time. The solid lines represent best fits obtained from computer simulation using a mechanism consisting of a series of six single nucleotide incorporations at 2.3 ± 0.1 (22-mer), 2.9 ± 0.2 (23-mer), 3.5 ± 0.4 (24-mer), 4.3 ± 0.6 (25-mer), 3.9 ± 0.6 (26-mer), and 3.4 ± 0.4 s⁻¹ (27-mer), respectively; and DNA dissociations at 0.26 ± 0.05 s⁻¹ (21/41-mer), 0.18 ± 0.11 s⁻¹ (22/41-mer), 0.19 ± 0.12 s⁻¹ (23/41-mer), 0.18 ± 0.16 s⁻¹ (24/41-mer), 0.18 ± 0.18 s⁻¹ (25/41-mer), 0.15 ± 0.11 s⁻¹ (26/41-mer), 0.20 ± 0.09 s⁻¹ (27/41-mer).
Figure 3.8. Image of the processive polymerization by Dpo4. A preincubated solution of Dpo4 (35 nM) and 5'-32P-labeled-D-1 (100 nM) was rapidly mixed with dTTP, dCTP, and dGTP (1.2 mM each) in the reaction buffer containing Mg2+. The reactions were quenched at various times with 0.37 M EDTA and analyzed by sequencing gel analysis. This image of the results from the processivity experiment shows the progress of the six single nucleotide incorporations at various reaction times.
Figure 3.9. Measurement of the DNA dissociation rate of the E•DNA•dNTP complex. A preincubated solution of Dpo4 (55 nM) and 5'-32P-labeled-D-1 (60 nM) was mixed with Mg$_{2+}$-dTTP$\alpha$S (1.2 mM) in the absence (○) or presence (●) of unlabeled trap D-1 (2.5 μM). The rate of incorporation in the both experiments was 11.0 s$^{-1}$ however, the burst amplitudes were 53.1 ± 0.6 nM and 51.2 ± 0.7 in the absence and presence of DNA trap, respectively. This reduction in amplitude (3.6%) in the presence of trap suggests a DNA dissociation rate from this ternary complex of 0.41 s$^{-1}$. 

87
Figure 3.10. Measurement of the DNA dissociation rate of the E’•DNA•dNTP complex. A preincubated solution of Dpo4 (1.2 nM) and 5’-32P-labeled-D-1 (200 nM) was mixed with Mg2+-ddTTP (1.2 mM) in the absence (●) or presence (○) of the next correct nucleotide dCTP (1.2 mM). The steady-state rate constant in the absence of dCTP was 0.028 ± 0.001 s⁻¹ and corresponds to the rate of dissociation of DNA from the E•DNA binary complex while the rate constant in the presence of dCTP was 0.0040 ± 0.0003 s⁻¹ and corresponds to the rate of dissociation of DNA from the E’•DNA•dNTP ternary complex.
Figure 3.11. Pulse-chase and pulse-quench experiment. A preincubated solution of Dpo4 (30 nM) and unlabeled D-1 (30 nM) was mixed with $[\alpha^{-32}\text{P}]-\text{dTTP}$ (50 μM) for time intervals ranging from 0.05 to 4.5 s. The reactions in these two experiments were either quenched directly with 1 M HCl (pulse-quench) or chased with 2.5 mM unlabeled dTTP (pulse-chase) for 30 s, followed by HCl quenching. Both experiments were fit to the single exponential equation (equation 4) yielding amplitudes of 20 ± 0.6 nM and 25.5 ± 1.0 nM, and observed rates of 0.62 ± 0.17 s$^{-1}$ and 0.62 ± 0.24 s$^{-1}$ for the pulse-quench (○) and pulse-chase (●) experiments, respectively.
### Table 3.1. Estimated Kinetic Constants of Dpo4 at 37 °C.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Constants</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$</td>
<td>$1.9 , \mu\text{M}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td>$k_{-1}$</td>
<td>$0.02 , \text{s}^{-1}$</td>
</tr>
<tr>
<td>$K_{d, \text{DNA}}$</td>
<td>$10.6 , \text{nM}$</td>
</tr>
<tr>
<td>$k_2$</td>
<td>$100 , \mu\text{M}^{-1}\text{s}^{-1}$</td>
</tr>
<tr>
<td>$k_{-2}$</td>
<td>$23,000 , \text{s}^{-1}$</td>
</tr>
<tr>
<td>$K_{d, \text{dNTP}}$</td>
<td>$230 , \mu\text{M}$</td>
</tr>
<tr>
<td>$k_3$</td>
<td>$9.4 , \text{s}^{-1}$</td>
</tr>
<tr>
<td>$k_7$</td>
<td>$0.41 , \text{s}^{-1}$</td>
</tr>
<tr>
<td>$k_8$</td>
<td>$0.004 , \text{s}^{-1}$</td>
</tr>
</tbody>
</table>
Scheme 3.1. Structures of the $S_p$ and $R_p$ isomers of dTTPαS.
Scheme 3.2. Detailed minimal kinetic mechanism for dNTP incorporation into a DNA substrate catalyzed by Dpo4.
3.6 References


CHAPTER 4

MECHANISM OF ABASIC LESION BYPASS CATALYZED BY SULFOLOBUS SOLFATARICUS DNA POLYMERASE IV AT 37 °C

4.1 Introduction

Numerous DNA-damaging agents continuously attack the cellular genome and generate a myriad of DNA lesions. Although a majority of these lesions are repaired by DNA repair pathways, some damage evades repair. Unrepaired lesions arrest DNA replication by inhibiting replicative DNA polymerases. However, a recently identified class of low fidelity enzymes, known as the Y-family DNA polymerases, can replicate through DNA damage and rescue cells from apoptosis. In humans, four Y-family DNA polymerases have been identified (1). Due to its thermostability and ease of purification (2) we have decided to use Sulfolobus solfataricus DNA polymerase IV (Dpo4) as a model Y-family DNA polymerase to study the mechanistic basis of DNA synthesis. Dpo4, which retains catalytic activity at 37 °C, has been shown to exhibit a low fidelity in the range of $10^{-3}$ to $10^{-4}$, or one error for every 1,000 to 10,000 nucleotides incorporated on undamaged DNA as revealed by our pre-steady state kinetic studies (3) and by both steady-state kinetics and a forward mutation assay performed by others (2,4). Notably, Dpo4 and a human Y-family member, DNA polymerase κ (hPolκ), are both DinB homologs and have been shown to possess very similar error rates (5,6), yet, a previous study (2) suggests that the lesion bypass properties of Dpo4 are more similar to eukaryotic DNA polymerase η.
As such, due to its functionally and structurally conserved features with respect to other Y-family enzymes, Dpo4 is considered to be the archetypical Y-family DNA polymerase for studies elucidating the mechanism and function of this novel class of enzymes.

Careful kinetic analysis of the mechanistic basis of nucleotide incorporation into undamaged DNA catalyzed by the Y-family DNA polymerases has been reported for yeast Polη (7), Dpo4 (3,8), and *Sulfolobus acidocaldarius* Dbh (9). However, besides studies reporting several kinetic parameters and mutation profiles for the bypass of 7,8-dihydro-8-oxodeoxyguanine (8-oxodG) by yeast Polη (10), and for the bypass of a cis-syn thymine-thymine (T-T) dimer catalyzed by yeast Polη (11), thorough mechanistic analysis of a Y-family polymerase promoting lesion bypass has not been reported. In this paper, we report the first comprehensive characterization of the mechanism of lesion bypass catalyzed by Dpo4 through rigorous pre-steady state kinetic analysis. It has been well established that apurinic/apyrimidinic (AP) sites resulting from the hydrolytic cleavage of the *N*-glycosidic bond are among the most abundant lesions encountered in a mammalian cell (12) with approximately 10,000 spontaneous AP sites generated in each cell every day (13,14). It is plausible that AP sites will be generated at a similar or higher frequency in *S. solfataricus* since this aerobic crenarchaeon propagates at roughly 80 °C (15). While replication past these noninstructive lesions is a mechanistic challenge, replicative and repair DNA polymerases have been shown to preferentially incorporate dATP, although inefficiently, opposite the AP lesion in a phenomenon known as the A-rule (16). Yet, recent structural studies have concluded that instead of the A-rule, Dpo4, and thus other DinB homologs, almost exclusively uses the template base 5' to the lesion (referred to as the 5'-rule) to instruct nucleotide incorporation during AP bypass (17). However, our thorough examination of the kinetic effects of AP bypass has clearly
illustrated a sequence dependent competition between these two proposed pathways in disagreement with the conclusions reached from these previous reported structural studies.

4.2 Materials and methods

*Reaction buffer R.* Buffer R contains 50 mM HEPES-NaOH (pH 7.5 at 37 °C), 5 mM MgCl₂, 50 mM NaCl, 5 mM DTT, 10% glycerol, 0.1 mM EDTA, and 0.1 mg/ml BSA (3). All concentrations reported here refer to the concentration of components after mixing. All reactions, unless noted, were carried out at 37 °C.

*Running start assay.* Experiments were carried out in a rapid chemical quench flow apparatus (KinTek) by rapidly mixing an aliquot (15 µl) of a solution containing 100 nM 5'-[³²P] DNA and 100 nM Dpo4 preincubated in buffer R with an aliquot (15 µl) containing all four dNTPs (200 µM each) for times ranging from milliseconds to minutes followed by quenching with 0.37 M EDTA. The nucleotide incorporation pattern was resolved by sequencing gel analysis. Assays performed with PolB1 were carried out using the same buffer and reaction conditions used for Dpo4.

*Determination of the kp and Kd of an incoming nucleotide.* A solution containing Dpo4 (120 nM) and 5'-[³²P] DNA (30 nM) preincubated in buffer R was mixed with increasing concentrations of a single nucleotide. Reactions were terminated by the addition of EDTA. Products were separated from substrate via sequencing gel electrophoresis (17%
acrylamide, 8 M urea) and quantitated using a PhosphorImager 445 SI (Molecular Dynamics). The time course of product formation was fit to equation 1:

$$[\text{Product}] = A[1 – \exp(-k_{\text{obs}}t)]$$  \hspace{1cm} (1)

for each concentration of dNTP to yield an observed rate constant ($k_{\text{obs}}$) and a reaction amplitude ($A$). The extracted $k_{\text{obs}}$ values were then plotted as a function of the concentration of dNTP and fit to equation 2:

$$k_{\text{obs}} = \frac{k_p [\text{dNTP}]}{[\text{dNTP}] + K_d}$$  \hspace{1cm} (2)

to give $k_p$ and $K_d$. The substrate specificity ($k_p/K_d$) was then calculated.

### Electrophoretic mobility shift assay

Mobility shift titrations were conducted using a 4.5% native polyacrylamide gel with a running buffer (50 mM Tris-Acetate, pH 7.5 at 23 °C, 5.5 mM Mg(OAc)$_2$, and 0.5 mM EDTA). Dpo4 (15 nM to 425 nM) was titrated into a solution of 5’-[32P] DNA (100 nM) in buffer R at 23 °C. Aliquots of the titration were loaded into the native gel and run at a constant voltage of 80 V for 30 minutes at 23 °C. The dried gels were exposed to a PhosphorImager and quantitated. The concentration of Dpo4•DNA was plotted against the concentration of Dpo4 and fit using equation 3:

$$[E\cdot\text{DNA}] = 0.5(K_d + E_o + D_o) - 0.5[(K_d + E_o + D_o)^2 - 4E_oD_o]^{1/2}$$  \hspace{1cm} (3)

where $E_o$ represents the active enzyme concentration and $D_o$ the DNA concentration.

### 4.3 Results

Previously we exploited transient-state kinetic analyses to measure the fidelity and establish a minimal kinetic mechanism for single nucleotide incorporation into an undamaged DNA substrate catalyzed by Dpo4 (3,8). These results demonstrated that
Dpo4 uses an induced-fit mechanism to select a correct nucleotide as observed for several replicative and repair polymerases (18) and another Y-family DNA polymerase, yeast Polη (7). This mechanism served as a foundation to establish the mechanistic basis for bypassing DNA lesions such as AP sites, catalyzed by Dpo4. Notably, natural AP sites exist as an uneven mixture of four species, with the equilibrium favoring the two anomeric hemiacetals over the two open chain aldehydes (19,20). However, the aldehydic forms are subject to β-elimination and subsequent scission of the DNA backbone, producing a heterogeneous mixture of DNA substrates. Due to this instability, we have chosen to perform our studies with the stable AP analog tetrahydrofuran (THF) since a homogeneous population of DNA substrates is requisite for rigorous kinetic studies. Most crystal structure studies and kinetic assays involving AP-containing DNA have been carried out using THF (17,21-26). Moreover, THF has been shown to retain the biological properties of the natural AP site *in vivo* (27).

**Bypassing an AP lesion.** To investigate the response of Dpo4 to an AP lesion, we designed a “running start” assay (Materials and Methods) such that the effect of the AP lesion on DNA synthesis could be determined via observation of the polymerization pattern over time at 37 °C for reasons described previously (3). Elongation of 5’-[32P]-labeled 17-mer/41AP (Table 4.1) proceeded rapidly up until the template AP site where the incorporation pattern showed two consecutive strong pause sites (Figure 4.1), corresponding to incorporation opposite the AP site and its extension. At these two strong pause sites, the nucleotide incorporation pattern showed a significant accumulation of intermediate products 21-mer and 22-mer (Figure 4.1), suggesting slow turnover. Nonetheless, the AP site was bypassed relatively efficiently by Dpo4 (full-length product was observed at 40 s) and subsequent downstream incorporation was not significantly perturbed by the embedded AP lesion. In comparison, Dpo4 pausing was not observed
and the full-length product was formed after 10 s (Figure 4.1) in the control reactions with an undamaged 17-mer/41CTL (Table 4.1). Interestingly, Dpo4 was able to catalyze blunt-end addition to the full length 41-mer to generate 42-mer as described previously (28). Notably, the same running start assay was performed with *S. solfataricus* replicative DNA polymerase PolB1, and showed no bypass of the AP site, instead stalling predominantly one nucleotide preceding the lesion (Figure 4.2).

**Measurement of the binding of Dpo4 to DNA containing an AP site.** It is well established that AP sites have a destabilizing effect on the DNA duplex stability (29,30) and conformation (31,32). It is thus plausible that the strong pause sites observed proximal to the lesion (Figure 4.1) could be due to reduced binding of Dpo4 for the AP-containing DNA. To evaluate this hypothesis, the electrophoretic mobility shift assay (EMSA) was employed to measure overall binding affinity of Dpo4 to a series of lesion containing DNA substrates. In these substrates, primers of increasing length were designed (Table 4.1) such that when annealed to the template 41AP they would structurally mimic the progression of DNA synthesis as Dpo4 incorporated one nucleotide at a time. These measurements allowed for the estimation of the equilibrium dissociation constant ($K_d$) of the Dpo4•DNA binary complex, as Dpo4 incrementally approached and proceeded downstream from the AP site. For example, a solution containing the 5’-[32P]-labeled 21-mer/41AP was titrated with increasing concentration of Dpo4 (Materials and Methods) and the $K_d$ of Dpo4•21-mer/41AP complex was estimated after gel analysis (Figure 4.3) to be 36 ± 2 nM (Figure 4.4). Similarly, the $K_d$ for the binding of Dpo4 to undamaged 21-mer/41CTL was estimated to be 10 ± 2 nM. This value was very similar to 11 ± 1 nM which was measured previously using active site titration (8) suggesting that EMSA was a reliable assay to measure the $K_d$ of the Dpo4•DNA binary complex.
Our kinetic analysis (see below) indicated that while incorporation opposite the AP site (first pause site) was ~210-fold slower than a matched incorporation, 90% of the incorporation events at this site were either dATP or dCTP, suggesting two dominant pathways to bypass the AP lesion. To observe the effect of an AP site on DNA binding to Dpo4, the DNA substrates assayed were designed to precisely represent the sequence contexts at the two strong pause sites as well as at both upstream and downstream nonpause sites (Figure 4.1) with either adenine or cytosine opposite the AP site when applicable (Table 4.2). Interestingly, the Dpo4-binding affinities at the two strong pause sites were less than 4-fold lower than their corresponding controls while at nonpause sites all the $K_d$ values were within 2-fold of their corresponding control substrates (Table 4.2). Thus the flexibility from the AP lesion did not significantly perturb Dpo4 binding and therefore was not the major contributor to the observed strong pausing of Dpo4 in Figure 4.1. However, the local structural flexibility at the AP site may hinder nucleotide incorporation. We subsequently hypothesized that these strong pause sites were generated by a significant decrease in the incorporation efficiency of an incoming nucleotide.

**Efficiency of nucleotide incorporation in the vicinity of the AP lesion.** To test the aforementioned hypothesis, single turnover kinetic experiments used previously (3) were employed to measure the maximum incorporation rate constant ($k_p$) and the $K_d$ of an incoming dNTP which are used to calculate catalytic efficiency ($k_p/K_d$). Initially we measured the $k_p$ and $K_d$ for nucleotide incorporation into 21-mer/41AP, corresponding to the DNA substrate at the first pause site where the AP site was the first ‘templating’ position. Notably, this substrate also contained a template guanine immediately 5’ to the AP site (Table 4.1). Assuming that Dpo4 would follow the A-rule for incorporation opposite the AP site, a solution containing Dpo4 (120 nM) was preincubated with 5’-radiolabeled 21-mer/41AP (30 nM) and reacted with increasing concentration of dATP
for various times (Materials and Methods). The resulting product concentrations were plotted against reaction time and fit to equation 1 (Materials and Methods) to yield individual $k_{obs}$ values at each corresponding dATP concentration (Figure 4.5). The extracted $k_{obs}$ values were then plotted against the dATP concentration and fit to equation 2 (Materials and Methods) to yield a $k_p$ of $0.015 \pm 0.002 \text{ s}^{-1}$ and a $K_d$ of $490 \pm 173 \text{ µM}$ (Figure 4.6). The $k_p/K_d$ of dATP was then calculated to be $3.1 \times 10^{-5} \text{ µM}^{-1}\text{s}^{-1}$ (Table 4.3), which was $\sim2,520$-fold slower than a matched nucleotide incorporation into the control substrate 21-mer/41CTL (3). Strikingly, we observed that the catalytic efficiency for this and the remaining three nucleotide incorporations (dCTP, dGTP, and dTTP) into the 21-mer/41AP substrate dropped to the level observed previously for incorrect nucleotides ($k_p/K_d \sim 10^{-5} \text{ µM}^{-1}\text{s}^{-1}$) into 21-mer/41CTL (3). Similar results were observed for all incorporations at the second pause site (Table 4.3) where incorporation efficiencies were in the range of $10^{-4} - 10^{-6} \text{ µM}^{-1}\text{s}^{-1}$. Since all the catalytic efficiencies for incorporations at these two strong pause sites were of similar magnitude (Table 4.3) and several orders of magnitude smaller than a correct incorporation into the undamaged DNA (3), these data kinetically justified our hypothesis that the strong pause sites were caused by a significant reduction in catalytic efficiency.

Many previous studies have demonstrated a proclivity of DNA polymerases to preferentially incorporate dATP opposite an AP site (A-rule). Table 4.3 indicated that Dpo4 followed the A-rule by favoring dATP over dGTP and dTTP into 21-mer/41AP, however, Dpo4 failed to incorporate dATP ($3.1 \times 10^{-5} \text{ µM}^{-1}\text{s}^{-1}$) with a higher efficiency than dCTP ($5.4 \times 10^{-5} \text{ µM}^{-1}\text{s}^{-1}$) in this sequence context. In contrast, the $k_p$, $K_d$, and calculated $k_p/K_d$ values for both dGTP and dTTP were more similar to those of the least efficient misincorporation into an undamaged 21-mer/41CTL (3), and were 39-fold and 6-fold less likely to be incorporated, respectively (Table 4.3). The kinetic preference (1.7-
fold) for dCTP over dATP incorporation may be due to the formation of a Watson-Crick basepair between dCTP and the template guanine 5’ to the AP site. To evaluate this hypothesis, we synthesized a template 22AP that was identical in sequence to 41AP, except that it lacked all the template bases 5’ to the AP site (Table 4.1). Upon measuring the catalytic efficiency for each single nucleotide incorporation into 21-mer/22AP, we observed dCTP was incorporated 33-fold less efficiently than dATP, indicating that the template information 5’ to the AP site significantly influenced nucleotide incorporation (Table 4.4). These results, coupled with structural analysis (17) and additional data from our laboratory (see Discussion) demonstrated the existence of two competing pathways for bypass of an AP site catalyzed by Dpo4; the A-rule and the previously described 5’-rule (17) which involves incorporation directed by the template base 5’ to the lesion in the context of an extrahelical AP site (Scheme 4.1).

Nucleotide incorporation downstream from the AP site. Under single turnover conditions, we also determined the incorporation efficiency of each dNTP for several downstream incorporation events from the AP site (Tables 4.5 and 4.6). Our data revealed that the effect from the embedded AP site on nucleotide incorporation decreased with distance, regardless of the pathway (A-rule or 5’-rule) (Figures 4.7 and 4.8). Interestingly, the relatively low $k_p/K_d$ values before normalization suggested that Dpo4 paused slightly at several downstream incorporation events. This observed ‘weak pausing’ was due to a lingering effect on the $k_p$ which incrementally diminished as Dpo4 moved farther from the embedded lesion, while all $K_d$ values varied as observed previously with undamaged DNA (3). For the pathway with adenine incorporated opposite the AP site, the kinetic effect of an embedded lesion normalized after six incorporations (Figure 4.7) while the pathway with cytosine opposite the AP site normalized after five incorporations (Figure
4.8). The ‘weak pause sites’ were much less noticeable (Figure 4.1) due to their reduced effect on further downstream synthesis, coupled with significantly longer reaction times.

Measurement of the possibility of primer realignment after bypassing the AP lesion. After Dpo4 bypassed the AP lesion via the 5’-rule, the embedded AP lesion may remain looped-out or the primer may realign with the DNA template to force the incorporated cytosine to be located opposite the AP site during downstream incorporation events (Scheme 4.1). To determine the dominant branch pathway in Scheme 4.1, we evaluated the $k_p$, $K_d$, and $k_p/K_d$ for all four possible nucleotide incorporations at each position downstream from the lesion (Tables 4.5 and 4.6). For each substrate listed in Table 4.7, only the two dNTPs relevant to the looped-out and realignment pathways were shown. The realignment% values indicated that the looped-out branch pathway was dominant to the realignment pathway especially when the embedded AP lesion was further from the nascent basepair.

Sequence dependence of the 5’-rule. To determine if the 5’-rule was sequence dependent, we measured the kinetic parameters (Table 4.8) for nucleotide incorporation into two additional DNA substrates, AP-1 and AP-8 (Table 4.9), where the template base immediately 5’ to the AP site was changed to an adenine and a cytosine respectively (thymine is irrelevant in this context). Interestingly, the competition between the A-rule and the 5’-rule pathways was similar in magnitude for both AP-1 and AP-8 as observed with the substrate containing a 5’ template guanine (21-mer/41AP, Table 4.3). For the AP-1 substrate, the competition between the A-rule and 5’-rule was indistinguishable while for the AP-8 substrate, the A-rule was favored by ~4-fold. All other incorporations, therefore incorporations other than dCTP (21-mer/41AP)/dTTP (AP-1)/dGTP (AP-8) or dATP into these three DNA substrates were on average 18-fold less likely to occur (Table
4.8). Incidentally, incorporation of dTTP (16.5%) had a similar probability to the incorporation of dGTP (17.3%) into AP-8 (Table 4.8), yet could not be explained based on these two bypass pathways. We had previously reported a similar observation where Dpo4 incorporated nucleotides onto a blunt-end DNA substrate with the following preference: dATP>dTTP>dCTP>dGTP (28). In addition, a previous report demonstrated a nearly equal probability of incorporation of dATP and dTTP opposite this lesion in vivo (33). Ultimately, we do not know the reason for the increased probability of dTTP incorporation into AP-8, but we suspect this trend is likely polymerase-dependent. Intriguingly, in these three sequence contexts (21-mer/41AP, AP-1, and AP-8), the preference for the A-rule was only observed in the presence of a 5’ template cytosine. Thus, due to the inherent ambiguity in the nomenclature for the 5’-rule (17) as revealed by our aforementioned sequence dependence studies, we have decided to rename the pathway the ‘lesion loop-out mechanism’ from hereafter.

_Incorporation of a nonhydrogen bonding nucleotide analog opposite an AP lesion._
Interestingly, previous studies have shown that yeast Polη requires Watson-Crick hydrogen bonds for efficient incorporation, unlike the geometric discrimination that helps govern the fidelity at the active sites of replicative DNA polymerases (34). To determine the effect of a nucleotide analog that lacked the ability to hydrogen bond with a template base on the incorporation opposite an AP site by Dpo4, we measured the $k_p/K_d$ of pyrene nucleoside 5’-triphosphate (dPTP, Figure 4.9) with 21-mer/41AP. dPTP served as a bifunctional probe due to its inability to hydrogen bond and its robust base-stacking abilities, attributed to its resonance stabilized aromatic ring system. The single turnover experiments were carried out as described above and we determined the $k_p$ to be 0.26 ± 0.02 s$^{-1}$, the $K_d$ to be 24 ± 4 μM (Figure 4.10), and the $k_p/K_d$ to be 1.1 x 10$^{-2}$ μM$^{-1}$s$^{-1}$ which was, strikingly, 358- and 205-fold higher than dATP and dCTP incorporation.
respectively (Table 4.3). The 20-fold higher ground-state binding affinity of dPTP over dATP (Table 4.3) indicated base stacking was the main contributor to the binding affinity of a nucleotide opposite an AP lesion at the Dpo4 active site.

**Biphasic kinetics of nucleotide incorporation at the first pause site.** To provide a more detailed analysis for the kinetics of nucleotide incorporation at the first strong pause site, a preincubated solution of Dpo4 (120 nM) and 5'-[^32P]-labeled 21-mer/41AP (30 nM) was mixed with a solution containing a non-radiolabeled 21-mer/41CTL DNA trap (5 µM) and either 1.2 mM dATP or 1.2 mM dCTP for various times before being quenched with 0.37 M EDTA. The product concentrations were calculated and fit to equation 4

\[
[\text{Product}] = E_o A_1 [1-\exp(-k_1 t)] + E_o A_2 [1-\exp(-k_2 t)]
\]

where \(E_o\) represents the total enzyme concentration, \(A_1\) and \(A_2\) represent the fast and slow phase reaction amplitude respectively, and \(k_1\) and \(k_2\) represent the rate constants for the fast phase and slow phase respectively. The trap DNA (~170-fold excess) functioned to remove all Dpo4 molecules that dissociated from the lesion-containing substrate. Interestingly, the results for both dATP and dCTP incorporation revealed biphasic kinetics, showing both a low amplitude, fast phase of nucleotide incorporation preceding a higher amplitude, slow phase. For dATP incorporation, the fast phase was determined to proceed with \(k_1\) of \(0.25 \pm 0.10\) s\(^{-1}\) and \(A_1\) of \(1.7 \pm 0.1\) nM (or 5.7%) while the slow phase was characterized by a \(k_2\) of \(0.00124 \pm 0.00009\) s\(^{-1}\) and \(A_2\) of \(22.2 \pm 0.8\) nM (or 74%) (Figure 4.11). For dCTP incorporation, the fast phase had \(k_1\) of \(0.24 \pm 0.06\) s\(^{-1}\) and \(A_1\) of \(7.1 \pm 0.3\) nM (or 23.7%) while the slow phase was determined to have \(k_2\) of \(0.0005 \pm 0.0004\) s\(^{-1}\) and \(A_2\) of \(18.2 \pm 5.9\) nM (or 60.7%) (data not shown). Although Dpo4 was in molar excess to initially ensure all DNA substrate was bound, neither time course reached full amplitude (80% for dATP, and 84% for dCTP), suggesting that a portion of Dpo4
bound to the DNA substrate either dissociated or was unable to change from a nonproductively bound state to a productively bound state competent for catalysis.

4.4 Discussion

The evolution of DNA repair pathways has provided cells with several different mechanisms to detect and eliminate DNA damage. However robust these systems may be, a population of DNA lesions will persist, halting the progression of the replication machinery and jeopardizing cell survival. With the recent discovery of a group of polymerases that promiscuously replicate through various types of DNA damage, we now know that lesion tolerance is a bona fide mechanism to rescue stalled replication forks. Here, we use the powerful techniques of pre-steady state kinetics to dissect the mechanism of lesion bypass by the Y-family member Dpo4 as it approaches, incorporates across from, and synthesizes downstream from an embedded AP site lesion. Bypass of an AP lesion by Dpo4 was accomplished very efficiently in comparison to the replicative DNA polymerase PolB1 from *S. solfataricus*, which was completely stalled by the same AP lesion even after ten minute incubation under the same reaction conditions (Figure 4.2). However, in the process of AP lesion bypass, intermediate product accumulation indicated that Dpo4 paused strongly at one nucleotide preceding the AP site, which has never been reported previously, and at the position opposite the lesion. This strong pausing (Figure 4.1) suggested a perturbation in the mechanism of nucleotide incorporation during bypass.
Kinetic studies reveal competition between two bypass pathways. Our kinetic data in Table 4.3 demonstrated that Dpo4 preferentially incorporated dATP and dCTP when it bypassed a template AP site containing a guanine 5′ to the lesion (Table 4.1). While dATP was likely incorporated via the A-rule, incorporation of dCTP was governed by a different mechanism. Due to the well established structural flexibility of AP site lesions in DNA (29,30,35), the preference for dCTP led us to hypothesize that this nucleotide could be incorporated via a mechanism directed by the downstream template guanine involving an extrahelical AP site. Further evidence showing: (i) loss of the preference of dCTP incorporation into 21-mer/22AP (Table 4.4) relative to 21-mer/41AP (Table 4.3); (ii) extension assays of Dpo4 with the 21-mer/41AP (Table 4.1) in the presence of dideoxynucleotides showing full-length products containing -1 deletions (data not shown); and (iii) crystal structure evidence showing the ability of Dpo4 to ‘loop-out’ an AP site (17), led us to conclude that incorporation of dCTP was due to the lesion loop-out mechanism (Scheme 4.1). Structure analysis showed the looped-out AP lesion can be accommodated in the cavity between the finger and little finger domains (17). In contrast, replicative DNA polymerases have been found to only follow the A-rule to bypass an AP site. This is due to their lack of the aforementioned structural cavity to accommodate the extrahelical AP site. Intriguingly, this kinetically observable intermediate in the A-rule pathway for Dpo4 was not able to be crystallized (17). This could be due to its weak thermostability or transient lifetime which makes it difficult to be captured by X-ray crystallography. Thus we conclude that AP bypass by Dpo4 follows both the A-rule and the lesion loop-out mechanism (Scheme 4.1) as opposed to strictly the latter as concluded previously (17).

AP lesion bypass via different pathways. Interestingly, the preference of these two pathways in Scheme 4.1 depended on the identity of the template base immediately 5′ to
the AP site (Table 4.9). When the 5’-base to the AP site was cytosine, the A-rule was preferred by 3.7-fold, indicating the mechanism of AP bypass was sequence dependent. Neither pathway was significantly favored when the 5’-base was adenine. Overall, our data in Table 4.9 indicated that Dpo4 and possibly all Y-family DNA polymerases use both the A-rule and the lesion loop-out mechanism to select an incoming nucleotide opposite an AP site.

After Dpo4 followed the lesion loop-out mechanism to incorporate dCTP into 21-mer/41AP to form 22-C-mer/41AP (Table 4.1), further elongation of this product could occur either retaining the looped-out AP site or after primer/template realignment (Scheme 4.1). The preferred nucleotide was expected to form a Watson-Crick basepair with the corresponding template base of each of these two intermediates, e.g. dGTP for the looped-out intermediate while dCTP for the realigned DNA substrate. Interestingly, such a realigned DNA substrate has been crystallized with Dpo4 and an incoming dCTP (see ternary structure Ab-3 (17)). From the measured catalytic efficiencies of the incorporation of these two nucleotides into 22-C-mer/41AP (Table 4.3), we calculated the efficiency ratio of 5.2 and a realignment percentage of 16.1% (Table 4.7). These values indicated that the looped-out branch pathway was preferred over the realignment branch pathway at this position. This preference was more significant for downstream nucleotide incorporations (Table 4.7), suggesting the looped-out AP lesion will remain extrahelical until the primer is fully extended. The A-rule intermediates, which possessed an intrahelical AP site, were also stabilized by downstream incorporation events (Table 4.5). On the basis of our kinetic data, we proposed several major pathways in Scheme 4.1 for the AP lesion bypass catalyzed by Dpo4. The pathways in Scheme 4.1 also apply to AP-1 and AP-8 although the pathway preference varied depending on the base 5’ to the AP site (Table 4.9). Notably, many other possible pathways were not included in Scheme 4.1 due
to their reduced kinetic partitioning. However, the totality of possible bypass pathways implied by the kinetic partitioning of each incorporation event strongly suggested that the AP lesion will cause numerous downstream mutations during primer elongation (see discussion below).

**Insignificant effect of an AP lesion on the affinity of the Dpo4•DNA binary complex.** A plausible explanation for the pausing observed in the vicinity of the AP lesion during processive synthesis by Dpo4, could involve a dramatically reduced affinity of Dpo4 for the AP-containing DNA substrate. Although an AP site does not affect the global B-form conformation of DNA (29,36) numerous studies have shown significant structural aberrations in the immediate vicinity of an AP lesion (29,31,36,37). The resulting bending and kinking of the duplex DNA and disruption of the hydrogen bonding network proximal to the lesion (38) may be significant enough to disrupt contacts within the Dpo4 binding pocket, potentially accounting for the observed pausing. However, the measured affinity for the Dpo4•DNA binary complex was only modestly affected (2- to 4-fold) at the two strong pause sites, while the $K_d$ values at the flanking nonpause/weak pause sites were indistinguishable from those of control DNA substrates (Table 4.2). Interestingly, our study concluded that the mechanism of bypass of an AP site by Dpo4 competed between incorporation of dATP (A-rule) and incorporation directed by the template base 5’ to the AP site (lesion loop-out mechanism). Notably, when bypass was directed by the latter mechanism, a looped-out AP site required accommodation in the Dpo4 active site. A series of ternary structures of Dpo4 showed that single-base bulges in the template can be inserted into the gap between the finger and little finger domain while extrahelical bases in the primer can fit in the DNA minor groove due to the unusually small finger and thumb domains (17). These structural observations are consistent with the
quantitative results from the gel mobility shift assay and help justify why Dpo4 binding to the AP site was not the cause of intermediate product accumulation seen in Figure 4.1.

**Kinetic mechanism for Dpo4 pausing.** Our systematic pre-steady kinetic analysis demonstrated that Dpo4 pausing was due to a significantly reduced catalytic efficiency ($k_p/K_d$) near the AP lesion. Comparisons of the $k_p/K_d$ for the two preferred incorporations opposite the AP site versus matched nucleotide incorporation into undamaged DNA (3) revealed decreases in $k_p/K_d$ of 2,560-fold (dATP, the A-rule) and 2,020-fold (dCTP, the lesion loop-out mechanism), indicating the extension of the DNA substrate 21-mer/41AP was slow. In contrast, the production of this substrate from 20-mer/41AP was as efficient as for undamaged DNA (Figure 4.7) thus resulting in 21-mer accumulation (Figure 4.1). Similarly, the elongation of 22-mer/41AP was difficult because even the most efficient nucleotide (dGTP) was incorporated into the most probable 22-mer/41AP substrate with 230-fold lower efficiency than into undamaged control DNA (Table 4.3, (3)), leading to strong accumulation of 22-mer in Figure 4.1. Following the two strong pause sites, the catalytic efficiencies of downstream incorporation events were also lowered, but gradually normalized, regardless of the bypass pathway (Figure 4.7 and 4.8).

The aforementioned reduction of nucleotide incorporation efficiency at the two strong pause sites was mainly due to significantly slower $k_p$ values, with respective decreases in the $k_p$ values via the A-rule and the lesion loop-out mechanism pathways of approximately 1,070-fold and 100-fold (Table 4.3) at the first pause site and 50-fold and 130-fold (Table 4.3) at the second pause site relative to undamaged DNA (3). Considering the two proposed mechanisms for AP site bypass in Scheme 4.1, the slow $k_p$ values observed opposite the lesion were justified. Incorporation via the A-rule most likely proceeded by means of nontemplated incorporation of dATP that was stabilized in
the ground-state by stacking interactions with the primer 3’-terminal base (adenine in our case). The lack of viable hydrogen bonding contacts between the incoming nucleotide (dATP) and this noncoding lesion precluded turnover rates on the same order as observed for matched nucleotide incorporation into undamaged DNA. In this respect, extension from this incorporation event (i.e. 22-A-mer/41AP) was also relatively slow (Table 4.3) because once incorporated, structure analysis shows the adenine at the primer terminus shifts toward the template strand and stacks with the preceding bases of both strands (17). This inward shift of adenine in the primer strand increases the distance between the 3’-OH and the α–phosphate of the next incoming nucleotide (dCTP in structure Ab-3 (17)) to 7 Å. This is beyond the 3.4 Å observed for an optimum catalytically active DNA polymerase ternary complex (39), therefore retarding catalysis. A similar rate reduction was observed for incorporation via the lesion loop-out mechanism which involved a shift of the downstream template strand into the active site generating an extrahelical AP site in which structural studies suggest the distance between the 3’-OH and the α–phosphate of the incoming dCTP is > 4 Å (17).

X-ray crystal structure studies reveal that Dpo4 can form different types of ternary structures when it bypasses an AP site (17), suggesting that damaged DNA may not be bound specifically at the Dpo4 active site. Further interrogation into this hypothesis revealed biphasic kinetics for the incorporation of both dATP and dCTP (21-mer/41AP DNA substrate) opposite the AP site. The biphasic kinetic trace (Figure 4.11) showed a fast phase of nucleotide incorporation with a low reaction amplitude followed by a slow phase of nucleotide incorporation with a significantly higher amplitude, thus suggesting the existence of two Dpo4-bound species: a productive ternary complex E•D_n^P•dNTP (D_n^P denotes productively bound DNA) that was competent for catalysis and a separate nonproductive ternary complex E•D_n^N•dNTP that was slowly converted to E•D_n^P•dNTP.
or simply dissociated (Scheme 4.2). In contrast, nucleotide incorporation into the nonpause site 20-mer/41AP substrate (data not shown) showed only a fast phase with a rate constant of 7.0 s\(^{-1}\), suggesting that the large slow phase and the significantly smaller ‘fast’ phase observed opposite the AP site contributed to strong Dpo4 pausing in Figure 4.1. Interestingly, since the assay in Figure 4.11 was performed in the presence of a large molar excess of non-radiolabeled DNA trap (>150-fold), the tightly bound ternary complexes E•D\(_n\)\(^{P}\)•dNTP and E•D\(_n\)\(^{N}\)•dNTP could only be converted from substrate D\(_n\) to product D\(_{n+1}\) in a single binding event. Thus, it was possible to calculate the individual contribution of each population of productive and nonproductively bound Dpo4 ternary complex to the overall reaction rate constant (\(k_p\)) measured under single turnover conditions. For example, when considering dATP incorporation, the summation of the contributions from the fast phase of nucleotide incorporation [(0.25 s\(^{-1}\) x (5.7% reaction amplitude)] and the slow phase [(0.00124 s\(^{-1}\) x (74% reaction amplitude)] gave an overall rate constant for the reaction of 0.0152 s\(^{-1}\), which was roughly equivalent to the \(k_p\) of 0.015 s\(^{-1}\) (Table 4.3). Similarly, the total contribution of the fast phase [(0.24 s\(^{-1}\) x (23.7% reaction amplitude)] and the slow phase [(0.0005 s\(^{-1}\) x (60.7% reaction amplitude)] generated a rate constant of 0.0571 s\(^{-1}\) which was similar to the \(k_p\) of 0.079 s\(^{-1}\) (Table 4.3) for dCTP incorporation. There was inherently more error in the latter calculation because the assay was performed at a subsaturating concentration of dCTP (1,200 µM) and as such the fast phase reaction amplitude was likely underestimated (\(K_d = 1,453 \mu M\) for dCTP incorporation into 21-mer/41AP). Since the \(k_p\) was the primary determinant in the overall contribution to the catalytic efficiency, and that the catalytic efficiency was responsible for Dpo4 pausing (see above discussion), these data suggested that since the rate constant for the fast phase (~0.25 s\(^{-1}\)) was similar for both dATP and dCTP incorporation, the relative population of productively bound complex (5.7% versus 23.7%) brought about the preference for dCTP over dATP incorporation opposite the AP
site (Table 4.9). This preference for productive ternary complex containing dCTP was most likely due to the fact that while extrahelical and intrahelical AP sites are in equilibrium in solution, a bulged AP site is more thermodynamically favorable especially when flanked by purines (40). In this example, the AP site was flanked by one pyrimidine and one purine (3’-PyXPu-5’) and in this sequence context, the lesion loop-out mechanism was preferred by ~2-fold. Interestingly, when we change the 5’ template base to adenine (3’-PyXPu-5’) and cytosine (3’-PyXPy-5’) we observed a preference for the A-rule (no AP collapse) only in the latter context (~4-fold), which contained two flanking pyrimidines which are known to poorly base stack. This suggested that the identity of the flanking template bases may play an important role in influencing the mechanism by which Dpo4 prefers to bypass an AP site. Although the second strong pause site was not assayed to determine if it too was characterized by biphasic kinetics, similar results were expected since this strong pause site was also characterized by a significant accumulation of intermediate product, again suggesting the existence of nonproductive complex.

Interestingly, the effect of the AP site on the ground-state binding affinity ($K_d$) of the two preferred dNTPs at these the first pause site (Table 4.3) only decreased 2.4- to 20.8-fold with respect to matched nucleotide incorporation (3). This decrease (on average, 11.6-fold) in $K_d$ corresponded to a free energy change ($\Delta\Delta G$) of 1.51 kcal/mol. While these $K_d$ differences were similar to the differences on undamaged DNA (3), they were also similar in magnitude to the decreases observed for preferred incorporation catalyzed by yeast Polη opposite a cis-syn T-T dimer (11) and an 8-oxodG (10) with respect to matched incorporation on undamaged DNA, where differences in $K_d$ values varied from 1.8- to 4.6-fold ($\Delta\Delta G = 0.41$ kcal/mol and 0.82 kcal/mol) opposite the 5’ thymine and 3’ thymine respectively and 2.1-fold ($\Delta\Delta G = 0.43$ kcal/mol) for incorporation opposite 8-oxodG. On the other hand, high fidelity replicative polymerases like human DNA
polymerase γ, have significantly higher values for ΔΔG (ΔΔG = 4.0 kcal/mol) due to the greater discrimination between matched and mismatched nucleotides in the ground-state attributed to more stringent fidelity checking mechanisms (41). These results suggested that the contribution from the active sites of Dpo4, yeast Polη, and most likely all Y-family polymerases, in selecting the incoming nucleotide in the ground-state when bypassing a DNA lesion is quite insignificant. The deficiency in specific interactions of Dpo4 with the incoming nucleotide in the ground-state suggested that nucleotide selection depended on its ability to stack with the base at the 3’ end of the primer (A-rule) or to form Watson-Crick basepairing with the downstream template base (lesion loop-out mechanism). The base stacking effect was justified based on a 20-fold decrease in the $K_d$ value from dPTP to dATP (Table 4.3) incorporation due to the superior aromatic π-stacking capabilities of dPTP relative to dATP, the best stacking natural nucleotide (26,42). Additionally, the stacking power of dPTP also facilitated catalysis by significantly increasing its $k_p$ value over those of dNTPs (Table 4.3).

*Strong pause sites are mutational hot spots.* Before Dpo4 encountered the AP site, our kinetic data with 20-mer/41AP (Table 4.10) suggested that the fidelity and likely the mechanism of nucleotide incorporation was equivalent to that reported for undamaged DNA (3,8) yet this mechanism vastly differed opposite the lesion (Scheme 4.2) and at several downstream bases. When opposite the AP site, our kinetic data (Table 4.3) predicted that Dpo4 can insert each of the four dNTPs although the incorporation percentages of dCTP (56.8%) and dATP (32.6%) were significantly higher due to the lesion loop-out mechanism and the A-rule, respectively. These data further suggested that bypass of an AP lesion by Dpo4 will cause either a ‘-1 frameshift’ or randomized substitutions. Moreover, our kinetic data suggested that substitution events occurred when Dpo4 extended these bypassed intermediates, especially 22-C-mer/41AP and 22-A-
mer/41AP in which two dNTPs were preferentially incorporated (Table 4.3). Thus, our pre-steady state kinetic analysis predicted that the two strong pause sites in Figure 4.1 were mutational hot spots and that the embedded AP lesion lowered downstream incorporation preference, but gradually normalized (Tables 4.5 and 4.6). If these kinetic predictions are confirmed in vivo, bypassing AP lesions by Dpo4 or any Y-family DNA polymerase member will be detrimental to the stability of the organism’s genome and thus would require tight cellular regulation to maintain the stable transmission of genetic information.
Figure 4.1. Running start nucleotide incorporation assay. A preincubated solution of Dpo4 (100 nM) and 17-mer/41-mer (100 nM) were mixed with all four dNTPs (200 µM each) for various reaction times before quenching with EDTA. The product lengths were labeled and the AP was designated. (A) Reaction with the 17-mer/41CTL substrate, (B) reaction with the 17-mer/41AP substrate.
Figure 4.2. Running start nucleotide incorporation assay. A preincubated solution of *S. solfataricus* replicative DNA polymerase PolB1 (100 nM) and 17-mer/41-mer (100 nM) were mixed with all four dNTPs (200 µM each) for various reaction times before quenching with EDTA. The product lengths were labeled and the AP was designated. (A) Reaction with the 17-mer/41CTL substrate, (B) reaction with the 17-mer/41AP substrate.
Figure 4.3. Measurement of the binding of Dpo4 to 21-mer/41AP. Reactions containing 100 nM 5’-[32P] 21-mer/41AP were incubated with increasing concentration of Dpo4 (35-425 nM) followed by native gel analysis to separate binary complex from unbound DNA substrate.
Figure 4.4. Measurement of the binding of Dpo4 to 21-mer/41AP. The solid line is a fit of the data (●) from Figure 4.3 to a quadratic equation giving a $K_d = 36 \pm 2$ nM.
Figure 4.5. Concentration dependence of dATP on nucleotide incorporation. A preincubated solution of Dpo4 (120 nM) and 5'-labeled 21-mer/41AP (30 nM) was rapidly mixed with increasing concentration of dATP (100 µM, ○; 300 µM, ●; 500 µM, ■; 800 µM, □; 1100 µM, ▲; 1400 µM, △; 1800 µM, ◆) for various reaction times. The solid lines are best fits to the single exponential equation.
Figure 4.6. Concentration dependence of dATP on nucleotide incorporation. The extracted observed rate constants from the above data fitting were plotted as a function of dATP concentration and fit to eq 2 to obtain a $k_p$ of $0.015 \pm 0.002$ s$^{-1}$ and a $K_d$ of $490 \pm 174$ µM.
Figure 4.7. Quantitative effect of the AP site on nucleotide incorporation. To determine the kinetic influence on AP site bypass, the catalytic efficiency for the favored nucleotide incorporation, \((k_p/K_d)_{\text{AP}}\), for each AP-containing DNA substrate was divided by the catalytic efficiency for the corresponding correct incorporation into each control substrate, \((k_p/K_d)_{\text{CTL}}\), and was then plotted for each substrate.
Figure 4.8. Quantitative effect of the AP site on nucleotide incorporation. To determine the kinetic influence on AP site bypass, the catalytic efficiency for the favored nucleotide incorporation, $(k_p/K_d)_{\text{AP}}$, for each AP-containing DNA substrate was divided by the catalytic efficiency for the corresponding correct incorporation into each control substrate, $(k_p/K_d)_{\text{CTL}}$, and was then plotted for each substrate.
Figure 4.9. Concentration dependence of observed dPTP incorporation rate constant. The chemical structure of dPTP.
Figure 4.10. Concentration dependence of observed dPTP incorporation rate constant. The extracted observed rate constants were plotted as a function of dPTP concentration and fit to eq 2 (Materials and methods) to obtain a $k_p$ of $0.26 \pm 0.02$ s$^{-1}$ and a $K_d$ of $24 \pm 4$ µM.
Figure 4.11. Incorporation of dATP into 21-mer/41AP catalyzed by Dpo4 in the presence of a DNA trap showed biphasic kinetics. The product concentration versus reaction time were fit to eq 4 which gave a fast phase reaction amplitude and rate constant of $1.7 \pm 0.1$ nM and $0.25 \pm 0.10$ s$^{-1}$ respectively while the slow phase was characterized by a rate constant of $0.00124 \pm 0.00009$ s$^{-1}$ and an amplitude of $22.2 \pm 0.8$ nM.
### Table 4.1. DNA primers and templates.

<table>
<thead>
<tr>
<th>Primers</th>
<th>17-mer</th>
<th>5′-CGCAGCCGTCCAACCAA-3′</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTC-3′</td>
</tr>
<tr>
<td></td>
<td>21-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA-3′</td>
</tr>
<tr>
<td></td>
<td>22-A-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{A} )-3′</td>
</tr>
<tr>
<td></td>
<td>23-A-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{AC} )-3′</td>
</tr>
<tr>
<td></td>
<td>24-A-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{ACG} )-3′</td>
</tr>
<tr>
<td></td>
<td>25-A-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{ACGT} )-3′</td>
</tr>
<tr>
<td></td>
<td>26-A-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{ACGTC} )-3′</td>
</tr>
<tr>
<td></td>
<td>27-A-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{ACGTCG} )-3′</td>
</tr>
<tr>
<td></td>
<td>22-C-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{C} )-3′</td>
</tr>
<tr>
<td></td>
<td>23-C-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{CG} )-3′</td>
</tr>
<tr>
<td></td>
<td>24-C-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{CGT} )-3′</td>
</tr>
<tr>
<td></td>
<td>25-C-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{CGTC} )-3′</td>
</tr>
<tr>
<td></td>
<td>26-C-mer</td>
<td>5′-CGCAGCCGTCCAACCAACTCA ( \text{CGTCG} )-3′</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Templates*</th>
<th>41AP</th>
<th>3′-GCGTCGGCAGGTTGGTGAGTXGCAGCTAGGTTACGGCAGG-5′</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>41CTL</td>
<td>3′-GCGTCGGCAGGTTGGTGAGTNGCAGCTAGGTTACGGCAGG-5′</td>
</tr>
<tr>
<td></td>
<td>22AP</td>
<td>3′-GCGTCGGCAGGTTGGTGAGTX-5′</td>
</tr>
</tbody>
</table>

* \( \text{X} \) designates an AP site. \( \text{N} \) designates a template T or G corresponding to primers containing A or C, respectively, located twenty-two nucleotides from the 5′ end of the primer.
Table 4.2. Binding affinity of Dpo4 to control and damaged DNA substrates at 23 °C.

<table>
<thead>
<tr>
<th>Damaged substrate</th>
<th>$K_d^{DNA}$ (nM)</th>
<th>Control substrate</th>
<th>$K_d^{DNA}$ (nM)</th>
<th>Affinity ratio$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-mer/41AP</td>
<td>6 ± 1</td>
<td>20-mer/41CTL</td>
<td>9 ± 1</td>
<td>0.6-fold</td>
</tr>
<tr>
<td>21-mer/41AP$^b$</td>
<td>36 ± 2</td>
<td>21-mer/41CTL</td>
<td>10 ± 2</td>
<td>3.6-fold</td>
</tr>
<tr>
<td>22-A-mer/41AP$^b$</td>
<td>17 ± 2</td>
<td>22-A-mer/41CTL</td>
<td>6 ± 2</td>
<td>2.6-fold</td>
</tr>
<tr>
<td>23-A-mer/41AP</td>
<td>8 ± 1</td>
<td>23-A-mer/41CTL</td>
<td>10 ± 1</td>
<td>0.8-fold</td>
</tr>
<tr>
<td>22-C-mer/41AP$^b$</td>
<td>24 ± 1</td>
<td>22-C-mer/41CTL</td>
<td>11 ± 2</td>
<td>2.2-fold</td>
</tr>
<tr>
<td>23-C-mer/41AP</td>
<td>14 ± 1</td>
<td>23-C-mer/41CTL</td>
<td>8 ± 2</td>
<td>1.8-fold</td>
</tr>
</tbody>
</table>

$^a$Calculated as $(K_d^{DNA})_{damaged}/(K_d^{DNA})_{control}$.

$^b$DNA substrates at the pause sites.
Table 4.3. Kinetic parameters for each nucleotide incorporation into damaged DNA at the two pause sites.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$</th>
<th>$k_p$</th>
<th>$k_p/K_d$</th>
<th>Incorporation%a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(µM)</td>
<td>(s⁻¹)</td>
<td>(µM⁻¹ s⁻¹)</td>
<td></td>
</tr>
<tr>
<td>The 1st pause site (Damaged 21-mer/41AP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>490±174</td>
<td>0.015±0.002</td>
<td>3.1x10⁻⁵</td>
<td>32.6%</td>
</tr>
<tr>
<td>dCTP</td>
<td>1453±287</td>
<td>0.079±0.009</td>
<td>5.4x10⁻⁵</td>
<td>56.8%</td>
</tr>
<tr>
<td>dGTP</td>
<td>780±165</td>
<td>0.0010±0.0001</td>
<td>1.4x10⁻⁶</td>
<td>1.5%</td>
</tr>
<tr>
<td>dTTP</td>
<td>1953±773</td>
<td>0.017±0.004</td>
<td>8.7x10⁻⁶</td>
<td>9.1%</td>
</tr>
<tr>
<td>dPTP</td>
<td>24±4</td>
<td>0.26±0.02</td>
<td>1.1x10⁻²</td>
<td>1.0%</td>
</tr>
<tr>
<td>The 2nd pause site (Damaged 22-C-mer/41AP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>303±103</td>
<td>0.075±0.008</td>
<td>2.4x10⁻⁴</td>
<td>82.6%</td>
</tr>
<tr>
<td>dATP</td>
<td>146±8</td>
<td>0.00020±0.00001</td>
<td>1.6x10⁻⁶</td>
<td>0.6%</td>
</tr>
<tr>
<td>dCTP</td>
<td>646±220</td>
<td>0.029±0.004</td>
<td>4.6x10⁻⁵</td>
<td>15.8%</td>
</tr>
<tr>
<td>dTTP</td>
<td>1253±205</td>
<td>0.0038±0.0004</td>
<td>3.0x10⁻⁶</td>
<td>1.0%</td>
</tr>
<tr>
<td>The 2nd pause site (Damaged 22-A-mer/41AP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>567±85</td>
<td>0.157±0.009</td>
<td>2.8x10⁻⁴</td>
<td>95.8%</td>
</tr>
<tr>
<td>dATP</td>
<td>316±38</td>
<td>0.001±0.0001</td>
<td>1.3x10⁻⁶</td>
<td>0.4%</td>
</tr>
<tr>
<td>dGTP</td>
<td>437±102</td>
<td>0.0043±0.0004</td>
<td>9.9x10⁻⁶</td>
<td>3.4%</td>
</tr>
<tr>
<td>dTTP</td>
<td>1246±319</td>
<td>0.0014±0.0002</td>
<td>1.1x10⁻⁶</td>
<td>0.4%</td>
</tr>
</tbody>
</table>

aIncorporation% = ($k_p/K_d$)_{dNTP} / Σ($k_p/K_d$)_{dNTP}
Table 4.4. Comparison of the substrate specificity for nucleotide incorporation into 21-mer/22AP.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$  (µM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (µM$^{-1}$s$^{-1}$)</th>
<th>Efficiency ratio$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dATP</td>
<td>634 ± 175</td>
<td>0.016 ± 0.002</td>
<td>$2.5 \times 10^{-5}$</td>
<td>1</td>
</tr>
<tr>
<td>dCTP</td>
<td>723 ± 141</td>
<td>0.0006 ± 0.0001</td>
<td>$7.7 \times 10^{-7}$</td>
<td>0.03</td>
</tr>
<tr>
<td>dGTP</td>
<td>906 ± 174</td>
<td>0.0006 ± 0.0001</td>
<td>$6.2 \times 10^{-7}$</td>
<td>0.02</td>
</tr>
<tr>
<td>dTTP</td>
<td>699 ± 94</td>
<td>0.0031 ± 0.0002</td>
<td>$4.4 \times 10^{-6}$</td>
<td>0.18</td>
</tr>
</tbody>
</table>

$^a$Efficiency ratio = ($k_p/K_d$)$_{dNTP}/(k_p/K_d)_{dATP}$
Table 4.5. Kinetic parameters for each nucleotide incorporation into damaged DNA downstream of the pause sites.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (µM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (µM$^{-1}$s$^{-1}$)</th>
<th>Incorporation%a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Damaged 23-A-mer/41AP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>368 ± 104</td>
<td>1.2 ± 0.1</td>
<td>3.4 x 10$^3$</td>
<td>98.6%</td>
</tr>
<tr>
<td>dATP</td>
<td>827 ± 275</td>
<td>0.007 ± 0.001</td>
<td>8.1 x 10$^6$</td>
<td>0.2%</td>
</tr>
<tr>
<td>dCTP</td>
<td>173 ± 36</td>
<td>0.0038 ± 0.0002</td>
<td>2.2 x 10$^5$</td>
<td>0.6%</td>
</tr>
<tr>
<td>dTTP</td>
<td>1454 ± 566</td>
<td>0.026 ± 0.006</td>
<td>1.8 x 10$^5$</td>
<td>0.5%</td>
</tr>
<tr>
<td><strong>Damaged 24-A-mer/41AP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>447 ± 119</td>
<td>2.2 ± 0.2</td>
<td>4.8 x 10$^3$</td>
<td>99.7%</td>
</tr>
<tr>
<td>dATP</td>
<td>496 ± 58</td>
<td>0.00097 ± 0.00004</td>
<td>2.0 x 10$^6$</td>
<td>0.1%</td>
</tr>
<tr>
<td>dCTP</td>
<td>612 ± 196</td>
<td>0.008 ± 0.001</td>
<td>1.4 x 10$^5$</td>
<td>0.3%</td>
</tr>
<tr>
<td>dGTP</td>
<td>334 ± 37</td>
<td>0.00039 ± 0.00001</td>
<td>1.2 x 10$^6$</td>
<td>0.0%</td>
</tr>
<tr>
<td><strong>Damaged 26-A-mer/41AP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>412 ± 48</td>
<td>2.8 ± 0.1</td>
<td>6.7 x 10$^3$</td>
<td>99.9%</td>
</tr>
<tr>
<td>dATP</td>
<td>1499 ± 377</td>
<td>0.004 ± 0.001</td>
<td>3.0 x 10$^6$</td>
<td>0.0%</td>
</tr>
<tr>
<td>dCTP</td>
<td>1436 ± 151</td>
<td>0.0016 ± 0.0001</td>
<td>1.1 x 10$^6$</td>
<td>0.0%</td>
</tr>
<tr>
<td>dTTP</td>
<td>1303 ± 489</td>
<td>0.005 ± 0.001</td>
<td>3.8 x 10$^6$</td>
<td>0.1%</td>
</tr>
<tr>
<td><strong>Damaged 27-A-mer/41AP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>260 ± 45</td>
<td>8.5 ± 0.5</td>
<td>3.3 x 10$^2$</td>
<td>99.9%</td>
</tr>
<tr>
<td>dCTP</td>
<td>524 ± 81</td>
<td>0.0018 ± 0.0001</td>
<td>3.4 x 10$^6$</td>
<td>0.0%</td>
</tr>
<tr>
<td>dGTP</td>
<td>578 ± 110</td>
<td>0.0083 ± 0.0006</td>
<td>1.4 x 10$^5$</td>
<td>0.1%</td>
</tr>
<tr>
<td>dTTP</td>
<td>329 ± 34</td>
<td>0.0029 ± 0.0001</td>
<td>8.9 x 10$^6$</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

aIncorporation$% = (k_p/K_d)_{dNTP}/\Sigma(k_p/K_d)_{dNTP}$
<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (µM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (µM$^{-1}$s$^{-1}$)</th>
<th>Incorporation%$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Damaged 23-C-mer/41AP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>267 ± 51</td>
<td>0.68 ± 0.05</td>
<td>2.6 x 10$^3$</td>
<td>98.5%</td>
</tr>
<tr>
<td>dATP</td>
<td>253 ± 23</td>
<td>0.0029 ± 0.0001</td>
<td>1.1 x 10$^5$</td>
<td>0.4%</td>
</tr>
<tr>
<td>dCTP</td>
<td>398 ± 58</td>
<td>0.0077 ± 0.0004</td>
<td>1.9 x 10$^5$</td>
<td>0.8%</td>
</tr>
<tr>
<td>dGTP</td>
<td>575 ± 145</td>
<td>0.0047 ± 0.0005</td>
<td>8.2 x 10$^6$</td>
<td>0.3%</td>
</tr>
<tr>
<td><strong>Damaged 24-C-mer/41AP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>156 ± 44</td>
<td>1.6 ± 0.1</td>
<td>1.0 x 10$^2$</td>
<td>99.9%</td>
</tr>
<tr>
<td>dATP</td>
<td>472 ± 47</td>
<td>0.00056 ± 0.00002</td>
<td>1.2 x 10$^6$</td>
<td>0.0%</td>
</tr>
<tr>
<td>dGTP</td>
<td>531 ± 198</td>
<td>0.0006 ± 0.0001</td>
<td>1.2 x 10$^6$</td>
<td>0.0%</td>
</tr>
<tr>
<td>dTTP</td>
<td>1076 ± 234</td>
<td>0.008 ± 0.001</td>
<td>7.7 x 10$^6$</td>
<td>0.1%</td>
</tr>
<tr>
<td><strong>Damaged 25-C-mer/41AP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>232 ± 93</td>
<td>2.9 ± 0.4</td>
<td>1.3 x 10$^2$</td>
<td>99.9%</td>
</tr>
<tr>
<td>dATP</td>
<td>339 ± 72</td>
<td>0.0024 ± 0.0002</td>
<td>7.2 x 10$^6$</td>
<td>0.1%</td>
</tr>
<tr>
<td>dCTP</td>
<td>1087 ± 141</td>
<td>0.0014 ± 0.0001</td>
<td>1.3 x 10$^6$</td>
<td>0.0%</td>
</tr>
<tr>
<td>dTTP</td>
<td>608 ± 126</td>
<td>0.0018 ± 0.0001</td>
<td>3.0 x 10$^6$</td>
<td>0.0%</td>
</tr>
<tr>
<td><strong>Damaged 26-C-mer/41AP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>129 ± 41</td>
<td>16.3 ± 1.5</td>
<td>1.3 x 10$^3$</td>
<td>99.8%</td>
</tr>
<tr>
<td>dCTP</td>
<td>607 ± 128</td>
<td>0.025 ± 0.002</td>
<td>4.2 x 10$^5$</td>
<td>0.0%</td>
</tr>
<tr>
<td>dGTP</td>
<td>817 ± 184</td>
<td>0.15 ± 0.02</td>
<td>1.9 x 10$^4$</td>
<td>0.1%</td>
</tr>
<tr>
<td>dTTP</td>
<td>786 ± 125</td>
<td>0.052 ± 0.004</td>
<td>6.6 x 10$^5$</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

$^a$Incorporation% = ($k_p/K_d$)$_{dNTP}$/$\Sigma(k_p/K_d)_{dNTP}$

**Table 4.6.** Kinetic parameters for each nucleotide incorporation into damaged DNA downstream of the pause sites.
<table>
<thead>
<tr>
<th>Looped-out</th>
<th>Realignment</th>
<th>Efficiency ratio</th>
<th>Realignment%$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dGTP 5′-CAC 3′-GTGCAAGCT-5′ ◀ ◀ 3′-GTGCAAGCT-5′</td>
<td>dCTP 5′-CAC 3′-GTGCAAGCT-5′</td>
<td>$\frac{(k_f/K_d)<em>{dGTP}}{(k_f/K_d)</em>{dCTP}} = 5.2$</td>
<td>16.1%</td>
</tr>
<tr>
<td>dTTP 5′-CACG 3′-GTGCAAGCT-5′</td>
<td>dGTP 5′-CACG 3′-GTGCAAGCT-5′</td>
<td>$\frac{(k_f/K_d)<em>{dTTP}}{(k_f/K_d)</em>{dGTP}} = 125$</td>
<td>0.8%</td>
</tr>
<tr>
<td>dCTP 5′-CAGT 3′-GTGCAAGCT-5′ ◀ 3′-GTGCAAGCT-5′</td>
<td>dTTP 5′-CAGT 3′-GTGCAAGCT-5′</td>
<td>$\frac{(k_f/K_d)<em>{dCCTP}}{(k_f/K_d)</em>{dTTP}} = 1250$</td>
<td>0.08%</td>
</tr>
<tr>
<td>dGTP 5′-CAGTC 3′-GTGCAAGCT-5′ ◀ 3′-GTGCAAGCT-5′</td>
<td>dCTP 5′-CAGTC 3′-GTGCAAGCT-5′</td>
<td>$\frac{(k_f/K_d)<em>{dGTP}}{(k_f/K_d)</em>{dCTP}} = 1860$</td>
<td>0.05%</td>
</tr>
</tbody>
</table>

$^a$Realignment% = [1/(1 + efficiency ratio)]x100%

Table 4.7. Realignment possibilities of the looped out DNA intermediates.
Table 4.8. Kinetic parameters for each nucleotide incorporation into damaged DNA substrates AP-1 and AP-8.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (µM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (µM$^{-1}$s$^{-1}$)</th>
<th>Incorporation%$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP-1$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>567 ± 201</td>
<td>0.044 ± 0.006</td>
<td>7.7 x 10$^{-5}$</td>
<td>48.0%</td>
</tr>
<tr>
<td>dCTP</td>
<td>1266 ± 155</td>
<td>0.0058 ± 0.0004</td>
<td>4.6 x 10$^{-6}$</td>
<td>2.9%</td>
</tr>
<tr>
<td>dGTP</td>
<td>680 ± 182</td>
<td>0.0034 ± 0.0004</td>
<td>5.0 x 10$^{-6}$</td>
<td>3.1%</td>
</tr>
<tr>
<td>dTTP</td>
<td>583 ± 132</td>
<td>0.043 ± 0.004</td>
<td>7.4 x 10$^{-5}$</td>
<td>46.0%</td>
</tr>
<tr>
<td>AP-2$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>623 ± 96</td>
<td>0.045 ± 0.003</td>
<td>7.3 x 10$^{-5}$</td>
<td>63.4%</td>
</tr>
<tr>
<td>dCTP</td>
<td>1373 ± 304</td>
<td>0.0043 ± 0.0005</td>
<td>3.2 x 10$^{-6}$</td>
<td>2.8%</td>
</tr>
<tr>
<td>dGTP</td>
<td>885 ± 129</td>
<td>0.018 ± 0.001</td>
<td>2.0 x 10$^{-5}$</td>
<td>17.3%</td>
</tr>
<tr>
<td>dTTP</td>
<td>853 ± 120</td>
<td>0.016 ± 0.001</td>
<td>1.9 x 10$^{-5}$</td>
<td>16.5%</td>
</tr>
</tbody>
</table>

$^a$Incorporation% = ($k_p/K_d$)$_{dNTP}$/Σ($k_p/K_d$)$_{dNTP}$

$^b$AP-1 and AP-8 are both modified 21-mer/41-AP substrates possessing a template adenine or cytosine, respectively, 5′ to the AP site

Table 4.8. Kinetic parameters for each nucleotide incorporation into damaged DNA substrates AP-1 and AP-8.
Table 4.9. Preference between the A-rule and the lesion loop-out mechanism for bypass of an AP site.

<table>
<thead>
<tr>
<th>DNA</th>
<th>Dominant mechanism</th>
<th>Magnitude of preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>21-mer/41AP</td>
<td>5’rule</td>
<td>1.7-fold</td>
</tr>
<tr>
<td>AP-1</td>
<td>A-rule</td>
<td>1.1-fold</td>
</tr>
<tr>
<td>AP-8</td>
<td>A-rule</td>
<td>3.7-fold</td>
</tr>
</tbody>
</table>

*X* designates the AP site in the following DNA substrate:
5’-CGCAGCCGTCCAACCAACTCA-3’
3’-GCGTCGGCAGGTTGGTTGAGTXYCAGCTAGGTACGCAGG-5’

*Y* denotes base G in 21-mer/41AP (see Table 4.1), A in AP-1, and C in AP-8.
Table 4.10. Kinetic parameters for each nucleotide incorporation into damaged DNA substrates 20-mer/41AP.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (µM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (µM$^{-1}$s$^{-1}$)</th>
<th>Fidelity$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-mer/41AP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>283 ± 73</td>
<td>7.7 ± 0.6</td>
<td>2.7 x 10$^2$</td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>1547 ± 218</td>
<td>0.014 ± 0.001</td>
<td>9.0 x 10$^{-6}$</td>
<td>3.3 x 10$^{-4}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>658 ± 69</td>
<td>0.101 ± 0.005</td>
<td>1.5 x 10$^{-4}$</td>
<td>5.6 x 10$^{-3}$</td>
</tr>
<tr>
<td>dTTP</td>
<td>1796 ± 999</td>
<td>0.10 ± 0.03</td>
<td>5.6 x 10$^{-5}$</td>
<td>2.0 x 10$^{-3}$</td>
</tr>
</tbody>
</table>

$^a$Fidelity = $(k_p/K_d)_{incorrect}/[(k_p/K_d)_{incorrect} + (k_p/K_d)_{correct}]$
Scheme 4.1. Proposed mechanisms to bypass and AP lesion.
Scheme 4.2. Proposed minimal kinetic mechanism of AP site bypass catalyzed by Dpo4.
4.6 References


CHAPTER 5

SLOPPY BYPASS OF AN ABASIC SITE LESION CATALYZED BY
*SULFOLOBUS SOLFATARICUS* DNA POLYMERASE IV

5.1 Introduction

Cellular DNA is continually damaged by agents both endogenous and exogenous to an organism. This damage structurally modifies DNA, arresting replication by stalling the replicative DNA polymerase. The inability to copy the genome leads to cell-cycle stalling and possibly cell death. Although a large portion of DNA lesions is restored by cellular repair pathways, a biologically significant amount of damage persists. Fortunately, organisms have evolved specialized enzymes known as the Y-family DNA polymerases which are capable of replicating damaged DNA in either an error-free or error-prone fashion (1,2), thus rescuing a stalled replisome.

Although the Y-family DNA polymerases collectively synthesize one to a few nucleotides per binding event, they all lack proofreading exonuclease activity and generate up to 100,000-fold more errors on undamaged DNA when compared to the astonishingly accurate replicative DNA polymerases (1,3,4). This high error rate is due to the loose active sites in the structures of the eukaryotic polymerases: human polymerase ι (5), yeast polymerase eta (6), human polymerase κ (Polκ) (7), and yeast Rev1 (8), as well as the thermophilic archaeal *Sulfolobus solfataricus* DNA polymerase IV (Dpo4) (9) and
S. acidocaldarius Dbh (10). These more permissive active sites impart unique functional attributes to the Y-family, including the ability to replicate through aberrant and bulky DNA lesions which is conferred by the limited contacts between the Y-family polymerase active site residues and DNA.

Among DNA damage, apurinic/apyrimidinic (AP) sites are one of the most common lesions in mammalian cells (11) with a steady-state frequency approaching 50,000 sites per genome (12). These non-coding lesions are generated enzymatically, via spontaneous hydrolysis, or by more complex mechanisms involving free radicals or alkylating agents (13,14). Since S. solfataricus grows at 75 °C to 85 °C, it is reasonable to expect more genomic AP sites compared to the relative amount present in humans and other mesophilic systems. While replicative polymerases like S. solfataricus PolB1 lack the ability to bypass AP sites (Chapter 4), they can inefficiently incorporate dATP opposite the lesion in a phenomenon known as the ‘A-rule’ (15). Thus, due to the large number of AP sites and the fact that Dpo4 is the only known Y-family polymerase in S. solfataricus, Dpo4 is an excellent model for mechanistic and structural studies of AP bypass. In Chapter 4, we showed that during efficient bypass, Dpo4 paused when incorporating nucleotides directly opposite and one position downstream from an AP lesion due to a decrease in catalytic efficiency. In addition, incorporation opposite the lesion partitioned primarily between two distinct pathways, one involving the ‘A-rule’ (Scheme 5.1, A) and the second involving the novel lesion loop-out mechanism (Scheme 5.1, B). This latter mechanism involved incorporation directed by the template base 5’ to the AP lesion via an extrahelical AP site that was extended in-frame without primer/template realignment generating a ‘-1 frameshift’. However, these kinetic predictions need to be verified by actual DNA sequence data. In addition, kinetic analysis cannot capture infrequent mutation events and thus cannot provide a complete picture about the mutagenic
spectrum, especially downstream from the abasic lesion. In this chapter, we determine the precise sequences of short DNA products synthesized by Dpo4 during AP lesion bypass in vitro, using a novel short oligonucleotide sequencing assay. These sequences will reveal the mutagenic impact of AP lesion bypass by the Y-family.

5.2 Materials and methods

Materials. Human AP endonuclease and Taq DNA polymerase were purchased from Trevigen and Invitrogen, respectively. Dpo4 was expressed and purified as described previously (16). The oligodeoxynucleotides listed in Table 5.1 were purchased from Integrated DNA Technologies and were purified, labeled and annealed as described previously (17).

Reaction Buffer R. The optimized buffer contained 50 mM HEPES (pH 7.5 at 37 °C), 5 mM MgCl2, 50 mM NaCl, 5 mM DTT, 10% glycerol, 0.1 mM EDTA, and 0.1 mg/ml BSA (16).

Short oligonucleotide sequencing assay. Dpo4 (120 nM) was incubated with 17-mer/41-AP (30 nM, Table 5.1) and mixed with four dNTPs (200 µM each) for 3 hours which was sufficient to generate full-length ‘AP bypass product’. Dpo4 was then extracted three times with buffer saturated phenol-chloroform-isoamyl alcohol (25:24:1) and the remaining DNA substrates were incubated with human AP endonuclease (50 units, Trevigen) to cleave the DNA template 41AP. The full-length ‘AP bypass product’ synthesized by Dpo4 was then purified from the pool of DNA complexes via 20%
denaturing polyacrylamide gel electrophoresis to separate ‘AP bypass product’ from cleaved template 41AP, followed by gel extraction and reverse-phase C18 column chromatography. The resulting purified ‘AP bypass product’ was then PCR amplified using the following primers, 17mer_AP_sequencing_for: 5’-CGCAGCCGTCCAACCAA-3’ and 15mer_AP_sequencing_rev: 5’-GGACGGCATTTGATC-3’, which were designed to provide sequence information for four nucleotides both upstream and downstream from the AP site in addition to events that score an incorporation directly opposite the lesion. Based on our results, the spectrum of mutations were essentially always resolved within this sequence window, however, there remains a small possibility that some bypass products contained additional mutations outside this window. The amplified ‘AP bypass product’ was then ligated into a pCRII-TOPO vector using a TOPO TA cloning kit (Invitrogen) followed by transformation into TOP10 *E. coli* strain (Invitrogen). Plasmid was isolated from fifty-five independent colonies using a miniprep kit (Promega) and were then sequenced using an automated 3730 DNA analyzer (Applied Biosystems, Inc.). The method was summarized in Scheme 5.2. Controls for human AP endonuclease show its ability to completely digest DNA containing either intrahelical or extrahelical AP sites, while showing no degradation of an undamaged DNA control (data not shown).

*BmgB I digestion assay.* A preincubated solution of Dpo4 (120 nM) and 5’-radiolabeled 17-mer/41AP (30 nM, Table 5.1) was mixed with four dNTPs (200 µM each) to generate full length ‘AP bypass product’ at 37 °C. The full-length ‘AP bypass product’ was purified from Dpo4 and the template 41AP via phenol-chloroform extraction, human AP endonuclease treatment, denaturing polyacrylamide gel electrophoresis, and C18 column chromatography as described above. The pool of purified ‘AP bypass product’ was annealed to 40Te (Table 5.1) in molar excess containing half of the BmgB I restriction
site, in 1 x NEB Buffer and then overdigested with BmgB I (20 units, NEB) for 3 hours at
37 °C. Digestion reactions were then heated to 65 °C for 30 minutes to thermally
denature BmgB I and subsequently run on a 20% native polyacrylamide gel for 2 hours to
resolve digested product from full-length duplex DNA. The percent digestion was
quantitated through PhosphorImager. Reactions at 75 °C were performed in Buffer R
with potassium phosphate instead of HEPES.

5.3 Results and Discussion

Our previous kinetic results (Chapter 4) reveal numerous branched pathways used by
Dpo4 to bypass an AP lesion. While elucidating the kinetics of each branched pathway
would be extremely tedious, various aspects of these kinetic assays require careful
consideration. Although these pre-steady state kinetic methods are mechanistically
informative and provide a relative measure of polymerization fidelity, they cannot
generate DNA sequence information or the mutation profile. The most common kinetic
assay for fidelity determination involves measuring and comparing the incorporation
preference of each individual deoxynucleoside triphosphate (dNTP) into a DNA substrate
via single nucleotide incorporation assays under either steady-state or pre-steady state
reaction conditions. Since the selection of dNTPs at the polymerase active site for
incorporation opposite a specific canonical template base forms a correct or an incorrect
Watson-Crick basepair, polymerase fidelity can be calculated under the assumption that
there will be exactly one nucleotide incorporated opposite each template base. While
almost exclusively true for a replicative polymerase which replicates undamaged DNA
with high fidelity, bypass of a DNA lesion by Dpo4, a low fidelity enzyme, has been
shown by us here and others previously (18-22), to generate a variety of deletion and addition mutations in the vicinity of the lesion, indicating an equilibrium between several conformationally distinct DNA species exploited for bypass. This equilibrium is particularly relevant at structurally flexible AP sites (23,24) and thus invalidates the assumption of ‘one dNTP for each template base’ for lesion bypass. As such, the definition of fidelity, while suitable for studies with undamaged DNA, cannot be strictly applied to these incorporation events. Furthermore, these single nucleotide incorporation assays are subject to pyrophosphorolysis, the reverse of forward synthesis, if the incoming dNTP does not form a correct basepair with the downstream template base after the first incorporation event. Dpo4 has been shown to have somewhat robust pyrophosphorolysis activity which is notably attenuated when an incorrect incorporation event can be extended by correct nucleotide incorporation (25). Thus the kinetic results of lesion bypass from single nucleotide incorporation assays may not accurately reflect all actual bypass events, especially those through non-canonical pathways (see below), and requires cautious interpretation. To extend our previous kinetic studies reported in Chapter 4 and provide a complete mutagenic spectrum, we designed an assay for sequencing short DNA products synthesized by Dpo4 during AP lesion bypass to elucidate the entire mutation profile.

Development of a novel assay for sequencing short DNA products. Traditional strategies for sequencing short DNA products include the Maxam-Gilbert method and the Sanger method. However both of these methods are limited by the inability to resolve sequence information from a DNA mixture of variable lengths and sequences (20). Since the kinetic results in Chapter 4 indicated Dpo4 generated a population of bypass products, these techniques could not be used. Although the LC-MS/MS technique used by Zang and coworkers is fundamentally better than the aforementioned methods, it still only
provides a semi-quantitative analysis of a population of DNA substrates and disregards the sequence contexts of minor species (20). In addition, the modified reversion assay used by Kokoska et al. (18), while elegant, only detects roughly 60% of single base errors made during in vitro lesion bypass. This is due to the lack of a completely ligated plasmid prior to transformation into E. coli. In addition, the assay itself is less likely to observe minor species since it lacks DNA product amplification before plasmid assembly. Due to the shortcomings of these assays, we developed a short oligonucleotide sequencing assay (SOSA) integrating PCR amplification and an efficient cloning technique to unambiguously determine the sequence of a population of short DNA bypass products synthesized by Dpo4 in the exact sequence context used in Chapter 4.

**Analysis of DNA products using SOSA at 37 °C.** We exploited this novel assay to determine the mutation profile for the bypass of an AP site by Dpo4 using a 17-mer/41AP (Table 5.1) which contained a template guanine 5’ to the lesion. Ultimately, the method (Scheme 5.2) was designed to isolate and ligate the ‘AP bypass products’ synthesized by Dpo4 into a vector for automated sequence analysis. Analysis of the DNA sequences from fifty-five independent colonies is summarized in Figure 5.1. The results of the assay performed at 37 °C revealed that upstream from the AP site all incorporation events catalyzed by Dpo4 was sequence dependent and all incorporation events (220 nucleotides) were replicated error-free and in-frame. Strikingly, once Dpo4 encountered the lesion, a spectrum of mutagenic bypass events was generated. Interestingly 13% of the time (7/55 colonies) Dpo4 synthesized a full-length (41-mer) ‘AP bypass product’ lacking any downstream mutations with dAMP located opposite the AP site (Figure 5.1). Other ‘AP bypass products’ contained only single/double base substitutions (9%, 5/55 colonies), only single/multiple base deletions (56%, 31/55 colonies), only base additions (2%, 1/55 colonies), or complex transactions involving a combination of deletion(s),
addition(s), and substitution(s) (13%, 7/55 colonies). The remaining 7% (4/55 colonies) of ‘AP bypass products’ were full-length and in-frame, but had dCMP, dGMP, or dTMP inserted opposite the AP site (Figure 5.1). Although our results demonstrated the ability to incorporate all nucleotides directly opposite the lesion, 82% of the time (45/55 colonies) Dpo4 incorporated either dAMP (via A-rule) or dCMP (mostly via the lesion loop-out mechanism) coinciding with our previous kinetic results (Chapter 4) showing that 90% of AP lesion bypass events occurred via these two pathways. Yet, this initial partitioning opposite the lesion did not represent the full extent of mutagenic bypass of the AP bypass. Our results showed twenty-one unique sequence contexts revealing that unregulated bypass of the lesion by Dpo4 was extremely sloppy.

The most logical mechanisms for each of these twenty-one DNA sequences were diagrammed in Schemes 5.3 to 5.6 and were based on previous structural studies of Dpo4. Roughly 52% (30/55 colonies) of these ‘AP bypass products’ were characterized by either dAMP incorporation opposite the AP site (A-rule) lacking downstream mutations or a ‘-1 deletion’ mechanism due to incorporation via the lesion loop-out mechanism (see above). The proposed mechanism for the former, as illustrated in Scheme 5.3 B, was supported by two crystal structures from Ling et al. (19) of Dpo4 with a DNA substrate containing an AP site. The first crystal structure, (structure Ab-4A), provides partial evidence for dAMP incorporation by Dpo4 opposite the lesion in an intrahelical fashion while structure Ab-3 shows Dpo4 adopts a conformation allowing in-frame extension from this incorporation event generating the observed full-length product. The latter mechanism for the generation of the frameshift mutation via the lesion loop-out mechanism (Scheme 5.4 E) was derived directly from two different structures in the same report. Structure Ab-1, which has the AP site at the first “templating” position, demonstrates that in the presence of an incoming nucleotide that can form a Watson-
Crick basepair with the template base 5’ to the AP site, Dpo4 can accommodate an extrahelical AP site in the gap between its finger and little finger domains, shifting the template base 5’ to the lesion into the active site to direct nucleotide incorporation. The second structure, Ab-2A, captures the ability of Dpo4 to extend from this incorporation event retaining the extrahelical AP lesion (19). The remaining nineteen ‘AP bypass products’ were each observed only 2 to 4% of the time and contain primarily deletion and substitution mutations. For example, the bypass event in Scheme 5.3 A occurred via incorporation of dAMP opposite the lesion followed by a substitution mutation that proceeded either by a simple mismatch incorporation of dTMP opposite the downstream 5’ template guanine, or via a mechanism reminiscent of the type II Dpo4 structure reported with undamaged DNA, where the incoming dTTP instead basepairs with the downstream template adenine (9), similar to that observed in the Ab-2B structure (19). Other proposed mechanisms involved more complicated DNA structural rearrangements as in Scheme 5.3 F where after dAMP incorporation opposite the AP site, the proposed pathway splits into two subpathways involving either; (i) a type II-like intermediate followed by a subsequent two-nucleotide loop-out in the template strand to generate the double deletion or (ii) an intermediate involving an extrahelical base in the primer similar to what was observed in Ab-4B (19) followed by extension and rearrangement of the primer-template to generate the ‘-2 deletion’. Importantly, while the assay temperature is not physiologically relevant for Dpo4, it is relevant for other mesophilic Y-family DNA polymerases, especially those from the DinB subfamily (i.e. human Polκ and E. coli DNA polymerase IV) which, like Dpo4, have all been shown to generate ‘-1 deletions’ even on undamaged DNA (26-28).

Analysis of AP bypass products at 75 °C. The same assay was performed at 75 °C, the in vivo temperature for S. solfataricus, to determine the effect of temperature on the
mutation profile for AP lesion bypass catalyzed by Dpo4. This assay was buffered with 50 mM potassium phosphate instead of 50 mM HEPES due to the pH stability of this buffer over a range of temperatures, while all other components in buffer R (Materials and Methods) remained unchanged. Notably, the change in buffer conditions did not affect the burst kinetics reported previously (data not shown, (29)). To prevent the melting of the 17-mer/41AP duplex at 75 °C prior to polymerization, the DNA substrate was preincubated with Dpo4 at room temperature for 10 min prior to incubation at 75 °C such that Dpo4 binding would stabilize the DNA duplex. For the reaction at 75 °C, we were able to extract sequencing information from 45 individual colonies using SOSA (Figure 5.2). Again, as observed at 37 °C, the results showed all four incorporations upstream of the AP site were replicated faithfully and lacked a single mutation (180 total incorporation events). However, once Dpo4 encountered the AP lesion, it generated a myriad of frameshift and substitution mutations resulting in sixteen distinct ‘AP bypass products’. The corresponding pathways were not illustrated as were those at 37 °C due to their modest redundancy. Yet, the precise mutational spectrum differed from that observed at 37 °C. While at 37 °C, 13% of the ‘AP bypass products’ were in-frame with dAMP located opposite the AP site (Figure 5.1), we similarly found 11% (5/45 colonies) had this sequence context at 75 °C (Figure 5.2). Likewise, at both 37 °C and 75 °C there were very few ‘AP bypass products’ that contained only base additions (2% at 37 °C, none at 75 °C). However, at 75 °C, 80% (36/45 colonies) contained only single/multiple base deletions (versus 56% at 37 °C), 7% (3/45 colonies) were complex transactions (versus 13% at 37 °C) and only 2% (1/45 colonies) had base substitutions (versus 9% at 37 °C). Thus, at higher temperature, Dpo4 generated more deletion mutations at the expense of a similar magnitude decrease in complex mutations, base substitutions, and in-frame incorporations with dCMP, dGMP or dTMP opposite the AP site (Table 5.2). In addition, regardless of specific mutations, the percentage of overall frameshift mutations
(both addition and deletion mutations) increased from 71% (39/55 colonies) at 37 °C to 87% (39/45 colonies) at 75 °C. It was not clear what caused the shift in the frameshift frequency. It is possible that the energy barriers for nucleotide incorporation into non-canonical ternary complexes could be overcome by more conformationally active Dpo4 at higher temperatures. We have observed an increase in the rate constant for an incorrect dGTP by 27-fold with a 30 °C increase in the reaction temperature (data not shown).

Yet, even more fascinating was the dramatic increase in bypass events characterized by dAMP incorporation opposite the AP site at 75 °C (69%, 31/45 colonies, Table 5.3) versus 37 °C (29%, 16/55 colonies, Table 5.3) and the concomitant decrease in dCMP incorporation events (13% or 6/45 colonies at 75 °C versus 53% or 29/55 colonies at 37 °C, Table 5.3). Further inspection of the ‘AP bypass products’ revealed a significant reduction of ‘-1 deletion’ mutations opposite the AP site via the lesion loop-out mechanism at higher reaction temperatures (51% or 28/55 colonies at 37 °C versus 13% or 6/45 colonies at 75 °C, Figures 5.1 to 5.3, Scheme 5.1 B). These results suggested that the equilibrium between an intrahelical and extrahelical AP site favored the former substrate at 75 °C. These intrahelical DNA substrates may be preferred at high temperatures due to a destabilization of the base stacking interactions required to facilitate the collapsed AP site or simply due to the increased thermal motion of Dpo4 at higher temperatures which may eliminate the cavity between its little finger and finger domains (19) required to accommodate the extrahelical lesion. Notably, a previous report of AP site bypass by Dpo4 at 70 °C similarly concluded that Dpo4 incorporated dAMP opposite the lesion approximately 75% of the time (18).

We feel that it is important to note the results obtained for SOSA performed at 75 °C indicate that 58% of the ‘AP bypass products’ (26/45 colonies) contain a mutation one
nucleotide upstream from where the PCR primer anneals. This is in contrast to the reaction performed at 37 °C, where only 9% (5/55 colonies) of the ‘AP bypass products’ contain such a mutation. As such, it is possible that there exists a small population of isolated ‘AP bypass products’ that could not be amplified at 75 °C as a result of several mispairs between the ‘AP bypass product and the PCR primer in this region. We have since evolved the assay to eliminate this problem in future studies with the human Y-family polymerases.

Comparison of SOSA results to previous kinetic observations. To determine whether the results from this novel SOSA agreed with our kinetic results reported in Chapter 4, we calculated the percentage of ‘AP bypass product’ containing each specific dNTP located at the 22nd position from the 5’-end, corresponding to incorporation ‘opposite’ the AP site (Table 5.1), and compared it to the calculated kinetic probabilities for the analogous incorporation with the 21-mer/41AP substrate (Table 5.1). Notably, this substrate contained a template guanine 5’ to the AP site, and both assays were performed at 37 °C. Mechanistically, the 22nd nucleotide from the 5’-end of the sequenced ‘AP bypass product’ was either incorporated directly opposite the lesion or via some other mechanism involving the local sequence context, e.g. lesion loop-out mechanism (see above). This was suggested by the error-free replication observed at all four upstream nucleotides (Materials and Methods), corresponding to the 18th-21st nucleotides, and the subsequent chaotic incorporation events downstream from the lesion, specifically the 22nd-26th nucleotides. We chose to compare this particular incorporation event opposite the lesion instead of events further downstream because, as mentioned above, Dpo4 generated twenty-one unique ‘bypass products’ each via a distinctive pathway. Thus, this first incorporation event opposite the lesion possessed the largest sample size and thus the greatest reliability over subsequent events for analysis. In addition, the anomalous nature
of downstream events precluded accurate comparison (Schemes 5.3 to 5.6). Interestingly, the results from our novel SOSA matched the results obtained from our kinetic studies. Specifically, the sequencing assay showed dATP (16/55 colonies, 29%) and dCTP (29/55 colonies, 53%) were the major incorporation events. These percentages coincided with those calculated from our kinetic studies (dATP: 33%, and dCTP: 57%) (Table 5.4). Furthermore, the probabilities for incorporation of dGTP (2/55 colonies, 4%) and dTTP (8/55 colonies, 15%) were much less frequent in this sequence context yet were very similar to those probabilities gleaned from our kinetic assays (1% and 9% for dGTP and dTTP respectively, Table 5.4). Thus, this novel assay correlated well to our kinetic studies of incorporation opposite the AP lesion and demonstrated strong competition between the A-rule and the lesion loop-out mechanism for AP lesion bypass.

Quantitative verification of the bypass percentage via the lesion loop-out mechanism by an enzymatic digestion assay. To provide an independent method for characterizing AP bypass catalyzed by Dpo4, an enzymatic digestion assay was also developed to quantitatively determine the percentage of ‘AP bypass product’ that resulted from the lesion loop-out mechanism (Scheme 5.1 B). Quantitative discrimination was based on the ability of a restriction enzyme, BmgB I, to digest a DNA substrate containing ‘AP bypass product’ annealed to a complementary 40-mer template (40Te, Table 5.1) possessing an intact BmgB I restriction site located precisely in the region of the ‘AP bypass product’ that corresponded to incorporation at the AP site. This complementary DNA substrate 40Te was designed so that when annealed to the pool of ‘AP bypass products’ synthesized by Dpo4, it would only anneal completely intrahelical and in the correct basepair registry with those sequence contexts generated by the lesion loop-out mechanism (Scheme 5.7, Pathway A, BmgB I restriction site is underlined and in bold type). All other sequence contexts would anneal with either mismatched basepairing or
extrahelical bases (Scheme 5.7, Pathway B, only one such sequence context shown) at the restriction site which would necessarily preclude digestion by BmgB I. Several control reactions were performed and run alongside the digested ‘AP bypass product’ synthesized by Dpo4 as described in Figure 5.3. The control reactions showed the restriction enzyme BmgB I was sufficiently sensitive to perform the desired assay accurately (Figure 5.3). The results for the BmgB I digestion assay, performed in triplicate, showed that 32% to 39% of the pool of ‘AP bypass products’ synthesized by Dpo4 at 37 °C contained the BmgB I restriction site and was therefore generated by the lesion loop-out mechanism described above (Figure 5.3, lane 5). These results are fairly consistent with results obtained from the SOSA which determined that 42% (23/55 colonies, Figure 5.1) of the population of ‘AP bypass products’ followed this mechanism. Likewise, we performed this digestion assay at 75 °C and found 7% to 8% of the population of ‘AP bypass products’ contained the BmgB I restriction site (data not shown) which was similar to 9% (4/45 colonies) obtained from the SOSA (Figure 5.2). Since this enzymatic digestion assay involved the entire pool of ‘AP bypass products’ synthesized by Dpo4, it verified the reliability of the SOSA technology.

_Broad impact of our studies._ Studies have shown that AP sites are lethal when introduced into biologically active DNA (30,31). The primary reason for this lethality is due to the inefficient incorporation and extension past these lesions by high fidelity replicative DNA polymerases. Evidence of this inability for replicative polymerases to bypass AP sites are abundant (32-35), _e.g._ _S. solfataricus_ PolB1 (Chapter 4). However, it is clear from our analysis and the series of published Dpo4 structures bound to DNA with an AP site that there are multiple mechanisms that Dpo4 uses to bypass and extend from this lesion (19). Similarly, results from analysis of Dpo4 replication of a 1,N2-ethenoguanine lesion also indicate the existence of multiple bypass mechanisms that generate a subset of
products even for a single sequence context (20). Additional crystal structures of Dpo4 with the bulky, covalently-modified \textit{cis-syn} thymine-thymine dimer (36) show that Dpo4 can accommodate this entire lesion into its active site, whereas structural analysis of Dpo4 bound to an adducted DNA containing benzo[a]pyrene suggests a mechanism by which Dpo4 can alternatively shift extremely bulky bases into the major groove to perform translesion synthesis (22). Dpo4 has also been shown to catalyze the error-free bypass of the bulky 8-oxodG through Watson-Crick basepairing with an incoming dCTP (37,38). Taken together, these results suggest structural plasticity and flexibility in the Dpo4 active site as was recently described by Mizukami \textit{et al.} (39). This flexibility translates into a more promiscuous mode of replication that is structurally justified by the limited contacts between Y-family members and the nascent basepair as seen in crystal structures of eukaryotic polymerases \(\eta, \iota, \kappa,\) and REV 1 as well as archaeal Dpo4 and Dbh. This plasticity bestows the Y-family with the exceptionally unique ability to perform DNA transactions forbidden by the replicative polymerases to facilitate lesion bypass.

So what does the observed chaotic bypass mean in the context of an \textit{in vivo} system? Our previous kinetic studies suggested two competing mechanisms for the bypass of an AP lesion by the Y-family polymerase Dpo4. However, unlike the conclusion reached by Ling \textit{et al.} (19) suggesting that Dpo4 does not use the AP site to instruct nucleotide insertion but rather uses the 5’ template base, our kinetic studies in Chapter 4 coupled with the two independent assays shown here involving direct and indirect sequencing methods, categorically demonstrate that there exist two primary mechanisms for AP site bypass. We believe that the basis for the preference of bypass involves the relative stability of the AP site (intrahelical or extrahelical) in the Dpo4 active site and is influenced by reaction temperature. Our assay showed that incorporation events
downstream from the AP site were far more anomalous than a previous report (18) which
did not specifically describe these events and were similar in this respect to results
described by Zang et al. (20) for bypass of a 1,N$_2$-e-G adduct by Dpo4. Yet, the relevance
of these complex downstream incorporation events is subject to debate in the context of
an in vivo system. Besides the inherently low fidelity of Dpo4 on undamaged DNA, the
observed spectrum of mutations downstream from the AP site indicates that it is
absolutely critical that replication by Dpo4 be limited to the lesion in order to guarantee
genome stability. Thus the role of polymerase switching in bypass synthesis, which has
been suggested to be mediated by the ubiquitination of PCNA (40,41), will likely prove
to be an important mechanism in regulating continued synthesis by Y-family members.

Overall, we have developed a novel short oligonucleotide sequencing assay that produced
results that demonstrate two dominant mechanisms used by Dpo4 to bypass an AP lesion
and provide a powerful new tool to directly sequence mixtures of small oligonucleotides
synthesized by any DNA polymerase or reverse transcriptase. We are currently using this
assay to characterize AP bypass by the eukaryotic Y-family polymerases. This
technically feasible method can be applied to the bypass of any DNA lesion and can be
used in multiple enzyme/protein systems.
Figure 5.1. Mutation spectrum of AP site bypass catalyzed by Dpo4 at 37 °C. See legend of Figure 5.2 (below) for color scheme explanation.
<table>
<thead>
<tr>
<th>Pathway</th>
<th>Adenine Incorporation</th>
<th>Cytosine Incorporation</th>
<th>Guanine Incorporation</th>
<th>Thymine Incorporation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) (1/45)</td>
<td>A</td>
<td>$5'$-gcagctcgaaccgtccagaattcgtc-3'</td>
<td>A</td>
<td>$5'$-gcagctcgaaccgtaatcgtc-3'</td>
</tr>
<tr>
<td>(B) (1/45)</td>
<td>A C</td>
<td>E: (C) (1/45)</td>
<td>C G</td>
<td>B: (2/45)</td>
</tr>
<tr>
<td>(C) (1/45)</td>
<td>A G</td>
<td>F: (G) (1/45)</td>
<td>C C</td>
<td>C: (4/45)</td>
</tr>
<tr>
<td>(D) (5/45)</td>
<td>ACGTC</td>
<td>(E) (6/45)</td>
<td>TC</td>
<td>(I) (5/45)</td>
</tr>
<tr>
<td>(E) (6/45)</td>
<td>CTCA</td>
<td>F: (H) (1/45)</td>
<td>GT</td>
<td>(J) (3/45)</td>
</tr>
<tr>
<td>(F) (6/45)</td>
<td>ACGTCA</td>
<td>G: (I) (1/45)</td>
<td>G</td>
<td>(K) (3/45)</td>
</tr>
<tr>
<td>(G) (1/45)</td>
<td>CGTC</td>
<td>TC</td>
<td>C</td>
<td>(L) (4/45)</td>
</tr>
<tr>
<td>(H) (1/45)</td>
<td>CG</td>
<td>GT</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>(I) (5/45)</td>
<td>CGTC</td>
<td>G</td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>(J) (3/45)</td>
<td>CG</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>(K) (3/45)</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>(L) (4/45)</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 5.2. Mutation spectrum of AP site bypass catalyzed by Dpo4 at 75 °C.** Results from the SOSA are shown separately based on specific nucleotide incorporated opposite the AP site. Only full-length ‘AP bypass product’ (FABP) is shown in its entirety. Sequences corresponding to primers used for amplification are shown in lower-case while sequenced nucleotides are upper-case. Individual base substitutions (blue) are shown above the expected FABP. Base deletions (red) and base additions (green) are shown below the FABP while complex mutations are color-coded based on specific mutation in order of occurrence and shown either above or below the expected FABP. Boldface letter corresponds to incorporation opposite the AP lesion. Relative ratios are shown at right in parentheses.
**Figure 5.3. BmgB I digestion assay.** $^{32}$P-labeled ‘AP bypass product’ was purified (Experimental Procedures) and digested with BmgB I for three hours and the resulting digestion fragments were resolved by native PAGE. The digestion reaction and controls were loaded as follows (all lanes digested with BmgB I unless mentioned): Lane 1, 40Pr/40Te duplex with intact BmgB I restriction site in the absence of BmgB I; Lane 2, 41Pr/40Te duplex with single unpaired base within BmgB I restriction site; Lane 3, 40Pr/40Te duplex with intact BmgB I restriction site; Lane 4, 41Pr/41Te duplex with single basepair addition in BmgB I restriction site; Lane 5, Dpo4 bypass product catalyzed at 37 °C. Assay performed in triplicate with reaction indicating 32% digestion shown in Lane 5.
<table>
<thead>
<tr>
<th>Primers</th>
<th>17-mer</th>
<th>5’-CGCAGCCGTCCAAACCA-3’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>21-mer</td>
<td>5’-CGCAGCCGTCCAAACCAACTCA-3’</td>
</tr>
<tr>
<td></td>
<td>40Pr</td>
<td>5’-CGCAGCCGTCCAAACCAACTCAGTCGATCCCAATGCCGTCC-3’</td>
</tr>
<tr>
<td></td>
<td>41Pr</td>
<td>5’-CGCAGCCGTCCAAACCAACTCAGTCGATCCCAATGCCGTCC-3’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Templates</th>
<th>41AP</th>
<th>3’-GCGTCGGCAGGTGTTGAGTXGCAGCTAGGTTACGGCAGG-5’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40Te</td>
<td>3’-GCGTCGGCAGGTGTTGAGTGTCAGCTAGGTTACGGCAGG-5’</td>
</tr>
<tr>
<td></td>
<td>41Te</td>
<td>3’-GCGTCGGCAGGTGTTGAGTGTCAGCTAGGTTACGGCAGG-5’</td>
</tr>
</tbody>
</table>

BmgB I restriction site is underlined in relevant DNA substrates. X designates an AP site.

Table 5.1. DNA Primers and Templates.
<table>
<thead>
<tr>
<th>Mutation Type</th>
<th>Reaction at 37 °C</th>
<th>Reaction at 75 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of 'AP</td>
<td>Percentage (%)^b</td>
</tr>
<tr>
<td></td>
<td>bypass products^a</td>
<td></td>
</tr>
<tr>
<td>1 base substitution</td>
<td>5</td>
<td>9.1</td>
</tr>
<tr>
<td>2 base addition</td>
<td>1</td>
<td>1.8</td>
</tr>
<tr>
<td>1 base deletion</td>
<td>27</td>
<td>49.1</td>
</tr>
<tr>
<td>2 base deletion</td>
<td>1</td>
<td>1.8</td>
</tr>
<tr>
<td>3 base deletion</td>
<td>2</td>
<td>3.6</td>
</tr>
<tr>
<td>4 base deletion</td>
<td>1</td>
<td>1.8</td>
</tr>
<tr>
<td>5 base deletion</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>complex transaction</td>
<td>7</td>
<td>12.7</td>
</tr>
<tr>
<td>full length^c</td>
<td>11</td>
<td>20.0</td>
</tr>
<tr>
<td>total</td>
<td>55</td>
<td>100</td>
</tr>
</tbody>
</table>

^aDesignates the number of colonies containing 'AP bypass products' with a particular mutation.
^bPercentage refers to percentage of a particular mutation calculated as [(Number of 'AP bypass products')/Total] x 100.
^cDesignates those colonies possessing full length DNA substrates containing no additional downstream mutations.

Table 5.2. Mutational spectrum of short oligonucleotide sequencing assay at 37 °C and 75 °C.
Table 5.3. Comparison of nucleotide incorporation opposite an AP site at 37 °C and 75 °C.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>Probability at 37 °C</th>
<th>Probability at 75 °C&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>dATP</td>
<td>29%</td>
<td>69%</td>
</tr>
<tr>
<td>dCTP</td>
<td>53%</td>
<td>13%</td>
</tr>
<tr>
<td>dGTP</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>dTTP</td>
<td>14%</td>
<td>2%</td>
</tr>
</tbody>
</table>

Note - Probability = (Number of colonies with particular incorporation)/(Total number of colonies).
<sup>a</sup>The remaining 14% (6/45 colonies) contained no incorporation opposite the AP site.

Table 5.3. Comparison of nucleotide incorporation opposite an AP site at 37 °C and 75 °C.
Table 5.4. Comparison of methods for the preference of nucleotide incorporation opposite an AP site at 37 °C.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>Kinetic probability(^b)</th>
<th>SOSA probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>dATP</td>
<td>33%</td>
<td>29%</td>
</tr>
<tr>
<td>dCTP</td>
<td>57%</td>
<td>53%</td>
</tr>
<tr>
<td>dGTP</td>
<td>1%</td>
<td>4%</td>
</tr>
<tr>
<td>dTTP</td>
<td>9%</td>
<td>14%</td>
</tr>
</tbody>
</table>

\(^{a}\)Note – AP site designated as “X”.
\(^{b}\)Kinetic probability = \(\frac{k_p}{K_d}\)\(\sum\)\(\frac{k_p}{K_d}\)\(dATP, dCTP, dGTP, dTTP\); taken from our previous studies (accompanying manuscript).
Scheme 5.1. Two primary pathways used by Dpo4 to bypass an abasic site lesion.

X: abasic lesion
Scheme 5.2. Diagram of the short oligonucleotide sequencing assay (SOSA).
Scheme 5.3. Proposed mechanisms for ‘AP bypass products’ with adenine opposite the AP site.
Adenine Incorporation

5’-ACTCA
3’-TGAGTXTCAG---5’

A
5’-ACTCAAT
3’-TGAGTXTCAG---5’

B
5’-ACTCAA
3’-TGAGTXTCAG---5’

C
5’-ACTCAAG
3’-TGAGTXTCAG---5’

D
5’-ACTCAAC
3’-TGAGTXTCAG---5’

E
5’-ACTCAAC-T
3’-TGAGTXTCAG---5’

F
5’-ACTCAAC-T
3’-TGAGTXTCAG---5’

G
5’-ACTCAAC-T
3’-TGAGTXTCAG---5’

H
5’-ACTCAAC-T
3’-TGAGTXTCAG---5’
Scheme 5.4. Proposed mechanisms for ‘AP bypass products’ with cytosine opposite the AP site.
Cytosine Incorporation

A

5'---ACTCA
3'---TGAGTXGCAG--5'

5'---ACTCA
3'---TGAGT_GCAG--5'

5'---ACTCA CA-C
3'---TGAGT_GCAG--5'
realignement

B

5'---ACTCA
3'---TGAGTXGCAG--5'

5'---ACTCA
3'---TGAGT_GCAG--5'

5'---ACTCA CATC--3'
3'---TGAGT_GCAG--5'

C

5'---ACTCA
3'---TGAGTXGCAG--5'

5'---ACTCA
3'---TGAGT_GCAG--5'

5'---ACTCA CGT
3'---TGAGT_GCAG--5'

D

5'---ACTCA
3'---TGAGTXGCAG--5'

5'---ACTCA
3'---TGAGT_GCAG--5'

5'---ACTCAC
3'---TGAGTXGCAG--5'

E

5'---ACTCA
3'---TGAGTXGCAG--5'

5'---ACTCA
3'---TGAGT_GCAG--5'

5'---ACTCA CGTC--3'
3'---TGAGT_GCAG--5'

F

5'---ACTCA
3'---TGAGTXGCAG--5'

5'---ACTCA
3'---TGAGT_GCAG--5'

5'---ACTCA--A--3'
3'---TGAGT_G_G--5'
Guanine Incorporation

\[
\begin{align*}
5' &\longrightarrow A \\
3' &\longrightarrow TGAGTXGCAGCTA\longrightarrow 5'
\end{align*}
\]

A  
\[
\begin{align*}
5' &\longrightarrow ACTCA \\
3' &\longrightarrow TGAGTXGCAG\longrightarrow 5' \quad \longrightarrow 5' &\longrightarrow ACTCAG \\
3' &\longrightarrow TGAGTXGCAG\longrightarrow 5' \quad \longrightarrow 3' &\longrightarrow TGAGTXGCAG\longrightarrow 5'
\end{align*}
\]

B  
\[
\begin{align*}
5' &\longrightarrow ACTCA \\
3' &\longrightarrow TGAGTXGCAG\longrightarrow 5' \quad \longrightarrow 5' &\longrightarrow ACTCAG \\
3' &\longrightarrow TGAGTXGCAG\longrightarrow 5' \quad \longrightarrow 3' &\longrightarrow TGAGTXGCAG\longrightarrow 5'
\end{align*}
\]

\[
\begin{align*}
5' &\longrightarrow ACTCAGCA \\
3' &\longrightarrow TGAGTXGCAG\longrightarrow 5' \quad \longrightarrow 5' &\longrightarrow ACTCAGCATC\longrightarrow 3' \\
3' &\longrightarrow TGAGTXGCAG\longrightarrow 5' \quad \longrightarrow 3' &\longrightarrow TGAGTXGCAG\longrightarrow 5'
\end{align*}
\]

Scheme 5.5. Proposed mechanisms for ‘AP bypass products’ with guanine opposite the AP site.
Scheme 5.6. Proposed mechanisms for ‘AP bypass products’ with thymine opposite the AP site.
Thymine Incorporation

A

5'-ACTCA
3'---TGAGTXGCAG---5'

5'-ACTCAT
3'---TGAGTXGCAG---5'

5'-ACTCATA
3'---TGAGTXGCAG---5'

B

5'-ACTCA
3'---TGAGTXGCAG---5'

5'-ACTCAT
3'---TGAGTXGCAG---5'

5'-ACTCATGTC---3'
3'---TGAGTXGCAG---5'

C

5'-ACTCA
3'---TGAGTXGCAG---5'

5'-ACTCAT
3'---TGAGTXGCAG---5'

5'-ACTCAT A
3'---TGAGTX G---5'

GCA

5'-ACTCAT A---3'
3'---TGAGTX G---5'

GCA

D

5'-ACTCA
3'---TGAGTXGCAG---5'

5'-ACTCAT
3'---TGAGTXGCAG---5'

5'-ACTCAT C
3'---TGAGTX G---5'

GCA

5'-ACTCAT C---3'
3'---TGAGTX G---5'

GCA

E

5'-ACTCA
3'---TGAGTXGCAG---5'

5'-ACTCAT
3'---TGAGTXGCAG---5'

5'-ACTCAT G
3'---TGAGTX CAG---5'

5'-ACTCAT-GTC---3'
3'---TGAGTXGCAG---5'
Scheme 5.7. Possible sequence contexts with regard to the BmgB I digestion assay.
5.5 References


6.1 Introduction

The DNA polymerases of the Y-family are known for their ability to bypass a variety of DNA lesions that, to varying degrees, impede the continuous synthesis of genomic DNA catalyzed by high-fidelity replicative DNA polymerases. In addition to these roles in lesion bypass, recent studies have suggested additional functions of these specialized polymerases in the cell (1,2). These Y-family enzymes are found in all three life domains and are notably devoid of 3’-5’ exonuclease activities which are genuine fidelity checking mechanisms inherent to replicative DNA polymerases. These lesion bypass DNA polymerases are also shown to possess conformationally flexible active sites (3-8) which both contribute to their observed lesion-bypass capabilities and low polymerization fidelities.

It has been shown previously that T7 DNA polymerase utilizes an induced-fit mechanism that is powered by the energy of nucleotide binding to select a correct nucleotide from a pool of similar nucleotide substrates (9). Similar results have been provided as convincingly for DNA polymerase I (Klenow fragment) (10) and HIV-RT (11). To our
knowledge, pre-steady state mechanistic studies have been performed for only three Y-
family DNA polymerases with an undamaged DNA substrate. Both Saccharomyces
cerevisiae polymerase η (polη) (12) and Sulfolobus solfataricus DNA polymerase IV
(Dpo4) (13) have been also shown to utilize an induced-fit mechanism for selecting a
correct nucleotide while studies of Sulfolobus acidocaldarius DinB homolog (Dbh) (14)
very unconvincingly concluded that Dbh does not follow such a mechanism. In the latter
study, Cramer and Restle surprisingly argue against an induced-fit mechanism for Dbh
based on the results gleaned from the determination of the elemental effect for both
correct and incorrect incorporation, which suggest that both incorporations are limited by
the chemistry step. However, these elemental effect studies are largely considered a poor
diagnostic for determining the rate-limiting step for DNA polymerases (15) and need to
be supplemented with additional data, not provided by Cramer and Restle. In addition,
they further report that the $k_{off}$ rate constant ($\sim 63$ s$^{-1}$) is at least 10-fold higher than the
polymerization rate ($k_{pol}$) yet are able to observe a burst in the first turnover due to a
possible reduced $k_{off}$ rate in the presence of high concentrations of nucleotide (14). But
this possibility suggests the existence of an induced-fit mechanism in order to account for
a reduced $k_{off}$ which would most reasonably be due to a conformational change to slow
the DNA dissociation. Clearly, more studies of the Dbh catalytic mechanism needs to be
performed before such a conclusion can be reached.

In two previous publications, we used Dpo4 and the methods of pre-steady state kinetics
to determine the fidelity (16) and mechanism (13) of single nucleotide incorporation into
undamaged DNA at 37 °C. These studies demonstrated that similar to reports for other Y-
family polymerases, Dpo4 possesses very low incorporation fidelity ($10^{-3}$ to $10^{-4}$) several
orders of magnitude below that observed for replicative DNA polymerases. However,
owing to the fact that the S. solfataricus species propagate in an environment where their
physiological temperature varies between 75 °C and 85 °C (17,18) it becomes important to address what effect temperature may have on the kinetic mechanism of Dpo4 we previously established at 37 °C (13). Thus we exploit the thermal stability of Dpo4 in order to probe key mechanistic parameters to determine the effect of temperature on the reaction mechanism. Here we provide evidence indicating that both the fidelity and induced-fit mechanism remain unchanged with an increase in the reaction temperature although microscopic rate constants are dramatically altered.

6.2 Materials and methods

*Materials.* These chemicals were purchased from the following companies: $[^\alpha \cdot 32\ P]dTTP$ and $[^\gamma \cdot 32\ P]ATP$, GE Healthcare (Picataway, NJ); dNTPs, Gibco-BRL (Rockville, MD); Sp-dTTP$_{\alpha S}$ and Sp-dGTP$_{\alpha S}$, Biolog, Life Science Institute (Bremen, Germany); ddTTP, Trilink Biotechnologies (San Diego, CA). Full-length Dpo4 fused to a C-terminal His$_6$ tag was overexpressed in *E. coli* as described in (16). The protein was stored in 20 μl aliquots and stored at -80 °C. The DNA substrate (D-1) listed in Figure 6.1 was prepared as described previously (16).

*Pre-steady state kinetic assays.* All experiments using Dpo4, if not specified, were performed in a reaction buffer D containing 50 mM HEPES (pH 7.5 at all temperatures), 5 mM MgCl$_2$, 50 mM NaCl, and 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA. All reactions were carried out at temperatures ranging from 2 °C to 56 °C using either a rapid chemical quench flow apparatus (KinTek, PA) as described
previously (16) or via manual quench. In some cases, we used conditions at subsaturating nucleotide concentration in order to mimic the assays performed previously at 37 °C, and due to the fact that the rate constants at 56 °C was very close to the upper limit that can be measured by instrument, and for reasons explained previously (16).

Active site titration assay. The equilibrium dissociation constant ($K_d$) of the Dpo4•DNA binary complex was determined by mixing a preincubated solution of Dpo4 (30 nM) and increasing concentrations of a 21/41-mer DNA substrate (D-1) with 1.2 mM dTTP and subsequently quenching the reaction after 26 ms (equivalent to seven half-times) in order to achieve the maximal first turnover amplitude. The resulting burst amplitudes were plotted against the concentration of the D-1 substrate and fit via quadratic regression (equation 5).

Measurement of the phosphorothioate elemental effect. A preincubated solution of Dpo4 (120 nM) and 5’-labeled D-1 (30 nM) was rapidly mixed with either dTTP (100 µM) or dTTPαS (100 µM, >95% purity) in buffer D at 56 ºC. The reactions were quenched with 0.37 M EDTA, analyzed via denaturing gel electrophoresis, and quantitated as described above. The data were fit using equation 1 to yield $k_{obs}$ values. For the incorporation of an incorrect nucleotide, reactions were initiated with either dGTP (100 µM) or dGTPαS (100 µM, >95% purity). For both correct and incorrect incorporation, the $S_p$ isomer was used as opposed to the $R_p$ isomer due to the stereoselectivity observed with rPolβ (19).

Pulse-chase and pulse-quench assays. To provide insight into whether or not a conformational change limited the rate constant for correct nucleotide incorporation in the first enzyme turnover, two time courses were performed by mixing a preincubated solution of Dpo4 (30 nM) and unlabeled D-1 (30 nM) with a solution containing 40 µM
[α-32P]dTTP in buffer D at 56 °C. Reactions for both time courses were mixed for times ranging from 20 ms to 2 s. After this variable period, reactions were either immediately quenched with 1N HCl (pulse-quench) or chased for 15 seconds (pulse-chase) with a large molar excess of unlabeled dTTP (2.5 mM), before being quenched with 1 N HCl. Dpo4 was subsequently extracted from each reaction using chloroform, followed by rapid neutralization with 1 N NaOH/0.1 M Tris. Reactions were quantitated via sequencing gel analysis using a highly crosslinked 20 % polyacrylamide matrix as described previously (20) to resolve product from unincorporated [α-32P]dTTP.

Circular dichroism (CD) spectroscopic studies. CD spectra were measured using an AVIV CD spectrometer model 62A DS (Lakewood, NJ). UV spectra were acquired at a fixed wavelength of 222 nm using a 1 mm pathlength cuvette. Samples were dissolved into a degassed buffer (25 mM NaPO4 pH 7.5, 50 mM NaCl, 5 mM MgCl2, 10% glycerol) and filtered with a 0.45 µm membrane to remove residual aggregation. Baseline spectra of buffer alone was obtained and subtracted from sample spectra. The resulting ellipticity (mdeg) was plotted as a function of temperature. To determine the ellipticity of the Dpo4, 30 µM of the protein was added to a solution containing the degassed buffer and was equilibrated at each temperature for ten minutes before scanning.

Product analysis. Unless noted, reaction products were analyzed by sequencing gel electrophoresis (17% acrylamide, 8 M urea, 1x TBE running buffer) and quantitated with a Phosphorimager 445 SI (Molecular Dynamics).

Data analysis. Data were fit by nonlinear regression using the program KaleidaGraph (Synergy Software). Data from the single turnover experiments were fit to equation 1:

\[ \text{[Product]} = A[1 - \exp(-k_{obs}t)] \]  \hspace{1cm} (1)
Data from the secondary plot for the single turnover experiments were fit to equation 2:

\[
[k_{obs}] = k_p[d\text{NTP}]/(K_d^{d\text{NTP}} + [d\text{NTP}])
\]  

(2)

Data from burst experiments were fit to equation 3:

\[
\text{[Product]} = A[1 – \exp(-k_1t) + k_2t]
\]  

(3)

Data from the steady-state kinetic experiments were fit to equation 4:

\[
\text{[Product]} = k_{ss}E_0t + E_0
\]  

(4)

Data from the active site titration were fit to equation 5:

\[
[E\cdot\text{DNA}] = \frac{1}{2}(K_d^{DNA} + E_0 + D_0) - \frac{1}{2} [(K_d^{DNA} + E_0 + D_0)^2 - 4E_0D_0]^{1/2}
\]  

(5)

Data from the Arrhenius activation energy treatment were fit to equation 6:

\[
k = A_r[\exp(-E_a/RT)]
\]  

(6)

Here, A represents the enzyme amplitude, \(k_1\) is the observed burst rate, and \(k_2\) is the observed steady-state rate, \(E_0\) the active enzyme concentration, \(k_{ss}\) is the steady-state rate constant of dNTP incorporation at initial active enzyme concentration of \(E_0\), \(k_{obs}\) the observed single turnover rate, \(K_d^{DNA}\) is the equilibrium dissociation constant for \(E\cdot\text{DNA}\) binary substrate, \(D_0\) the DNA concentration, \(K_d^{d\text{NTP}}\) is the equilibrium dissociation constant of dNTP from the \(E\cdot\text{DNA}\cdot d\text{NTP}\) complex, \(A_r\) is a proportionality constant, \(E_a\) is the activation energy, R is the universal gas constant, and T is reaction temperature in kelvin.
6.3 Results

**Effect of temperature on Dpo4 incorporation fidelity.** To determine the effect of temperature on the efficiency and fidelity of nucleotide incorporation catalyzed by Dpo4, single nucleotide incorporation studies were performed under single turnover conditions in order to isolate first-order kinetics without complications from subsequent turnovers. The ground-state binding affinity ($K_d$) for each individual nucleotide was measured through the dNTP concentration dependence of the observed single turnover rate constant. Dpo4 (120 nM) and 5’-[32P]-D-1 (30 nM) was first preincubated at room temperature, then at 37 °C, and finally at 56 °C prior to be mixed with increasing concentrations of a single dNTP to initiate nucleotide incorporation at 56 °C. Reactions were quenched at various times and analyzed as described previously (16). For the correct incorporation of dTTP into the D-1 substrate (Figure 6.1), the observed rate constants were obtained by fitting the data to a single exponential equation (Materials and Methods) for each individual time course (Figure 6.2). The resulting rate constants ($k_{obs}$) were then plotted against the concentration of nucleotide and fit via equation 2 (Materials and Methods) to yield the maximum nucleotide incorporation rate constant ($k_p$) which was determined to be 189 ± 13 s$^{-1}$, and the equilibrium dissociation constant ($K_d$) of 383 ± 65 μM for the dissociation of dTTP (Figure 6.3) from the Dpo4•D-1 binary complex. A similar assay was carried out at 38 °C for reasons explained below under the same conditions, where we obtained a $k_p$ of 17.7 ± 0.6 s$^{-1}$ and the $K_d$ of 221 ± 15 μM (Table 6.1, data not shown), which were both within 2-fold of the corresponding values at 37 °C (16). This indicated that with an 18 °C increase in reaction temperature, the rate constant increased 11-fold to a rate constant that approached the upper limitation for accurate measurement by our rapid chemical quench flow instrument. Yet, the $K_d$ for dTTP
binding increased (decreased affinity) by only 1.7-fold. The resulting substrate specificity \( (k_p/K_d) \) at 56 °C was subsequently calculated to be 0.493 \( \mu \text{M}^{-1}\text{s}^{-1} \), a value approximately 6-fold higher than that calculated at 38 °C. In addition, under the assumption that the incoming nucleotide (dNTP) freely diffuses into the Dpo4•D-1 active site (13), the \( k_{on} \) was assumed to be equivalent to diffusion control \((1 \times 10^8 \text{ M}^{-1}\text{s}^{-1})\), and as such, the \( k_{off} \) (or \( k_2 \) in Scheme 6.1) of dNTP from the Dpo4•D-1•dNTP ternary complex can be estimated \((k_{off} = k_{on}K_d)\) to be 38,300 s\(^{-1}\). Notably, we observed full reaction amplitude at all tested temperatures which suggested the integrity of the DNA duplex was not compromised at these higher reaction temperatures. A previous study assessing the overall mutational frequency of Taq DNA polymerase from Thermus aquaticus (21) also suggests that duplex stability had little effect on nucleotide discrimination at 70 °C. We believe the stability of the duplex is due to preincubating Dpo4 with DNA before exposing the mixture to elevated temperatures. Dpo4 binds DNA tightly (13) and therefore was able to stabilize the duplex DNA in its binding pocket.

In addition, each of the three incorrect dNTP incorporations into D-1 were performed at both 38 and 56 °C as described above, to determine the corresponding kinetic parameters (Table 6.1, data not shown). On average, there was approximately a 19-fold difference (9- to 35-fold) between the \( k_p \) values for incorrect incorporations at 56 °C to those obtained at 38 °C, yet only, on average, a 1.2-fold (0.9- to 1.5-fold) increase between the \( K_d \) values at these two temperatures. In an attempt to obtain \( k_p \) data over a larger range of temperatures, \( k_p \) estimations were performed at 2 °C and 14 °C by varying the temperature by 6 °C increments or multiples therein, to cover a 54 °C range. Not surprisingly, \( k_p \) determination at the lower reaction temperatures were severely limited by extremely slow turnover. The \( k_p \) estimated for correct (dTTP) and incorrect (dGTP) incorporation at 14 °C was 0.189 ± 0.008 s\(^{-1}\) and \((4.8 ± 0.8) \times 10^{-5} \text{s}^{-1}\), respectively (data
not shown). Due to the insignificant product formation for nucleotide incorporation at these lower temperatures, the fidelity could not be determined. However, based on these rate constants we estimated an increase of 1,000-fold and 5,000-fold for correct and incorrect nucleotide incorporation respectively over this 42 °C temperature range. Moreover at 2 °C the $k_p$ for correct incorporation was estimated to be $0.010 \pm 0.001 \, \text{s}^{-1}$ (data not shown) while the incorrect incorporation was not observed after three hours of incubation. Thus upon decreasing the temperature by 54 °C, the rate constant for correct nucleotide incorporation decreases by an amazing 18,900-fold.

In order to determine the change in the fidelity over this range of temperatures, the substrate specificities were calculated for all incorporations between 26 to 56 °C (Table 6.1). Interestingly, upon calculating the fidelity at each temperature (Table 6.1) we observed only a 3-fold difference over this 30 °C range.

**Biphasic kinetics of nucleotide incorporation at 56 °C.** We previously established the mechanism of single nucleotide incorporation catalyzed by Dpo4 at 37 °C (Scheme 6.1) (13). Yet, the question as to whether or not this mechanism changes with an increase in reaction temperature remains to be determined. In an effort to elucidate the mechanism at a more physiologically relevant temperature for Dpo4, these same studies were performed at 56 °C and not at 80 °C for the following reasons: (i) the instrument is not suitable for usage at temperatures above 70 °C (16); (ii) based on our kinetic results above for correct dNTP incorporation at 56 °C, the rate constant at 80 °C will exceed the capabilities of the instrument since the reaction will completed within the instrument’s dead time of mixing; (iii) at 80 °C the 21/41-mer D-1 DNA substrate will be largely melted requiring the use of a considerably longer primer-template that would preclude comparison to our previous results (13,16) not to mention the difficulty in separating the
product from substrate via gel electrophoresis for this longer substrate. To determine if Dpo4 follows the same biphasic kinetics that were reported at 37 °C (13), a pre-steady state kinetic assay at 56 °C for the correct incorporation of dTTP into the D-1 substrate (Figure 6.1) was performed by mixing a preincubated solution of 5’-[32P]-D-1 (120 nM) and Dpo4 (30 nM) with dTTP (100 µM) in buffer D (Materials and Methods). Reactions were then quenched with 0.37 M EDTA at times ranging from 5 ms to 4.5 s, analyzed via denaturing PAGE, and quantitated using a Phosphorimager. The resulting data were fit to equation 3 (Materials and Methods) and demonstrated the same biphasic nature observed at 37 °C, characterized by a fast-phase burst rate constant of 33.6 ± 5.0 s⁻¹ preceding a slow-phase rate constant of 0.11 ± 0.01 s⁻¹ (Figure 6.4). Thus, following the first turnover of the enzyme, the ensuing turnovers occurred at a rate constant over 300-fold slower due to the dissociation of the DNA product (DNAn+1, Scheme 6.1) from the binary complex as observed with other enzymes (9,13,22-26). The observed slow phase rate constant was verified via a subsequent assay performed under steady-state conditions where the DNA substrate was in molar excess (~200-fold) over the concentration of Dpo4 (Materials and Methods). Under these conditions, the observed steady-state rate which was determined by fitting the data to equation 4 (Materials and Methods), to be 0.23 s⁻¹ (Figure 6.5), which is approximately 2-fold different from that observed in the burst experiment. However, the rate-limiting step in the first turnover remained to be elucidated (see below).

Equilibrium dissociation constant of the E•DNA complex. Before probing the rate-limiting step in the first turnover of Dpo4, we needed to verify that the 19 °C increase in temperature did not significantly perturb the ability of Dpo4 to bind DNA as we reported previously (13). For reasons explained previously (13), we can measure the dependence of the burst amplitude on D-1 concentration to ascertain the equilibrium dissociation
constant ($K_d$) of the Dpo4•DNA binary complex at 56 °C. As seen in Figure 6.6, the burst amplitude increased with the addition of the D-1 substrate, reaching maximum amplitude at 400 nM D-1. The solid line (Figure 6.6) was a fit of the data to equation 5 (Materials and Methods) and yielded a $K_d$ value of 39.9 ± 6.3 nM. Using the DNA dissociation rate constant ($k_{-1}$) of 0.11 s⁻¹ determined above, the apparent second-order association rate constant of the Dpo4•D-1 complex was then calculated to be $k_1$ (or $k_{on}$) = $k_{off}/K_d = 2.8 \times 10^6$ M⁻¹s⁻¹, a binding rate constant which was below the diffusion limit and similar to the $k_1$ previously reported for Dpo4 at 37 °C as well as for T7 DNA polymerase (9), T4 DNA polymerase (27), and HIV-RT (11). These results indicate that the equilibrium dissociation constant of the Dpo4•D-1 complex at 56 °C was less than four-fold weaker than what was observed at 37 °C. Similar results showing a modestly insignificant change in the equilibrium dissociation constant have been reported for Taq DNA polymerase over a larger range of temperatures (28). Thus our single turnover conditions as described in Materials and Methods are valid based on these results.

**Determination of the rate-limiting step.** We have provided evidence previously that a conformational change limits the correct nucleotide incorporation rate constant, $k_p$, for Dpo4-catalyzed reactions at 37 °C. Yet, the identification of the rate-limiting step at a more physiologically relevant temperature remains to be determined. It is possible that as the temperature increases, the rate of this proposed rate-limiting conformational change ($k_3$) may in fact approach or overcome the rate of the chemical step ($k_4$) due to increased conformational dynamics. Thus, to determine if the $k_p$ at 56 °C is limited by a conformational change, the chemistry step, or a combination of both, we measured the value of the $\alpha$-thio elemental effect. Although this assay has been described as an unreliable diagnostic due to the possible steric clashes of the proposed pentacoordinated sulfur-containing intermediate with the enzyme, we believe that these studies may be
relevant in the context of Dpo4, due to its significantly less restrictive active site (3). The concept was originally proposed by Herschlag, Piccirilli, and Cech after observing that a rate-limiting chemical step involving the making or breaking of a phosphate bond in the hydrolysis of phosphate diesters shows a phosphorothioate elemental effect in the range of 4- to 11-fold (29). We performed single turnover studies for the correct incorporation of either dTTP or S_p-dTTP\(\alpha\)S (Materials and Methods). These reactions were quenched at various times and the resulting data were fit with equation 1 (Materials and Methods) yielding \(k_{obs}\) values of 29.3 ± 2.1 s\(^{-1}\) and 29.1 ± 1.9 s\(^{-1}\) for dTTP and S_p-dTTP\(\alpha\)S, respectively (Figure 6.7). The \(\alpha\)-thio elemental effect (\(k_{obs}^{dTTP}/k_{obs}^{dTTP\alphaS}\)) for this correct incorporation was determined to be 1.0, which suggested that the chemistry step was not rate-limiting for the incorporation of correct nucleotides by Dpo4. This conclusion was consistent with the experimental results shown below and moreover, was consistent with the results (value of the \(\alpha\)-thio elemental effect) gleaned at 37 ºC (13).

In addition, we measured the \(\alpha\)-thio elemental effect for the incorporation of an incorrect nucleotide, dGTP into the same D-1 substrate (Materials and Methods). After fitting these data to equation 1, the observed single turnover rate constants (\(k_{obs}\)) for dGTP and S_p-dGTP\(\alpha\)S respectively were 0.12 ± 0.01 s\(^{-1}\) and 0.024 ± 0.001 s\(^{-1}\) (Figure 6.8). The corresponding elemental effect of 5.0, which is in the range of 4-11 (29), suggested that the chemistry step likely limited the rate constant for incorrect nucleotide incorporation (29). This value was almost identical to the value (5.8) derived at 37 ºC for the same incorrect incorporation of dGTP (13), however, in lieu of the inherent ambiguity surrounding this assay, additional evidence is required to confirm this observation.

_Determination of the rates of dissociation from the E’•DNA•dNTP ternary complex._ To probe further into whether or not a conformational change is rate-limiting at 56 ºC, we
were able to determine the rates of dissociation of DNA from the E’•DNA•dNTP ternary complex under steady-state conditions. In two separate assays, a preincubated solution of Dpo4 (2.4 nM) and 5’-[32P]-D-1 (250 nM) was rapidly mixed with either dideoxyTTP (ddTTP) (1.2 mM) alone or with ddTTP (1.2 mM) and the next correct nucleotide dCTP (1.2 mM). The latter assay was used to measure the rate of DNA dissociation from a ternary complex that was incapable of performing chemistry due to the incorporation of ddTTP yet could still bind the next correct nucleotide, while the former assay, lacking dCTP, directly measures the DNA dissociation rate from the E•DNA binary complex. In the presence of ddTTP alone the steady-state rate (0.063 ± 0.001 s⁻¹) (Figure 6.9) was within 2-fold of that measured in our biphasic burst experiment (Figure 6.1). Conversely, when assayed in the presence of ddTTP and dCTP, a 22-fold slower dissociation rate (0.0028 ± 0.0004 s⁻¹) was observed (Figure 6.9) and suggested that DNA dissociation was inhibited by the presence of dCTP. This slow dissociation provided strong evidence for a conformational change from the E•DNA•dNTP ternary complex to a tighter binding E’•DNA•dNTP ternary complex, which released DNA with a significantly slower rate constant (k₈ in Scheme 6.1).

**Pulse-chase experiments.** To further explore the possibility that the previously suggested conformational change limited nucleotide addition at 56 °C, we used a well-established approach to attempt to identify an intermediate along the reaction pathway that would indicate the existence of a conformational change. In this set of experiments, a preincubated solution of Dpo4 and unlabeled D-1 was mixed with [α-32P]-dTTP for various time intervals (Materials and Methods). Reactions were immediately quenched by the addition of 1 N HCl in the first experiment, while in the second experiment, reactions were instead chased with an excess of unlabeled dTTP before quenching with 1 N HCl. The nucleotide trap in the pulse-chase reactions, precludes dissociated [α-32P]-
dTTP rebinding to the Dpo4•D-1 complex while allowing any competently bound Dpo4•D-1•[α-32P]-dTTP complexes to partition between dissociation and product formation. As such, additional product formation observed between the pulse-chase relative to the pulse-quench assay has been considered strong evidence for the existence of at least one distinct ternary complex (E’•DNA_n•dTTP) between the ground-state complex (E•DNA_n•dNTP) and product resulting from the chemistry step (E•DNA_{n+1}•PP_i). Our results were fit to equation 3 to yield a pulse-chase and pulse-quench amplitudes of 26.9 ± 1.8 nM and 23.3 ± 1.3 nM respectively (Figure 6.10). This corresponds to 3.6 nM of intermediate complex that was quenched in the latter assay yet could be chased to product, indicating the existence of an intermediate prior to the chemistry step. As argued previously, this intermediate complex was most likely the E’•DNA_n•dTTP (13).

**Measurement of the DNA dissociation rate from the E•DNA•dNTP ternary complex.** The ternary complex E•DNA•dNTP partitions between the conversion of DNA to product (E•DNA_{n+1}•PP_i) and the dissociation of the DNA substrate to E + DNA + dNTP. To measure this DNA dissociation rate (k_7), we used the nucleotide analog Sp-dTTPαS, to take advantage of the slower incorporation rate constant in order to increase the kinetic partitioning. A preincubated solution of Dpo4 (50 nM) and 5’-[32P]-D-1 (60 nM) was reacted with Sp-dTTPαS (100 µM) alone or with Sp-dTTPαS (100 µM) and a large molar excess of unlabeled D-1 trap (2.5 µM) to ensure the observation of a single binding event. In these time courses, the k_{obs} values were determined to be 18.6 s^{-1}, but the burst amplitudes were 44.2 ± 1.5 nM and 42.9 ± 1.7 nM respectively (Figure 6.11). Although small, the difference in the burst reaction amplitudes was confirmed by repeating these kinetic experiments. Since the relative amplitude (42.9/44.2) is equivalent to k_{obs}/(k_{obs} + k_{off}), the dissociation rate k_{off} (k_7) of the E•DNA•dNTP complex was estimated to be 0.58
s$^{-1}$. This rate was similar to the value reported at 37 °C, again indicating the observed low processivity of Dpo4 at 56 °C (data not shown) was due to fast DNA dissociation from the E•DNA•dNTP ternary complex.

**Thermal stability.** CD spectroscopy was performed to analyze the secondary structure stability of Dpo4 over a range of temperatures measured via the observation of the ellipticity at a fixed wavelength (222 nm) which corresponds to a region of the spectrum which detects the protein backbone conformation. Interestingly, the ellipticity slowly increases from 14 °C to 86 °C, where it then begins to increase sharply suggesting that the protein, or specific protein domains begin to thermally unfold (Figure 6.12). This correlates fairly well with the physiological temperature of the organism which varies between 75 °C and 85 °C. In addition, we performed a stability assay whereby we incubated solutions containing 55 nM Dpo4 at the following temperatures for ten minutes: 2 °C, 14 °C, 26 °C, 32 °C, 38 °C, 44 °C, 50 °C, 56 °C, 62 °C, 70 °C, 80 °C, 90 °C, 95 °C, and 100 °C. Subsequent to the incubation, 10 µl aliquots of the above solutions were then added to 5'-[³²P]-D-1 (50 nM) and following a ten minute incubation at 37 °C, were rapidly mixed with 1 mM dTTP for one minute. A plot of relative extension as a function of temperature (Figure 6.13) indicated that Dpo4 was stable over a range of temperatures from 2 °C to 90 °C, but then abruptly became inactivated at 95 °C. These observations correlate very well with results from our CD studies and strongly suggest that the decrease in activity at 95 °C was due to thermal denaturation of the structure of Dpo4. Although one may expect a decrease in stability of Dpo4 at 90 °C based on our CD results, we only observed a 5% decrease in relative product formation. This could be due to the slight molar excess of Dpo4 over DNA in this assay, but is also a function of the less significant conformational uncoupling of Dpo4 secondary structure observed in Figure 6.12 at 90 °C compared to 95 °C. However, when coupled with our observation of
a significant increase in $k_p$ with temperature, the increase in ellipticity between 14 °C and 86 °C suggests an increase in the dynamics of the conformation of Dpo4 as discussed below.

### 6.4 Discussion

Thermostable enzymes perform enzyme turnover at much higher rates as the temperature of the medium increases (21,28,30-37). These higher enzymatic activities are due to faster conformational dynamics of the enzyme at higher reaction temperatures (34,37) thought to be a function of a balanced relationship between molecular stability and structural flexibility (38). Although the nucleotide incorporation rate has not been quantitatively determined for Dpo4 at 80 °C, Dpo4 has been shown to remain active over a temperature range of 37 °C to 95 °C (39). We naturally expect that the microscopic rate constants (Scheme 6.1) will increase as the reaction temperature approaches the physiologically relevant temperature of Dpo4. Yet the pre-steady state kinetic parameters of single nucleotide incorporation catalyzed by Dpo4 on an undamaged DNA substrate have only been reported at 37 °C. Therefore, the effects of temperature on the fidelity and mechanism of Dpo4 catalysis requires elucidation.

*Fidelity.* We have incrementally determined the fidelity of nucleotide incorporation catalyzed by Dpo4 at several temperatures ranging from 26 °C to 56 °C using the same single turnover experiments described previously (16). We were able to extract the maximum rate constant of nucleotide incorporation ($k_p$) and ground-state binding affinity
\((K_d)\) for each nucleotide incorporation event into the undamaged DNA substrate, D-1 (Figure 6.1). These parameters along with the substrate specificities and overall fidelity were calculated and listed in Table 6.1. The resulting \(k_p^{\text{correct}}\) values (Figure 6.14) and average \(k_p^{\text{incorrect}}\) values (Figure 6.15) were then plotted as a function of temperature including those \(k_p\) values estimated at 2 °C and 14 °C. In addition, the average overall fidelity over this same temperature range was plotted (Figure 6.14). We found that with an increase in the reaction temperature of 30 °C (26 °C to 56 °C), the \(k_p\) for correct and incorrect nucleotides increase by 30-fold and 27- to 143-fold respectively (Table 6.1).

Similar results were observed for another thermophilic Y-family member Dbh, a homolog from \(S.\ acidocaldarius\), which demonstrated a 40-fold higher rate constant at 65 °C compared to 22 °C (31). Conversely, the differences in ground-state binding affinities (\(K_d\)) for both correct and incorrect nucleotides differ by only 2- to 3-fold over this temperature range, therefore not significantly altering the \(\Delta \Delta G\) value nor the conclusion we reported previously for Dpo4 in regard to its inability to discriminate between correct and incorrect nucleotides in the ground-state (16). Due to the lack of a significant change in the \(K_d\) with temperature coupled with the observation of a similar magnitude change of \(k_p\) for both correct and incorrect incorporations, the overall fidelity (Table 6.2) from 26 °C to 56 °C decreased by a very modest 2.7-fold (Figure 6.14). Interestingly, the error frequency of \(Taq\) DNA polymerase, a thermostable enzyme from \(Thermus aquaticus\), has been shown to increase only 2-fold over a temperature range of 15 °C, from 55 °C to 70 °C using a base substitution reversion assay (21). These differences could be a function of the experimental error inherent to these assays or due to the relative insensitivity of this assay, since a forward mutation assay for \(Taq\) polymerase showed no observable difference in the error frequency with the same increase in temperature (21). A similar magnitude (~2-fold) difference in error frequency was also observed for \(Thermococcus litoralis\) DNA polymerase with a temperature difference of 17 °C (40). Unfortunately,
our quest to perform analogous fidelity studies at temperatures exceeding 56 °C employing the same single turnover experiments are limited by the equipment, as described in Results, at temperatures approaching 65 °C to 70 °C. Thus we hypothesize that the fidelity of Dpo4 at 80 °C will not be significantly different than its fidelity determined at 56 °C in this report. Incidentally, a combination of steady-state kinetics and studies using a forward mutation assay reveal that the substitution error rate of Dpo4 with undamaged DNA (6.5 x 10^-3) at 70 °C (41) does not differ significantly from the misinsertion fidelity (10^-3 to 10^-4) estimated at 37 °C (39).

Conformational flexibility and mechanism. As the temperature of a reaction is increased, there is an ensuing exponential increase in the rate constant for catalysis until the enzyme structure succumbs to thermal denaturation which is thought to be brought about by local or global unfolding of the catalytically competent conformation (38). We have observed a similar exponential-like increase in the rate constant of Dpo4 for both correct and incorrect nucleotide incorporation when temperature is increased (Figures 6.14 and 6.15). In both cases this exponential increase is preceded by a small angle linear increase in the rate constant. In this linear phase, which occurs at the lower reaction temperatures, the rate constants remain similar in magnitude and are relatively slow presumably due to the lack of sufficient energy for a relatively large portion of Dpo4 molecules to overcome the activation energy ($E_a$) barrier to catalyze the reaction at these temperatures (38). This decrease in the ability of an enzyme performing the steps and conformational changes requisite for catalysis has been correlated with increased rigidity at these low nonphysiological temperatures (42,43). The data in Figures 6.14 and 6.15 indicate that the rate constants for both correct and incorrect incorporations slowly transition to an exponential phase at roughly 40 °C. However, the highest temperature that we have assayed for enzyme activity is 56 °C, which is roughly 20 °C to 30 °C below its
physiological temperature. This trend suggests that the rate constant for correct incorporation at its physiological temperature is on the order of several hundred per second. Moreover we observe a dramatic decrease of 18,900-fold in the rate constant for correct incorporation when the temperature decreases from 56 °C to 2 °C. Likewise, a 42 °C decrease in temperature decreases the rate constant for incorrect incorporation 5,000-fold and eliminates the observation of these events completely at 2 °C under our assay conditions. These dramatic decreases are not due to protein denaturation based on the results from our CD spectroscopic analysis and stability assay (Figure 6.12 and 6.13) but instead are a function of the rigid nature of Dpo4 at low nonphysiological temperatures. However, the increased conformational dynamics and decreased rigidity of Dpo4 at 56 °C did not change the minimal mechanism for single nucleotide incorporation (Scheme 6.1) which is analogous to that reported at 37 °C and indicates Dpo4 still utilizes an induced-fit mechanism for the selection of correct nucleotides. Free Dpo4 and DNA associate with a second order rate constant of 2.8 x 10⁶ M⁻¹s⁻¹ to form the Dpo4•DNA binary complex. Once competently bound, free dNTP diffuses into the Dpo4 active site at an association rate constant assumed to be equivalent to diffusion control. Observation of the biphasic nature for correct incorporation from the burst experiment indicates that the slowest step in the overall reaction scheme occurs subsequent to the chemistry step and is due to DNA dissociation from the enzyme since the slow linear phase is quantitatively equivalent to our steady-state results (Figures 6.4 and 6.5). Further kinetic interrogation of the burst phase revealed several lines of evidence suggesting that the incorporation of a correct nucleotide was limited by a protein conformational change that preceded the chemistry step. First, we failed to observe an elemental effect for the incorporation of a correct dNTP analog containing a sulfur substitution at the α-phosphate position. Although regarded as a controversial argument with respect to elucidation of the rate-limiting step in such mechanistic studies, it certainly remains plausible that the less
restrictive Dpo4 active site is less likely to succumb to the steric effects that compromise the reliability of this assay. Secondly, using steady-state kinetics, we determined a significantly slower DNA dissociation rate from a ternary complex formed in the presence of both dDTTP and dCTP (0.0028 ± 0.0004 s\(^{-1}\)) and a ternary complex only in the presence of dDTTP (0.063 ± 0.001 s\(^{-1}\)). This observation provides compelling evidence for the existence of two distinct ternary complexes with different affinities for DNA and suggests that the ternary complex in the presence of the next correct nucleotide (dCTP) undergoes some change in conformation to decrease its DNA dissociation rate constant while unable to perform the chemistry step. Finally, analysis of the reaction amplitudes of the pulse-quench/pulse-chase experiments have indicated that there exists a population of ternary complexes (E•DNA\(_n•\)dNTP) that can proceed to product under chase conditions yet are not observed when quenched in the former assay. This provides direct evidence for the slow formation of an enzyme-bound ternary complex preceding the chemistry step as has been also observed with T7 DNA polymerase (9), the Klenow fragment (10), HIV-RT (11), and pol\(\eta\) (12). The said conformational change was observed to occur at a rate constant of 189 s\(^{-1}\) at 56 °C and was 30-fold faster than the proposed rate-limiting conformational change measured at 37 °C. The argument for the nature of this rate-limiting conformational change was described previously (13) based on structural arguments and will not be discussed in detail here. In comparison to the rate constants gleaned at 37 °C, all the rate constants at 56 °C we determined to be faster (Table 6.2).

**Activation energy.** We would like to present additional evidence for the existence of a rate-limiting protein conformational change for Dpo4 using results from our dNTP concentration dependence studies described above. The \(k_p\) values extracted for the incorporation of a correct nucleotide at these varying temperatures were treated and fit
via the Arrhenius equation (equation 6) as shown in Figure 6.16 to yield an activation energy ($E_a$) of 32.9 kcal/mol. This $E_a$ for the incorporation of a correct nucleotide theoretically represents the energy barrier for either $k_3$, $k_4$ or a combination of both in Scheme 6.1 since this assignment is dependent upon which step in the first turnover is rate-limiting. Now it is well known that numerous, if not all DNA and RNA polymerases use a two-metal ion mechanism to catalyze the addition of a nascent nucleotide to a DNA or RNA substrate (44). Although the nucleophilic attack of the pentacovalent transition state can produce a completely associative, completely dissociative, or an intermediate semi-associative transition state complex, Herschlag et al. (29) and references therein in addition to recent crystal structure analyses of β-phosphoglucomutase (45) and a group I intron (46) have indicated the existence of a partially associative (semi-associative) transition state. Although modeled on an associative mechanism and under the assumption of a rate-limiting chemistry step, Florián et al. used computer simulation to conclude that for T7 DNA polymerase, a chemistry step involving the transfer of a proton to activate the 3’ hydroxyl nucleophile, accounts for an activation energy barrier of 12.3 kcal/mol (47). In addition, Radhakrishnan and Schlick used quantum mechanics/molecular mechanics dynamics simulations and quasi-harmonic free energy calculations to show that the rate-limiting chemistry step for correct incorporation catalyzed by DNA polymerase β occurred with a free energy of activation of 17 kcal/mol (48). Moreover, for uncatalyzed phosphodiester bond formation in solution, the $E_a$ for its rate-limiting chemistry step is estimated to be 21.1 kcal/mol (47). This value should be significantly lower when this rate-limiting reaction occurs in an enzyme active site based on Pauling’s transition state theory (49) due to the likely stabilization of the ionic transition state from favorable interactions conferred from the enzyme. Our calculated $E_a$ of 32.9 kcal/mol for the correct incorporation of dTTP into D-1 is considerably larger than that derived via these various simulation methods which each measured the $E_a$ of the
chemistry step and as such, suggests that $k_3$ (protein conformational change) rather than $k_4$ (chemistry step) is limiting the correct nucleotide incorporation catalyzed by Dpo4. Although the precise nature of the proposed rate-limiting conformational change is unknown, this conclusion seems reasonable because even local structural rearrangements within an enzyme active site will involve movements of many chemical bonds and should have a relatively large activation energy barrier. In addition, the large $E_a$ reported here suggests that the structure of Dpo4 is rigid at low temperature and becomes more dynamic at higher temperature (34,37,50), leading to increasingly higher $k_p$ values at more physiological temperatures (see above discussion). In addition, our results above gleaned an elemental effect of 5.0 for the incorrect incorporation of dGTP into D-1 which drove us to suggest that the chemistry step limited incorrect nucleotide incorporation. Interestingly, we calculated an $E_a$ of 24.2 kcal/mol for this incorrect incorporation which was very similar to that rate-limiting chemistry steps derived from simulation (see above). Thus we propose that calculation of $E_a$ is a new method of determining the rate-limiting step of such DNA polymerase-catalyzed mechanisms. However, there clearly needs to be more studies performed to prove the reliability of this approach.

*Effect of temperature on DNA lesions.* An interesting question arises when considering the effect of temperature on the physiological role of Dpo4. Here we have provided evidence that indicates the fidelity and mechanism for incorporation into an undamaged DNA substrate remains unchanged at 56 °C. But can this be translated into the lesion bypass responsibilities of Dpo4? We recently reported that the mutational profile for the bypass of an abasic site lesion catalyzed by Dpo4 was influenced by reaction temperature (51). Using a novel short oligonucleotide sequencing assay (SOSA) we were able to directly sequence four upstream and four downstream nucleotides in addition to nucleotides incorporated directly opposite the lesion. Interestingly we showed that the
undamaged upstream incorporations were unaffected by increasing the temperature from 37 °C to 75 °C while incorporation opposite and downstream of the abasic site were significantly affected. Specifically we observed a significant increase in the -1 frameshift fidelity directly opposite the lesion at 75 °C (51% of bypass products at 37 °C versus 13% at 75 °C) coupled with a simultaneous increase in dAMP and decrease in dCMP incorporation opposite the lesion. These results are consistent with a change in the reaction mechanism which at 37 °C partitions more favorably toward the lesion loop-out mechanism, while at 75 °C the mechanistic preference is for an intrahelical abasic site thus favoring dAMP incorporation, regardless of downstream incorporation events (51). These results suggested that at higher temperatures the equilibrium between the intrahelical and extrahelical abasic site favored an intrahelical lesion. This could be due to the preclusion of an extrahelical abasic site in the gap between the little finger and finger domain due to the increased thermal motion of Dpo4 at higher temperatures or due to a destabilization of possible base-stacking interactions that are required to extrude the abasic site. Yet, the overall frameshift fidelity decreased at 75 °C compared to 37 °C suggesting that the activation energy barriers for incorporation into nonconventional ternary complexes could be overcome at 75 °C with greater thermal energy and a more conformationally flexible Dpo4 active site (see above) that prevent such mechanisms at lower temperatures. Incidentally, results from our SOSA at 37 °C matched our pre-steady state kinetic results for abasic bypass also performed at 37 °C (51,52) verifying the integrity of the assay. In addition, another recent report showed that Dpo4 was more capable of bypassing a 10[S-(+)-trans-anti-[BP]-N2,dG lesion at 55 °C than at 37 °C (53). Although their quantitative results are based solely on the unreliable kinetic parameter $k_{cat}$ as opposed to substrate specificity ($k_{cat}/K_m$), they do show an increase in the ability to incorporate the “mismatched” nucleotides dATP, dGTP, and dTTP, with respect to incorporation of dCTP. Molecular dynamics simulations demonstrated that this change in
“preference” was due to a shift in the conformation of the lesion from \textit{anti} to \textit{syn}, a process possibly influenced by reaction temperature, which preferentially incorporated the three “mismatched” dNTPs over the “matched” dCTP. The direct implication of these results is that the conformational flexibility of Dpo4 does not impinge upon the faithfulness of Dpo4 to incorporate nucleotides into undamaged DNA at higher temperatures. Instead the conformational dynamics increases the incorporation rate significantly while establishing alternative pathways and mechanisms to bypass lesions that cannot be invoked by high fidelity replicative DNA polymerases.

In this paper we have probed the effects of reaction temperature on the fidelity and elementary steps of nucleotide incorporation into undamaged DNA catalyzed by a Y-family DNA polymerase. Our pre-steady state kinetic results reveal that Dpo4 uses an induced-fit mechanism to select correct nucleotides at 56 °C. We also have shown that the overall fidelity remains unchanged over a range of 30 °C, suggesting that the fidelity at 80 °C will also be in the range of $10^{-3}$ to $10^{-4}$. In addition, we propose a kinetic mechanism shown in Scheme 6.1 for the incorporation of a single nucleotide catalyzed by Dpo4 which is the same as proposed previously at 37 °C, but the microscopic rate constants are all faster than those reported previously.
6.5 Figures, Tables, Schemes

5′-CGCAGCCGTCCAACCAACTCA-3′

3′-GCGTCGGCAGGTTGGTTGAGTAGCAGCTAGGTTACGGCAGG-5′

Figure 6.1. The D-1 DNA substrate.
Figure 6.2. Concentration dependence on the pre-steady state rate of dTTP incorporation into the D-1 substrate at 56 °C. A preincubated solution of Dpo4 (120 nM) and 5'-32P-labeled-D-1 (30 nM) was mixed with increasing concentrations of dTTP (25 μM, ●; 50 μM, ○; 100 μM, ■; 200 μM, □; 400 μM, ▲; 800 μM, △; 1200 μM, ▼) for various time intervals. The solid lines are the best fits to the single exponential equation.
Figure 6.3. Concentration dependence on the pre-steady state rate of dTTP incorporation into the D-1 substrate at 56 °C. The single exponential rate constants ($k_{\text{obs}}$) were plotted as a function of dTTP concentration. The rate data were then fit via hyperbolic regression to yield a $k_p$ of $189 \pm 13 \text{ s}^{-1}$ and a $K_d$ of $383 \pm 65 \text{ M}$. 
Figure 6.4. Pre-steady state and steady-state kinetics of dTTP incorporation into D-1 at 56 °C. A preincubated solution of Dpo4 (30 nM) and 5'-32P-labeled-D-1 (120 nM) was mixed with dTTP (100 μM) for various times followed by quenching with 0.37 M EDTA. The data were fit by nonlinear regression to a biphasic equation (equation 3) with rate constants of 33.6 ± 5.0 s⁻¹ and 0.11 ± 0.01 s⁻¹ for the exponential and linear phases respectively.
Figure 6.5. Steady-state kinetics of dTTP incorporation into D-1 at 56 °C. dTTP incorporation into D-1 was independently measured under steady-state conditions by preincubating Dpo4 (2.4 nM) and 5'-32P-labeled-D-1 (250 nM) and then starting the reactions with the addition of dTTP (0.10 mM). The data were fit to equation 4 giving a steady-state rate constant of 0.23 s⁻¹.
Figure 6.6. Active site titration of Dpo4 at 56 °C. Dpo4 (30 nM) was preincubated with increasing concentrations of 5’-32P-labeled-D-1 and subsequently mixed with a solution containing dTTP. The reactions were quenched after 26 msec and the products were analyzed by sequencing gel electrophoresis. The resulting burst amplitudes were plotted as a function of substrate concentration and fit to the quadratic equation (equation 5) which gave a $K_d$ for the Dpo4•D-1 complex of 39.9 ± 6.3 nM.
Figure 6.7. Elemental effect on the rate constant for correct nucleotide incorporation at 56 °C. Dpo4 (120 nM) and 5'-32P-labeled-D-1 (30 nM) was mixed with either 100 µM dTTP (●) or Sp-dTTPαS (○) in parallel time courses. The data were fit using a single exponential equation (equation 1) yielding $k_{obs}$ values of 29.3 ± 2.1 s$^{-1}$ and 29.1 ± 1.9 s$^{-1}$ for dTTP and Sp-dTTPαS, respectively giving an elemental effect of 1.0 for correct nucleotide incorporation into D-1.
Figure 6.8. Elemental effect on the rate constant for incorrect nucleotide incorporation at 56 °C. Dpo4 (120 nM) and 5'-32P-labeled-D-1 (30 nM) was mixed with either 100 µM dGTP (●) or Sp-dGTPαS (○) in parallel time courses. The data were fit by nonlinear regression to the single exponential equation (equation 1) yielding $k_{obs}$ values of $0.12 \pm 0.01$ s$^{-1}$ and $0.024 \pm 0.001$ s$^{-1}$ for dGTP and Sp-dGTPαS, respectively giving an elemental effect of 5.0 for incorrect nucleotide incorporation into D-1.
Figure 6.9. Measurement of the DNA dissociation rate constant for the F′•DNA•dNTP complex at 56 °C. A preincubated solution of Dpo4 (2.4 nM) and 5′-32P-labeled-D-1 (250 nM) was mixed with ddTTP (1.2 mM) in the absence (●) or presence (○) of the next correct nucleotide dCTP (1.2 mM). The data were fit to equation 4 to give a steady-state rate constant in the absence and presence of dCTP of 0.063 ± 0.001 s⁻¹ and 0.0028 ± 0.0004 s⁻¹ respectively.
Figure 6.10. Pulse-chase and pulse-quench experiment at 56 °C. A preincubated solution of Dpo4 (30 nM) and D-1 (30 nM) was mixed with [α-32P]-dTTP (40 μM) for various times. The reactions were either quenched with 1 M HCl (pulse-quench) or chased with 2.5 mM unlabeled dTTP (pulse-chase) for 15 s before the HCl quench. Both experiments were fit to the burst equation yielding amplitudes of 26.9 ± 1.8 nM and 23.3 ± 1.3 nM, for the pulse-quench (⊙) and pulse-chase (●) experiments, respectively.
Figure 6.11. DNA dissociation from E•DNA•dNTP complex at 56 °C. A preincubated solution of Dpo4 (50 nM) and 5'-[32P]-D-1 (60 nM) was mixed with Sp-dTTPαS (100 µM) alone (●) or with Sp-dTTPαS (100 µM) and unlabeled D-1 trap (2.5 µM) (○). The $k_{obs}$ values were determined to be 18.6 s$^{-1}$ for both time courses, but the burst amplitudes were 44.2 ± 1.5 nM and 42.9 ± 1.7 nM respectively. This reduction in amplitude is equivalent to a dissociation rate of 0.58 s$^{-1}$. 
Figure 6.12. Thermal denaturation curve of Dpo4. The ellipticity of Dpo4 (30 µM) at 222 nm was monitored in the temperature range from 14 ºC to 100 ºC.
Figure 6.13. Dpo4 stability assay. A solution containing Dpo4 (55 nM) was incubated for ten minutes at a range of temperatures from 2 °C to 100 °C followed incubation with D-1 (50 nM) and then subsequent rapid mixing with dTTP (1 mM) at 37 °C for one minute to determine the thermal stability of Dpo4 as a function of temperature.
Figure 6.14. Temperature dependence of correct incorporation rate constant and fidelity on temperature. Plot of extracted $k_p$ values for correct nucleotide (●) and fidelity (○) as a function of temperature.
Figure 6.15. Temperature dependence of incorrect incorporation rate constant on temperature. The average values for incorrect nucleotide incorporation ($k_p$) were plotted as a function of temperature.
Figure 6.16. Activation energy for correct nucleotide incorporation. The extracted $k_p$ values were plotted as a function of reaction temperature according to equation 6 to yield an activation energy ($E_a$) of 32.9 kcal/mol.
Table 6.1. Kinetic parameters of Dpo4 for an undamaged D-1 substrate with varying temperature.
<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (μM$^{-1}$s$^{-1}$)</th>
<th>Fidelity$^a$</th>
<th>Average Fidelity$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>26 °C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>825</td>
<td>0.003</td>
<td>4.0 x 10$^{-6}$</td>
<td>1.9 x 10$^{-4}$</td>
<td>2.9 x 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>380</td>
<td>0.003</td>
<td>7.0 x 10$^{-6}$</td>
<td>3.3 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>1214</td>
<td>0.009</td>
<td>7.2 x 10$^{-6}$</td>
<td>3.4 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>304</td>
<td>6.4</td>
<td>2.1 x 10$^{-2}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>38 °C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>990</td>
<td>0.012</td>
<td>1.2 x 10$^{-5}$</td>
<td>1.5 x 10$^{-4}$</td>
<td>4.2 x 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>162</td>
<td>0.012</td>
<td>7.3 x 10$^{-5}$</td>
<td>9.1 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>1740</td>
<td>0.026</td>
<td>1.5 x 10$^{-6}$</td>
<td>1.9 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>221</td>
<td>17.7</td>
<td>8.0 x 10$^{-2}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>44 °C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>392</td>
<td>0.042</td>
<td>1.1 x 10$^{-4}$</td>
<td>9.2 x 10$^{-4}$</td>
<td>5.5 x 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>403</td>
<td>0.021</td>
<td>5.2 x 10$^{-4}$</td>
<td>4.4 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>1257</td>
<td>0.044</td>
<td>3.5 x 10$^{-4}$</td>
<td>2.9 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>483</td>
<td>59.7</td>
<td>0.12</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>50 °C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>502</td>
<td>0.13</td>
<td>2.6 x 10$^{-4}$</td>
<td>5.8 x 10$^{-4}$</td>
<td>4.4 x 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>517</td>
<td>0.064</td>
<td>1.2 x 10$^{-4}$</td>
<td>2.6 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>682</td>
<td>0.15</td>
<td>2.2 x 10$^{-4}$</td>
<td>4.8 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>258</td>
<td>118</td>
<td>0.45</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>56 °C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>1173</td>
<td>0.42</td>
<td>3.6 x 10$^{-4}$</td>
<td>7.4 x 10$^{-4}$</td>
<td>7.9 x 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>251</td>
<td>0.16</td>
<td>6.3 x 10$^{-4}$</td>
<td>1.3 x 10$^{-3}$</td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>1473</td>
<td>0.24</td>
<td>1.6 x 10$^{-4}$</td>
<td>3.3 x 10$^{-4}$</td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>383</td>
<td>189</td>
<td>0.49</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Calculated as \( (k_p/K_d)_{\text{incorrect}}/[(k_p/K_d)_{\text{correct}} + (k_p/K_d)_{\text{incorrect}}] \)

$^b$Calculated as \( \sum ((k_p/K_d)_{\text{incorrect}}/[(k_p/K_d)_{\text{correct}} + (k_p/K_d)_{\text{incorrect}}])^{dATP, dCTP, dGTP}/3 \)
<table>
<thead>
<tr>
<th>Parameters</th>
<th>Constants 37 ºC</th>
<th>Constants 56 ºC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$</td>
<td>1.9 μM$^{-1}$s$^{-1}$</td>
<td>2.8 μM$^{-1}$s$^{-1}$</td>
</tr>
<tr>
<td>$k_{-1}$</td>
<td>0.02 s$^{-1}$</td>
<td>0.10 s$^{-1}$</td>
</tr>
<tr>
<td>$K_{d, DNA}$</td>
<td>10.6 nM</td>
<td>39.9 nM</td>
</tr>
<tr>
<td>$k_2$</td>
<td>100 μM$^{-1}$s$^{-1}$</td>
<td>100 μM$^{-1}$s$^{-1}$</td>
</tr>
<tr>
<td>$k_{-2}$</td>
<td>23,000 s$^{-1}$</td>
<td>38,300 s$^{-1}$</td>
</tr>
<tr>
<td>$K_{d, dNTP}$</td>
<td>230 μM</td>
<td>383 μM</td>
</tr>
<tr>
<td>$k_3$</td>
<td>9.4 s$^{-1}$</td>
<td>189 s$^{-1}$</td>
</tr>
<tr>
<td>$k_7^*$</td>
<td>0.41 s$^{-1}$</td>
<td>0.58 s$^{-1}$</td>
</tr>
<tr>
<td>$k_8$</td>
<td>0.004 s$^{-1}$</td>
<td>0.0028 s$^{-1}$</td>
</tr>
</tbody>
</table>

*Results taken from (16) and (13)
*Results performed at different concentrations dTTP

Table 6.2. Kinetic Constants of Dpo4 at 37 ºC and 56 ºC.
Scheme 6.1. Proposed minimal kinetic mechanism for dNTP incorporation into DNA catalyzed by Dpo4.
6.6 References


7.1 Introduction

Since the discovery of *Escherichia coli* DNA polymerase I in 1950s, six families of DNA polymerases, which participate in diverse physiological roles, have been discovered in all three domains of life. A DNA polymerase, according to its orthodoxical definition, is an enzyme which catalyzes the formation of polynucleotides in a template-dependent manner in the presence of four deoxynucleotides (dNTPs). Although this ability to synthesize nucleotide polymers is affected by numerous factors including the overall mechanism, which is polymerase specific, the selection of a correct incoming dNTP is governed by proper Watson-Crick basepairing between the base of the dNTP and the template base opposite the incoming dNTP as well as geometric selection (1,2).

However, many archaeal, bacterial, eukaryotic, and viral DNA polymerases are also found to catalyze non-templated nucleotide additions to the 3’-termini of blunt-end DNA. These enzymes include thermostable polymerases (*Taq* DNA polymerase from *Thermus aquaticus, Tf*I DNA polymerase from *Thermus filiformis, Tth* DNA polymerase from *Thermus thermophilus*) (3), *E. coli* DNA polymerase I (4), DNA polymerase α from chick embryo (5), rat DNA polymerase β (5), yeast DNA polymerases I (5) and η (6),
and reverse transcriptases (RT) from avian myeloblastosis virus (5) and human immunodeficiency virus 1 (HIV-1) (7). Although the precise biological significance of non-templated blunt-end additions has not been resolved, studies of HIV-1 RT suggest the incorporation of additional bases before (-) strand DNA transfer contribute to the hypermutability of the HIV-1 genome (7). Interestingly, the non-templated blunt-end nucleotide incorporations catalyzed by most of the aforementioned enzymes are dominated by a single dATP incorporation. This activity has been cleverly exploited by Promega Corporation (pGEM-T easy vector system) and Invitrogen (TOPO TA PCR cloning system) in the development of TA cloning kits. In these kits, a thermostable DNA polymerase, like Taq DNA polymerase, is used to generate PCR products with a 3’-protruding base A, which forms a Watson-Crick basepair with the 5’-protruding base T of a vector to facilitate the ligation between the PCR product and vector. Surprisingly, however, the kinetic and structural basis for the non-templated blunt-end polymerization has yet to be established.

In this chapter, we used Sulfolobus solfataricus DNA polymerase IV (Dpo4) as a model enzyme to investigate the mechanistic basis for non-templated blunt-end additions. S. solfataricus is an aerobic crenarchaeon that grows optimally at 80 °C and pH 2 to 4 (8). Dpo4, a model lesion-bypass polymerase from the Y-family, has been well characterized kinetically (9,10), biochemically (11-18), and structurally (14-23) in the presence of either damaged or undamaged primer/template DNA substrates. The results of our pre-steady state kinetic studies demonstrated that the adenine of an incoming dATP stacked against the 5’-base of the opposite DNA strand in a blunt-end DNA substrate, leading to one predominant dATP incorporation. Using pyrene 5’-triphosphate (dPTP) as a dATP analog, we demonstrated that base stacking was the major factor in facilitating nucleotide
binding in the ground-state. These kinetic predictions were supported by the solved crystal structure of the ternary complex reported here of Dpo4•blunt-end DNA•ddATP.

7.2 Materials and methods

**DNA substrates.** All DNA oligonucleotides except those in X-1 (Table 7.1) were purchased from Integrated DNA Technologies and were purified by denaturing polyacrylamide gel electrophoresis. Their concentrations were determined by UV spectroscopy with calculated extinction coefficients. The elongating DNA strands were 5’-radiolabeled using Optikinase (USB) and [γ-32P]ATP (GE Healthcare) and annealed as described previously (24). The DNA substrate X-1 containing a BPDE-adenine adduct was synthesized and prepared as described previously (21).

**Pyrene nucleoside 5’-triphosphate.** dPTP was prepared from β-pyrene nucleoside (25) as previously described (26) with a slight modification. Following purification on a DEAE-cellulose column and prior to conversion to the sodium salt, the triphosphate was further purified by HPLC on a Thermo Hypersil AQUACIL C18 column (4.6 × 250 mm, 5 μM particle size) using a 25 min linear gradient of 1.6-36.6% acetonitrile in 0.1 M TEAB (triethylammonium bicarbonate) at a flow rate of 1 ml/min with the triphosphate eluting at 20 min. The purity and identity of the sodium salt of the triphosphate was confirmed by 31P NMR in D2O containing Tris-HCl (pH 7.5, 50 mM) and EDTA (2 mM): δ 22 (t), 10.6 (d), 7.6 (d). The concentration of dPTP was determined by UV spectroscopy with a molar extinction coefficient of 33,390 M⁻¹cm⁻¹.
Reaction buffer R. Buffer R contains the following components optimized previously (24): 50 mM HEPES-NaOH (pH = 7.5 at 37 °C), 5 mM MgCl₂, 50 mM NaCl, 5 mM DTT, 10% glycerol, 0.1 mM EDTA, and 0.1 mg/ml BSA. All concentrations reported in this paper refer to the concentration of components after mixing. All reactions, unless otherwise noted, were carried out at 37 °C.

Measurement of the $k_p$ and $K_d$ for nucleotide incorporation. A preincubated solution containing Dpo4 (120 nM) and DNA (30 nM) in buffer R was mixed with an increasing concentration of dNTP•Mg²⁺. Reactions were terminated at various times by the addition of 0.37 M EDTA. Product was separated from remaining substrate via sequencing gel electrophoresis (17% Acrylamide, 8 M urea) and quantitated using a Phosphorimager 445 SI (Molecular Dynamics). The time course of product formation was fit to equation 1 by nonlinear regression using Kaleidagraph (Synergy Software) for each concentration of dNTP•Mg²⁺ to yield an observed rate constant ($k_{obs}$). A represents the reaction amplitude. The abstracted values for $k_{obs}$ were then plotted as a function of the concentration of dNTP•Mg²⁺ and fit using equation 2 to give the $k_p$ and $K_d$. The substrate specificity ($k_p/K_d$) of this nucleotide incorporation was then calculated.

$$[\text{Product}] = A[1 - \exp(-k_{obs}t)] \quad (\text{equation 1})$$

$$k_{obs} = k_p[d\text{NTP}]/([d\text{NTP}] + K_d) \quad (\text{equation 2})$$

Gel mobility shift assay. The $K_d$ for the Dpo4•DNA binary complex was determined by titrating Dpo4 (7.5- to 200 nM) into a solution containing 5’-[^32P]DNA (100 nM) in buffer T which is identical to buffer R except its pH was 7.5 at 23 °C. After equilibrating for 15 minutes at 23 °C, the binary complex was then separated from free DNA using a 4.5% native polyacrylamide gel. The gel was dried and then quantitated with a PhosphorImager 445. The concentration of binary complex (Dpo4•DNA) formation was
plotted against the concentration of Dpo4 and fit to equation 3 to yield the $K_d$ value. $E_0$ and $D_0$ represent the active enzyme and total DNA concentrations, respectively.

$$[E\cdot DNA] = \frac{1}{2}(K_d + E_0 + D_0) - \frac{1}{2}[(K_d + E_0 + D_0)^2 - 4E_0D_0]^{1/2}$$

(equation 3)

**Crystallization and structure determination of a blunt-end DNA ternary complex.** Wild-type Dpo4 was overexpressed and purified as described (19). Dpo4 and X-1 (Table 7.1) were mixed at a 1:1.2 molar ratio with the final protein concentration of 8 mg/ml in the buffer of 20 mM HEPES (pH 7.0), 0.1 mM EDTA, 1mM DTT and 100 mM NaCl. With the addition of 1 mM ddATP, crystals were produced by the hanging drop method at 20 °C by using a precipitant solution of 100 mM HEPES (pH 7.0), 100 mM calcium acetate, 12% polyethylene glycol (PEG) 3350, and 2% glycerol. Crystals were transferred to a cryoprotectant solution containing the mother liquor with 25% PEG 3350 and 15% ethylene glycol and flash frozen in liquid nitrogen. Diffraction data were collected at -178 °C using a R axis IPII image plate mounted on an Rigaku RU 200 generator and processed by HKL (27). The structures were solved by molecular replacement with CNS (28) using the type I structure as a search model (19). The structure was iteratively adjusted with O (29) and refined using CNS and Refmac5 (30,31) with all residues in the allowed regions of the Ramachandran plot (Table 7.2).

**Coordinates.** The coordinates and structural factors of the ternary complex Dpo4•blunt-end X-1•ddATP have been deposited in the Protein Data Bank (accession code 2IMW).
7.3 Results

**Determination of equilibrium dissociation constant of Dpo4•blunt-end DNA.** To determine whether or not the binding affinity of Dpo4 to DNA was affected by the absence of an extruding 5’ template region, the gel-mobility shift assay was used to determine the equilibrium dissociation constant ($K_d$) of the binary complex of Dpo4 and B-1, a 5’-[32P] labeled blunt-end DNA substrate (Table 7.1). A solution containing a fixed concentration of B-1 was incubated with increasing concentrations of Dpo4 (Materials and Methods). After equilibration, the binary complex Dpo4•D-1 and free B-1 were separated using native polyacrylamide gel electrophoresis (Figure 7.1). Notably, each B-1 molecule containing two blunt-ends was not simultaneously bound by two molecules of Dpo4. This is likely due to the limited size of B-1 (21/21-mer) which only allowed one protein molecule to bind. The binary complex concentration was calculated and plotted against Dpo4 concentration and the data were fit to equation 1 (Materials and Methods) to yield a $K_d$ of 38 ± 3 nM (Figure 7.2). In parallel, a similar gel-mobility shift assay (data not shown) was employed to measure the $K_d$ of the binary complex of Dpo4 and D-1, a recessed primer/template DNA (Table 7.1). The Dpo4•D-1 complex dissociated with a $K_d$ of 10 ± 2 nM. Thus, Dpo4 binds ~4-fold weaker to blunt-end DNA. To compensate for the modestly weaker binding affinity of Dpo4•blunt-end DNA, we decided to measure blunt-end nucleotide incorporation efficiency under single-turnover conditions in which the enzyme was in molar excess over DNA in order to eliminate complications from multiple enzyme turnovers.

**Efficiency of nucleotide incorporation onto blunt-end DNA.** Initially, we tested whether each dNTP could be incorporated onto a blunt-end DNA by Dpo4. A time course of
individual dNTP incorporation onto B-1 was performed and the results from sequencing
gel analysis of the reactions showed that Dpo4 did incorporate each dNTP (Figures 7.3 to
7.6) but with different efficiencies. In addition, each dNTP could be inefficiently
incorporated a second time, however the 1st incorporation dominated.

To determine the preference of Dpo4 catalyzed nucleotide incorporation onto blunt-end
DNA, single-turnover experiments were designed to quantify the substrate specificity or
efficiency ($k_p/K_d$), by measurement of the maximum rate of incorporation ($k_p$) and
equilibrium dissociation binding constant ($K_d$) for the incorporation of each dNTP onto
B-1 (Materials and Methods). Typically, after titration of the Dpo4•B-1 binary complex
with an incoming nucleotide (e.g. dATP•Mg$^{2+}$), the reactions were stopped by addition of
EDTA, which inactivated Dpo4 by chelating its metal cofactor Mg$^{2+}$. The reaction
mixtures were analyzed via denaturing gel electrophoresis. Product concentrations were
plotted against reaction time and fit to equation 1 (Materials and Methods) to yield an
observed incorporation rate constant ($k_{obs}$) at each dATP•Mg$^{2+}$ concentration (data not
shown). The $k_{obs}$ values were then plotted against the concentration of dATP•Mg$^{2+}$ and
the data were fit to equation 2 (Materials and Methods) to yield $k_p$ and $K_d$ (Table 7.3).
With B-1, dATP was incorporated with a $k_p$ of $0.0035 \pm 0.0003 \text{ s}^{-1}$ and a $K_d$ of $571 \pm 132$
µM (Figure 7.9). The substrate specificity ($k_p/K_d$) was then calculated to be $6.1 \times 10^{-6} \mu\text{M}^{-1}\text{s}^{-1}$ (Table 7.3). Similar single-turnover experiments were carried out for the
incorporations of dCTP, dGTP, and dTTP onto B-1 (data not shown) and the kinetic
parameters were listed in Table 7.3. To evaluate which nucleotide was the most efficient
substrate, the efficiency ratio, $(k_p/K_d)_{\text{dNTP}}/(k_p/K_d)_{\text{dATP}}$, was then calculated and was in the
range of 0.02 to 1 (Table 7.3). Based on the efficiency ratios, Dpo4 selects nucleotides
for blunt-end addition in the order of dATP > dTTP > dCTP > dGTP, with dATP favored
by 5- to 50-fold over the other nucleotides. Unexpectedly, dTTP was incorporated only 5-
fold less efficiently than dATP, while dGTP had the smallest substrate specificity of all four incorporations. These results suggested that dATP was incorporated 79% of the time onto B-1 compared to 16%, 3%, and 2% for dTTP, dCTP, and dGTP, respectively. Interestingly, the substrate specificity for blunt-end dATP incorporation ($6.1 \times 10^{-6} \mu M^{-1} s^{-1}$) was almost identical to the incorporation efficiency of a mismatched dATP ($9.9 \times 10^{-6} \mu M^{-1} s^{-1}$, Table 7.3) into a recessed primer/template D-1 (Table 7.1). Moreover, the $K_d$ and $k_p$ values for the blunt-end dATP incorporation (Table 7.3) were also similar to the corresponding values of the mismatched dATP incorporation. Thus, the blunt-end dATP incorporation was kinetically indistinguishable from the “template directed” incorrect dATP incorporation.

However, the DNA sequence of the B-1 substrate facilitated the possibility of an incoming dATP forming a Watson-Crick basepair with the 5’-T of the opposite strand thus promoting its incorporation if the terminal basepair between the 5’-T and the 3’-A of the elongating strand “melted”. To examine this hypothesis, we measured the nucleotide incorporation efficiency with another blunt-end DNA substrate B-2 (Table 7.1) under identical single-turnover conditions. Notably, B-1 and B-2 are identical in sequence context, yet differ with respect to which oligonucleotide strand is 5’-radiolabeled. The kinetic results with B-2 (Table 7.3) demonstrated that Dpo4 had the same nucleotide selection hierarchy as B-1, and thus, incorporation of dATP onto B-2 was favored over all other nucleotide incorporations by a kinetically similar 10- to 33-fold. The incorporation probabilities onto B-2 were calculated to be 81%, 8%, 8%, and 2% for dATP, dTTP, dCTP, and dGTP, respectively. Moreover, the incorporation efficiency of dGTP, which was notably the lowest among all four dNTPs onto B-2 (Table 7.1), strongly invalidated any hypothesized influence from terminal basepair “melting” to blunt-end additions catalyzed by Dpo4.
Efficiency of pyrene nucleoside 5’-triphosphate incorporation onto blunt-end and recessed DNA. Watson-Crick basepairing between an incoming nucleotide and a template base observed with recessed primer/template DNA was shown to be irrelevant for nucleotide selection by Dpo4 when it catalyzed blunt-end additions. The kinetic preference for dATP incorporation onto both B-1 and B-2 was thus likely due to its favorable base-stacking with the 3’-terminal base of the elongating strand. To test this hypothesis we measured the incorporation efficiency of pyrene nucleoside 5’-triphosphate (dPTP, Scheme 7.1), a dNTP analog, onto B-1. dPTP was chosen as a probe because its large nonpolar base precludes hydrogen bonding with any natural bases, while possessing superior aromatic π-stacking capabilities relative to dATP (25,26). Figure 7.7 showed that Dpo4 only incorporated dPTP onto B-1 once. Using the same single-turnover assay described above, we then determined a $k_p$ of $(8.5 \pm 0.4) \times 10^{-3}$ s$^{-1}$ and strikingly, a low $K_d$ of $14 \pm 2$ µM (Figure 7.10) for the incorporation of dPTP onto B-1 (Table 7.3). The incorporation efficiency of dPTP incorporation onto B-1 was calculated to be $6.1 \times 10^{-4}$ µM$^{-1}$s$^{-1}$, which was 100-fold higher than dATP incorporation efficiency (Table 7.3).

To examine whether or not the dramatic kinetic difference between the incorporations of dPTP and dATP was affected by a template base, we determined the pre-steady state kinetic parameters for dPTP incorporation into D-1, a recessed primer/template DNA substrate (Table 7.1), and listed the kinetic parameters in Table 7.3. Overall, dPTP bound 20-fold tighter, was incorporated 9.5-fold faster, and possessed 200-fold higher incorporation efficiency than mismatched dATP into D-1 (Table 7.3). In contrast, dPTP was incorporated 21-fold less efficiently than a matched dTTP into D-1 although it
possessed 8-fold higher ground-state binding affinity. The reason for the decreased efficiency was due to a 165-fold slower rate constant for dPTP incorporation (Table 7.3).

Efficiency of the 2nd nucleotide incorporation onto blunt-end DNA. Figures 7.3 to 7.6 suggested that Dpo4, like other DNA polymerases, was inefficient at catalyzing the 2nd nucleotide incorporation onto blunt-end DNA. To confirm this observation, similar single-turnover experiments were performed (see above) to determine the incorporation efficiency onto B-1A (Table 7.1). B-1A contained a 3’-overhang dAMP and represented the DNA product after Dpo4 ‘incorporated’ the preferred nucleotide onto B-1. Single nucleotide incorporation assays with dCTP, dGTP, and dTTP only showed trace product formation onto B-1A even after 2 hours and were thereby considered kinetically irrelevant (data not shown). However, we were able to observe the incorporation of dATP onto B-1A (Figure 7.8) and determined the following kinetic parameters: $k_p = (3.6 \pm 0.2) \times 10^{-5} \text{ s}^{-1}$, $K_d = 464 \pm 72 \mu\text{M}$, and $k_p/K_d = 7.7 \times 10^{-8} \mu\text{M}^{-1}\text{s}^{-1}$ (Table 7.3). Interestingly, the $k_p/K_d$ value for the incorporation of dATP onto B-1A (Table 7.1) was 79-fold lower than the incorporation efficiency of dATP onto B-1, and was due to a 97-fold decrease in $k_p$, rather than a change in $K_d$. These results indicated that the 2nd dATP incorporation was significantly less efficient than the first dATP incorporation onto blunt-end DNA.

Crystal structure of ternary complex Dpo4•blunt-end DNA•dATP. Previously, we have published two ternary crystal structures of Dpo4, dATP and a 13/16-mer DNA substrate which contained a benzo[a]pyrene diol epoxide (BPDE)-adenine adduct at the last basepair between the template and primer strand (21). In these structures, Dpo4 functions as a typical template-dependent DNA polymerase, binding to the primer-template junction with the incoming dATP positioned opposite the templating dT. Interestingly, when BPDE-dA adduct was placed at the templating position in X-1 (Table 7.1), we
obtained crystals of a ternary complex, in which Dpo4 binds to the blunt-end of 13/17-mer with an incoming ddATP occupying the active site (Figures 7.11 and 7.12). This high resolution (2.05 Å) structure showed that the single-stranded portion of the 17-mer containing the BPDE-dA adduct was fully exposed to bulk solvent and thus disordered. The overall structure of Dpo4 (residues 1-348) was similar to the previously published type I structure (19) without significant protein structural changes with the DNA substrate remaining B-type. One minor difference was that seven additional residues at the C-terminus of Dpo4 were ordered and formed a flexible loop. The striking difference was that the base of the incoming ddATP was shifted toward the center of the helical duplex due to the absence of a templating base, and its six-membered ring stacked with the 5’-terminal base (dG) of the opposite strand. As a result of the ddATP shifting, its deoxyribose interacted with the 3’-terminal base (Figures 7.11 and 7.12) in addition to its interaction with amino acid residue Y12 as observed in other Dpo4 ternary complexes (19-22). In addition, only one divalent metal ion (Ca$^{2+}$) was found in the active site of Dpo4. The distance between the 3’-OH (nucleophile) and the α-phosphate of ddATP was 5.71 Å, which was beyond the 3.4 Å observed for an optimum catalytically active DNA polymerase ternary complex (32).

7.4 Discussion

Base-stacking governs blunt-end addition. Dpo4 is an archaeal Y-family DNA polymerase that can bypass a variety of DNA lesions including abasic sites, cisplatin-DNA adducts, cis-syn TT dimers, 6-4 TT dimers, acetyl aminofluorene-DNA adducts (AAF), and benzo[a]pyrene hydrocarbon adducts (11,18,21). Since it naturally lacks
proof-reading exonuclease activity, the fidelity of Dpo4 ($10^{-3}$-$10^{-4}$) with undamaged DNA at 37 °C (9) is much lower than the fidelity ($10^{-7}$-$10^{-8}$) of typical replicative DNA polymerases such as T7 phage DNA polymerase (33,34) and human mitochondrial DNA polymerase (35,36). Both DNA lesion tolerance and low fidelity suggest that Dpo4 may possess the template-independent nucleotidyl transferase activity which has been observed with other template-dependent DNA polymerases (3-5,7). This hypothesis was confirmed by the elongation of B-1 in the presence of each of four dNTPs (Figures 7.3 to 7.6).

The incorporation efficiency values in Table 7.3 confirmed that the most efficient substrate for the blunt-end addition onto both B-1 and B-2 catalyzed by Dpo4 was dATP, followed by dTTP, dCTP, and dGTP. Although B-1 and B-2 have identical DNA sequences, the nearest neighbor basepair for blunt-end additions, A:T in B-1 and G:C in B-2, were different, suggesting the selection of dATP for Dpo4-catalyzed blunt-end additions was sequence independent. dATP has also been found to be the most favorable nucleotide for blunt-end additions catalyzed by other template-dependent polymerases, although the selection order of the remaining dNTPs varies (3-5,7), e.g. dGTP > dTTP > dCTP with HIV-1 RT (7). Due to the lack of template information in B-1 and B-2 opposite an incoming nucleotide to form Watson-Crick basepairs, the ground-state binding affinity of this nucleotide was likely governed by its base-stacking with the blunt-end bases in addition to its van der Waals interactions with the active site residues of Dpo4 (A44, A57, and M76) (14,19). Previous thermodynamic studies of short DNA duplexes containing 5’-dangling ends by Guckian et al. (37) have determined the base-stacking free energies (kcal/mol) for the 5’-protruding base: A (1.0), G (0.7), T (0.6), and C (0.5). Other thermodynamic studies (38-40) have also revealed that adenine, regardless of 5’- or 3’-protuding, stacks better than the other three natural bases with DNA
duplexes. Thus, we conclude that base-stacking plays a major role in the selection of dATP by Dpo4 and other DNA polymerases for template-independent blunt-end additions. This conclusion was substantiated by the 41-fold higher ground-state binding affinity of the nucleotide analog dPTP (Scheme 7.1) to Dpo4•B-1 than dATP (Table 7.3) as well as the strong correlation between the order of base stacking energies (dPTP>dATP>dGTP>dTTP>dCTP) and the $k_p$, $K_d$, and $k_p/K_d$ with the exception of dGTP. The measured blunt-end base-stacking free energy of the pyrene base (1.7 kcal/mol) is significantly higher than adenine (1.0 kcal/mol) (37). The stronger stacking of dPTP, due to its four conjugated benzene rings (Scheme 7.1), over dATP with the terminal base(s) of B-1 significantly lowered the $K_d$ for the blunt-end incorporation of dPTP. In addition, we speculate that the increased stacking energies of each successive dNTP lead to a faster turnover rate constant ($k_p$) due to the inherent tighter binding of the particular dNTP in the ground-state that increases the probability of incorporation.

*Base-stacking controls $K_d$ while Watson-Crick hydrogen bonding governs $k_p*.* Strikingly, the strong stacking of dPTP also facilitated its binding to the binary complex of Dpo4 and recessed D-1 (Table 7.1). dPTP bound to Dpo4•D-1 in the ground-state with 8- and 20-fold higher affinity than matched dTTP and mismatched dATP, respectively (Table 7.3). The binding of dTTP likely involved Watson-Crick hydrogen bonds with the template base A, yet it was only about 2-fold tighter than the binding of mismatched dATP. The difference in the binding affinity of dPTP and dTTP further suggested that base-stacking was more critical to the ground-state binding affinity of an incoming nucleotide than hydrogen-bonding. In contrast, dPTP incorporation into D-1 was only slightly faster than typical mismatched dNTPs, but was 165-fold slower than the incorporation of matched dTTP into Dpo4•D-1. Thus, the Watson-Crick hydrogen bonds formed between an incoming nucleotide and a template base have a larger influence on the maximum
incorporation rate constant $k_p$ than base-stacking. This conclusion was supported by the fact that mismatched and matched dNTPs, which only differ in Watson-Crick hydrogen bonds, are incorporated with a larger difference in $k_p$ (100- to 1,000-fold) than in $K_d$ (1- to 100-fold) for all kinetically characterized DNA polymerases (9). The reason for this is due to the proper spatial orientation of the incoming dNTP with respect to the DNA 3’-OH guided by the Watson-Crick hydrogen bonds in a matched basepair. The tight binding of a nucleotide at a polymerase active site due to strong base-stacking interactions may not perfectly position this nucleotide for catalysis because it requires subtle adjustments by the Watson-Crick hydrogen bonds. However, both Watson-Crick hydrogen bonding and base-stacking mathematically contribute to the incorporation efficiency ($k_p/K_d$) of a DNA polymerase. There are many examples to support this general conclusion and importantly, one of these two parameters ($k_p$ or $K_d$) may have a larger influence on the substrate specificity than the other. With respect to the importance of hydrogen bonds, difluorotoluene nucleoside 5’-triphosphate, a nucleotide analog which cannot form Watson-Crick hydrogen bonds but has similar size, shape, and conformation as dTTP, is incorporated opposite a template base A with 220-fold and 10⁶-fold lower substrate specificity than dTTP catalyzed by yeast DNA polymerase η (41) and human DNA polymerase κ (42), respectively. However, as shown in Table 7.3, matched dTTP was incorporated 20-fold more efficiently than dPTP into D-1, suggesting that the Watson-Crick hydrogen bonds in the basepair dTTP:A contribute more significantly than the base-stacking interactions in the basepair dPTP:A to the efficiency of nucleotide incorporation catalyzed by Dpo4. The importance of hydrogen bonding shown here is similar to what has been revealed for incorporation catalyzed by human DNA polymerase β (X-Family polymerase) yet differs significantly from the lack of importance of hydrogen bonds for incorporation catalyzed by the A-Family DNA polymerase T7 phage DNA polymerase (43). Additionally, the fact that dPTP was more efficiently incorporated
into D-1 than mismatched dNTPs indicated that the large base of dPTP (surface area, 217 Å²) (37) did not perturb its incorporation efficiency and the active site of Dpo4 is flexible and/or spacious. Thus, both hydrogen bonding and base-stacking contribute disproportionally based on the specific catalyzed reaction.

Mechanism of blunt-end additions. The above discussion demonstrated that a polymerase selected dATP for blunt-end additions due to superior stacking with the terminal bases. What is the nature of dATP stacking? The ternary structure in Figures 7.11 and 7.12 showed that the six-membered aromatic ring of the base of ddATP surprisingly stacked against the six-membered aromatic ring of the 5’-G of the opposite strand while the sugar of ddATP stacked and possessed hydrophobic interactions with the 3’-terminal base C of the strand that will be extended. The vertical distance between the base planes of ddATP and 5’-G (3.4 Å) is equal to the average rise per basepair in B-type DNA, suggesting the base-stacking interactions were intense. However, in the case of dPTP incorporation, because the base of dPTP is so large, it likely base-stacked with both 3’-A and 5’-T of B-1 simultaneously, thus significantly increasing its binding affinity. The presence of an opposing template base (D-1) will likely limit the base-stacking of dPTP to the 3’-A only, possibly accounting for the observed 2-fold lower affinity with D-1 than with B-1 (Table 7.3). Thermodynamically, the base-stacking energy difference (0.7 kcal/mol) between 5’-dangling bases pyrene (1.7 kcal/mol) and adenine (1.0 kcal/mol) in a DNA duplex was not large enough to explain the binding free energy differences, \( \Delta \Delta G = -RT \ln \left( \frac{K_{d,PTP}}{K_{d,dATP}} \right) \), between dPTP and dATP with both B-1 (2.3 kcal/mol) and D-1 (1.8 kcal/mol). This indicated that other sources contributed to the higher stabilization of dPTP binding. One such source was favorable van der Waals interactions between pyrene and the active site residues (Ala43, Ala44, Ala57, Ala58, and Met76) of Dpo4 (14,19). The ternary structure shown in Figures 7.11 and 7.12 suggested that this possibility was
likely due to the large size of pyrene which is comparable to the total size of a canonical basepair (26). In addition, Figure 7.7 showed only one dPTP incorporation event. This was not surprising considering that consecutive dPTP incorporations into recessed DNA has yet to be observed in any DNA polymerase, e.g. yeast DNA polymerase η (6). This might be due to the lack of minor groove contacts between incorporated dPMP and polymerase active site residues (44), or more likely, due to the formation of a tightly bound dead-end ternary complex resulting from exceptional stacking of two consecutive pyrene bases (Figure 7.13). In addition, it is plausible that if the mechanism of blunt-end addition for the second dNTP incorporation strictly follows the ‘base-flipped’ mechanism as illustrated in Figure 7.14, the exceptional base stacking ability of the first incorporated dPMP would inhibit base-flipping and thus a second consecutive incorporation of dPTP.

For the second dATP incorporation onto B-1, Table 7.3 showed that the $K_d$ value was very similar to the ground-state nucleotide binding affinity for the first dATP blunt-end addition. This indicated that a similar base-stacking pattern as observed for the first dATP possibly stabilized the binding of the second dATP. In order to maintain such a stacking scheme between the second dATP and the blunt-end basepair, the first incorporated adenine on the 3’-terminus of the elongating strand must become extrahelical to make the terminal blunt-end basepair accessible. The base flipping phenomenon has been observed previously in the ternary structures of truncated human DNA polymerase λ (45) and Dpo4 (Ab-4B) (22). Consequently, the 3’-OH of the extrahelical unpaired 3’-A will be improperly positioned for efficient in-line nucleophilic attack of the α-phosphate of the incoming dNTP. This hypothesis was consistent with the very low $k_p$ value for dATP incorporation onto B-1A, which was 97-fold slower than the incorporation of dATP onto B-1 (Table 7.3), and the dominant first blunt-end addition of a single dNTP (Figures 7.3 to 7.8). For subsequent additions after the second blunt-end dATP incorporation, more
difficult active site adjustments of the DNA substrate were likely required, either precluding or rendering these incorporations extremely rare (Figures 7.3 to 7.8). Yet, one cannot rule out the possibility that the second incoming dATP simply stacked with the 3’-A without generation of an extrahelical nucleotide. This was due to similar $K_d$ values for the incorporation of dATP onto the blunt-end substrates (B-1 and B-2), the recessed substrate (D-1), and the 3’-protruding substrate (B-1A) which overall only differ from each other by 1.3-fold. Although the ground-state binding was not significantly perturbed, the lack of template information to help orient subsequent incoming nucleotides coupled with competition between these two such mechanisms may lead to the slow observed turnover.

However, the mechanisms in Figures 7.13 and 7.14 do not unambiguously explain the slow rate constants of blunt-end additions. Table 7.3 has shown that the $k_p$ values for the blunt-end incorporation of dATP and dPTP are two to three orders of magnitude slower than those of matched incorporations into D-1. This may be due to the distance between the 3’-OH of the elongating blunt-end DNA substrate and the $\alpha$-phosphate of the nucleotide (5.71 Å, Figures 7.11 and 7.12) which is longer than the distance (3.4 Å) observed for correct template-dependent nucleotide incorporation (32). With a recessed DNA substrate, our previous kinetic studies have concluded that the chemistry step is rate-limiting for incorrect nucleotide incorporation while the protein conformational change limits correct nucleotide incorporation (10). Thus, it is plausible that blunt-end additions, like misincorporations into recessed DNA, are also limited by phosphodiester bond formation. However, our kinetic and structural data cannot exclude the possibility that a local structural rearrangement at the Dpo4 active site to reorient the bound incoming nucleotide with respect to the blunt-end DNA, especially after the binding of the second metal ion, could limit blunt-end additions.
Ternary crystal structure of blunt-end complex. The crystal structure showed the incoming ddATP was mainly stabilized by aromatic stacking. However, instead of stacking with the elongating strand, the incoming ddATP was shifted toward the center of the DNA helix by ~2 Å when compared to the position of ddADP in the type I complex (19) (Figure 7.15). In this configuration, the adenine of ddATP was extensively stacked against the 5’-terminal base of the opposite strand while its deoxyribose made van der Waals interactions with the same strand 3’-terminal base. In addition, the stubby finger domain of Dpo4 established very few van der Waals contacts with ddATP, which were between the adenine and the small hydrophobic residues A44, A57, M76, and the main chain atoms of A43 and G58 (Figure 7.11 and 7.12). The interactions with A43 and G58 were not observed in Figure 7.15 (a), due to the inward movement of the incoming nucleotide in Figure 7.15 (b). The enhanced stacking and van der Waals interactions compensated for the lack of Watson-Crick basepairing and accounted for the similar $K_d$ with blunt-end B-1 and the recessed D-1 (Table 7.3). However, the significant decrease in the $k_p$ was most likely due to the lack of Watson-Crick hydrogen bonds that would normally help properly orient the incoming nucleotide in the active site of Dpo4 for catalysis. This effect on $k_p$ can be seen with each subsequent incorporation onto B-1 (Table 7.3). Clearly the optimized stacking of ddATP with blunt-end DNA was achieved at the expense of misalignment of DNA, dNTP, and metal ions in the active site, leading to a greatly reduced $k_p$. In fact, one would imagine that the conformation trapped in the crystal structure, which we assume to be the minimal energy state, was not suitable for the nucleotide incorporation. This is most likely the reason for the lack of additional dNTP and dPTP incorporation.
In the ternary structure with a T/G mismatch in the Dpo4 active site (14), the incoming nucleotide dGTP shifts towards the minor groove of the DNA duplex while the template base T wobbles towards the major groove due to lack of Watson-Crick hydrogen bonding (Figure 7.15 (c)). The shifting in T/G mismatch is thus different from the center-inward movement observed in Figure 7.15 (b) and the dGTP does not stack with the template base. This shifting led to extremely slow misincorporation, even in the case of the wobble basepair dGTP:T. In comparison, a similar center-inward movement was observed previously with Dpo4 and a DNA substrate containing an abasic site (22). Here, the primer strand, containing a 3’-terminal dAMP, tilted toward the gap in the template strand left by the abasic lesion, yet the 3’-OH was still within 5 Å of the α-phosphate of the incoming nucleotide (22). This structure and Figure 7.15 (c) are clearly different from the unprecedented base-stacking scheme in Figure 7.15 (b).

In summary, our kinetic and structural studies have established a mechanism for blunt-end additions catalyzed by Dpo4 and other DNA polymerases. We also revealed an unprecedented base-stacking pattern between an incoming nucleotide and the entire blunt-end basepair which stabilized the binding of this nucleotide. Our results should encourage and facilitate the investigation of any evolutionary advantage for retroviruses caused by the blunt-end addition activity of their reverse transcriptases.
7.5 Figures, Tables, Schemes

Figure 7.1. Gel mobility shift assay to determine Dpo4•B-1 equilibrium dissociation constant. Reactions containing 100 nM 5'-[^32P]B-1 were incubated with increasing concentration of Dpo4 as indicated in the gel picture followed by native gel analysis to resolve binary complex from unbound B-1.
Figure 7.2. Gel mobility shift assay to determine Dpo4•B-1 equilibrium dissociation constant. Binary complex formation [Dpo4•B-1] was plotted against Dpo4 concentration and the resulting data points were fit via quadratic regression (equation 3) to yield a $K_d$ of $38 \pm 3$ nM.
Figure 7.3. Series of polyacrylamide gel pictures showing time courses of product formation with 200 µM dATP onto B-1. Remaining primer, 21-mer, is shown at the bottom of each gel image with extension products located sequentially above it. Reaction time (minutes) is denoted below the corresponding lane.
Figure 7.4. Series of polyacrylamide gel pictures showing time courses of product formation with 1300 µM dCTP onto B-1. Remaining primer, 21-mer, is shown at the bottom of each gel image with extension products located sequentially above it. Reaction time (minutes) is denoted below the corresponding lane.
Figure 7.5. Series of polyacrylamide gel pictures showing time courses of product formation with 1300 μM dGTP onto B-1. Remaining primer, 21-mer, is shown at the bottom of each gel image with extension products located sequentially above it. Reaction time (minutes) is denoted below the corresponding lane.
Figure 7.6. Series of polyacrylamide gel pictures showing time courses of product formation with 400 µM dTTP onto B-1. Remaining primer, 21-mer, is shown at the bottom of each gel image with extension products located sequentially above it. Reaction time (minutes) is denoted below the corresponding lane.
Figure 7.7. Series of polyacrylamide gel pictures showing time courses of product formation with 2 µM dPTP onto B-1. Remaining primer, 21-mer, is shown at the bottom of each gel image with extension products located sequentially above it. Reaction time (minutes) is denoted below the corresponding lane.
Figure 7.8. Series of polyacrylamide gel pictures showing time courses of product formation with 1300 µM dATP onto B-1A. Remaining primer, 22-mer, is shown at the bottom of each gel image with extension products located sequentially above it. Reaction time (minutes) is denoted below the corresponding lane.
Figure 7.9. Concentration dependence on the rate of dATP incorporation onto B-1. A preincubated solution of Dpo4 (120 nM) and 5'-[\(^{32}\)P]B-1 (30 nM) was mixed with an increasing concentration of dATP\(\cdot\)Mg\(^{2+}\) for various times. The individual reactions were then quenched by 0.37 M EDTA. The single exponential rates for each individual time course were plotted as a function of dATP concentration. The rate data were then fit to the hyperbolic equation (Eq. 2) yielding a \(k_p\) of 0.0035 ± 0.0003 s\(^{-1}\) and a \(K_d\) of 571 ± 132 μM.
Figure 7.10. Concentration dependence on the rate of dPTP incorporation onto B-1. A preincubated solution of Dpo4 (120 nM) and 5'-[^32P]B-1 (30 nM) was mixed with an increasing concentration of dPTP•Mg²⁺ for various times. The individual reactions were then quenched by 0.37 M EDTA. The single exponential rates for each individual time course were plotted as a function of dPTP concentration. The rate data were then fit to the hyperbolic equation (Eq. 2) yielding a $k_p$ of 0.0085 ± 0.0004 s⁻¹ and a $K_d$ of 14 ± 2 μM.
Figure 7.11. Crystal structure of Dpo4•blunt-end X-1•ddATP (2.05 Å). (a) Overall ternary structure. Dpo4 was shown in grey ribbons while DNA and ddATP were shown as ball-and-stick models. The ddATP is highlighted in magenta. The Ca$^{2+}$ ion was shown in a green sphere. (b) The zoomed active site including ddATP and the blunt-end basepair. The residues in contact with ddATP were shown as ball-and-stick models (grey for atom C, red for atom O, yellow for atom S). Only the side chain and main chain atoms involved were shown.
Figure 7.12. Crystal structure of Dpo4•blunt-end X-1•ddATP (2.05 Å). 2Fo - Fc electron density map contoured at 1.2 σ (light-blue) of the active site. The amino acid residues, two blunt-end DNA basepairs, and incoming ddATP were shown as ball-and-stick models.
Figure 7.13. Proposed mechanisms for blunt-end additions of dPTP. dPTP is represented by P in different colors, respectively. The Watson-Crick hydrogen bonds were drawn as dashed lines while the base-stacking interactions were shadowed in green. The stacking interactions between the 2’-deoxyribose (R) of an incoming nucleotide and the 5’-terminal base A are displayed in a green box. The van der Waals interactions between an incoming nucleotide and Dpo4 active site residues were not shown for clarity.
Figure 7.14. Proposed mechanisms for blunt-end additions of dATP. dATP is represented by A in different colors, respectively. The Watson-Crick hydrogen bonds were drawn as dashed lines while the base-stacking interactions were shadowed in green. The stacking interactions between the 2'-deoxyribose (R) of an incoming nucleotide and the 5'-terminal base A are displayed in a green box. The van der Waals interactions between an incoming nucleotide and Dpo4 active site residues were not shown for clarity.
Figure 7.15. Comparison of the incoming nucleotide (magenta) positions in the active sites of three different Dpo4 ternary structures. The nucleotides were shown in the side (upper panel) and top (lower panel) views. (a) Type I (ddADP:T) (19); (b) Blunt-end (ddATP:DNA); and (c) Mismatched T/G-1 (dGTP:T) (14).
Table 7.1. DNA substrates.

<table>
<thead>
<tr>
<th>DNA Substrate</th>
<th>5’-Sequence</th>
<th>3’-Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-1</td>
<td>5´-CGCAGCCGTCCAACCAACTCA-3´</td>
<td>3´-GCGTCGGCAGGTTGTTGAGT-5´</td>
</tr>
<tr>
<td>B-2</td>
<td>5´-TGAGTTGGTTGGACGGCTGCG-3´</td>
<td>3´-ACTCAACCAACCTGCGACGC-5´</td>
</tr>
<tr>
<td>B-1A</td>
<td>5´-CGCAGCGTCCAACCAACTCAA-3´</td>
<td>3´-GCGTCGGCAGGTTGTTGAGT-5´</td>
</tr>
<tr>
<td>D-1</td>
<td>5´-CGCAGCGTCCAACCAACTCA-3´</td>
<td>3´-GCGTCGGCAGGTTGAGTAGCAGCTAGGTTACGGCAGG-5´</td>
</tr>
<tr>
<td>X-1</td>
<td>5´-TCT_AGATCCTCTCCCC-3´</td>
<td>3´-CTTAGGAAGGGGG-5´</td>
</tr>
</tbody>
</table>

All DNA substrates except X-1 contained a 5´-radiolabeled top strand. The underlined/italicized A is a BPDE-adenine adduct.
Space group \( \text{P2}_1\text{2}_1\text{2} \)
Unit cell \((a,b,c)\) (Å) \(98.1\ 101.9\ 52.4\)
No. of complex in AU 1
Non-hydrogen atoms 3697
Resolution range\(^a\) 26.7-2.05 Å (2.09-2.05 Å)
\(R_{\text{merge}}\)\(^a, b\) 0.063 (0.340)
Unique reflection 31879
Completeness (%)\(^b\) 97.5 (93.7)
\(R\)-value\(^c\) 0.230
\(R_{\text{free}}\)\(^d\) 0.284 (954 reflections)
r.m.s.d. bond length (Å) 0.011
r.m.s.d. bond angle (°) 1.30
Ave. B-value (Wilson) (Å\(^2\)) 45.3 (47.3)

\(^a\) Data completeness in the highest resolution shell is shown in parenthesis.

\(^b\) \(R_{\text{merge}} = \sum_h \sum_i |I_{hi} - <I_h>| / \sum <I_h>\), where \(I_{hi}\) is the intensity of the \(i\)th observation of reflection \(h\), and \(<I_h>\) is the average intensity of redundant measurements of the \(h\) reflections.

\(^c\) \(R\)-value = \(\sum |F_o| - |F_c| | / \sum |F_o|\), where \(F_o\) and \(F_c\) are the observed and calculated structure-factor amplitudes.

\(^d\) \(R_{\text{free}}\) is monitored with the reflections excluded from refinement in parenthesis.

Table 7.2. Summary of X-ray crystallographic data.
<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (µM)</th>
<th>$k_p$ (s⁻¹)</th>
<th>$k_p/K_d$ (µM⁻¹s⁻¹)</th>
<th>Efficiency Ratio a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporation onto B-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>571 ± 132</td>
<td>0.0035 ± 0.0003</td>
<td>6.1 x 10⁻⁶</td>
<td>1</td>
</tr>
<tr>
<td>dCTP</td>
<td>1581 ± 344</td>
<td>0.00042 ± 0.00005</td>
<td>2.7 x 10⁻⁷</td>
<td>0.04</td>
</tr>
<tr>
<td>dGTP</td>
<td>1983 ± 508</td>
<td>0.00026 ± 0.00004</td>
<td>1.3 x 10⁻⁷</td>
<td>0.02</td>
</tr>
<tr>
<td>dTTP</td>
<td>506 ± 85</td>
<td>0.00062 ± 0.00004</td>
<td>1.2 x 10⁻⁶</td>
<td>0.20</td>
</tr>
<tr>
<td>dPTP</td>
<td>14 ± 2</td>
<td>0.0085 ± 0.0004</td>
<td>6.1 x 10⁻⁴</td>
<td>100</td>
</tr>
<tr>
<td>Incorporation onto B-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>462 ± 98</td>
<td>0.0062 ± 0.0005</td>
<td>1.4 x 10⁻⁵</td>
<td>1</td>
</tr>
<tr>
<td>dCTP</td>
<td>744 ± 344</td>
<td>0.0009 ± 0.0001</td>
<td>1.3 x 10⁻⁶</td>
<td>0.10</td>
</tr>
<tr>
<td>dGTP</td>
<td>894 ± 134</td>
<td>0.00031 ± 0.00002</td>
<td>3.5 x 10⁻⁷</td>
<td>0.03</td>
</tr>
<tr>
<td>dTTP</td>
<td>835 ± 142</td>
<td>0.0011 ± 0.0001</td>
<td>1.4 x 10⁻⁶</td>
<td>0.10</td>
</tr>
<tr>
<td>Incorporation onto B-1A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>464 ± 72</td>
<td>(3.6 ± 0.2) x 10⁻⁵</td>
<td>7.7 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>Incorporation into D-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP b</td>
<td>578 ± 188</td>
<td>0.006 ± 0.001</td>
<td>9.9 x 10⁻⁶</td>
<td>2.4 x 10⁻⁴</td>
</tr>
<tr>
<td>dCTP b</td>
<td>1036 ± 95</td>
<td>0.013 ± 0.001</td>
<td>1.3 x 10⁻⁵</td>
<td>3.1 x 10⁻⁴</td>
</tr>
<tr>
<td>dGTP b</td>
<td>1150 ± 312</td>
<td>0.007 ± 0.001</td>
<td>6.0 x 10⁻⁶</td>
<td>1.5 x 10⁻⁴</td>
</tr>
<tr>
<td>dTTP b</td>
<td>230 ± 17</td>
<td>9.4 ± 0.3</td>
<td>4.1 x 10⁻²</td>
<td>1</td>
</tr>
<tr>
<td>dPTP</td>
<td>29 ± 3</td>
<td>0.057 ± 0.003</td>
<td>2.0 x 10⁻³</td>
<td>0.048</td>
</tr>
</tbody>
</table>

aEfficiency ratio = ($k_p/K_d$)dNTP/($k_p/K_d$)dATP with B-1 and B-2 or ($k_p/K_d$)dNTP/($k_p/K_d$)dTTP with D-1

bData are from reference (9)

Table 7.3. Comparison of nucleotide incorporation efficiencies for blunt-end DNA (B-1 and B-2) and recessed primer/template DNA (D-1).
Scheme 7.1. Structure of dPTP.
7.6 References


CHAPTER 8

PRE-STEADY STATE KINETIC STUDIES OF THE FIDELITY AND
MECHANISM OF POLYMERIZATION CATALYZED BY TRUNCATED DNA
POLYMERASE LAMBDA AT 37 ºC

8.1 Introduction

The base-excision repair (BER) pathway is one of the major mechanisms that removes
damaged base residues in DNA (1). This mechanism involves the excision of modified
DNA bases by DNA glycosylases, leaving noncoding apurinic/apyrimidinic (AP) sites in
DNA. These lesions are further processed and repaired by 5’-acting AP endonucleases,
5’-deoxyribose-5-phosphate lyases (dRPases), DNA polymerases, and DNA ligases (2-
5). In mammalian systems, the role of DNA polymerase β (Polβ) in BER has been well
established (3). Polβ has two independent domains, an N-terminal dRPase domain (8
kDa) and a C-terminal polymerase domain (31 kDa) (Figure 8.1) (6,7). The dRPase
activity removes the 5’-deoxyribose phosphate moiety (6,8) while the polymerase
domain, which has the general right-hand fold common to all replicative DNA
polymerases defined by subdomains named fingers, palm, and thumb (9), catalyzes gap-
filling synthesis in BER (6).

The recently discovered DNA polymerase λ (Polλ) along with Polβ, are members of the
X-family DNA polymerases and share 33% sequence identity (10-13). Biochemical
analysis has demonstrated that Polλ possesses intrinsic dRPase (14) and template-dependant DNA polymerase activities, but lacks 3’→5’ exonuclease activity (Figure 8.1) (10-12). The processivity of Polλ is low with normal template/primer DNA but is relatively high in short gaps having a 5’-phosphate group (13). Sequence alignment and three-dimensional structural modeling predict that the Polλ core contains the four conserved subdomains present in Polβ (11). The NMR structure of the dRPase domain solved recently displays a high degree of similarity with the corresponding domain in Polβ (15). Unlike Polβ, the N-terminal 132 amino acid residues of Polλ form a nuclear localization signal motif and a BRCT domain (Figure 8.1). BRCT domains are known to mediate protein/protein and protein/DNA interactions in DNA repair mechanisms or cell cycle check point regulation upon DNA damage (16). A proline-rich region (residues 133-244) links the BRCT domain to the dRPase domain. Truncated human Polλ (tPolλ) lacking the N-terminal 244 residues (Figure 8.1) has a 2.9- and 17.7-fold higher DNA polymerization activities than the full-length protein in the presence of Mg\(^{2+}\) and Mn\(^{2+}\), respectively (17). Analysis of deletion mutants suggests that the proline-rich domain functionally suppresses the polymerase activity (17). In the presence of Mn\(^{2+}\), both full-length Polλ and tPolλ display terminal deoxyribonucleotidyl transferase activity with sequence preference to pyrimidine nucleotides (18). Moreover, Polλ is shown to bypass an abasic site (19,20).

The biological function of Polλ is not clear yet although Polλ\(^{-/-}\) mice display hydrocephalus, situs inversus, chronic sinusitis, and male infertility (21). The gene encoding Polλ is mapped to mouse chromosome 19 and shown to be expressed at high level in the developing mouse testis, suggesting a possible function of Polλ in DNA repair synthesis associated with meiosis (11). It has been reported that proliferating cell nuclear antigen (PCNA) interacts with Polλ, increasing the processivity of Polλ in DNA
synthesis without affecting either the rate of nucleotide incorporation or discrimination efficiency (20). These results suggest Pol λ may be involved in the PCNA-dependent “long patch” BER pathway (19). In addition to the potential role in BER, human DNA Pol λ is recently shown to generate single-base deletions at average rates substantially higher than its base substitution rates (22). The high deletion frequency may rule out a significant role for Pol λ in translesion synthesis and in somatic hypermutation, but may suggest that Pol λ is involved in repair of double-stranded breaks (DSBs) through non-homologous end-joining (NHEJ) pathways. This hypothesis is supported by the results from immunodepletion studies suggesting that Pol λ, rather than other family-X polymerases, is primarily responsible for the gap-filling associated with NHEJ in human nuclear extracts (23).

The fidelity of Pol λ has been estimated to be (1.3 - 9.0) x 10^-4 by an M13mp2 forward mutation assay (13,19,22) and an M13mp2 reversion system (13,19,22). However, these mutation assays cannot provide quantitative kinetic constants to establish the mechanism of fidelity and the structure-fidelity relationship (24). In this chapter, we use pre-steady-state kinetic methods to measure the fidelity (10^-2 to 10^-4) of tPol λ based on all 16 possible nucleotide incorporations into single-nucleotide gapped DNA substrates. The reason we used tPol λ rather than the full-length protein was due to strong dNTP substrate inhibition on nucleotide incorporation by the full-length Pol λ (13,17). Our results showed that tPol λ catalyzed the T:dGMP misincorporation with substantially higher rates than all other mismatches. We also established a minimal mechanism of DNA polymerization catalyzed by tPol λ. The potential in vivo role of Pol λ was discussed.
8.2 Materials and methods

*Materials.* These chemicals were purchased from the following companies: $[^{32}\text{P}]\text{ATP}$, Perkin Elmer Life Sciences (Boston, MA); Biospin columns, Bio-Rad Laboratories (Hercules, CA); calf intestine alkaline phosphatase, Fermentas (Hanover, MD); dNTPs, Gibco-BRL (Rockville, MD); *Pfu* turbo, Stratagene (La Jolla, CA); T4 polynucleotide kinase, USB (Cleveland, OH).

*Cloning and purification of N-terminal truncated polymerase $\lambda$.* The human gene encoding tPol$\lambda$ (residues 245-575) was PCR amplified from a plasmid pET28b-Pol$\lambda$ encoding the full-length Pol$\lambda$ (K. A. Fiala and Z. Suo, unpublished results) using the following primers: betalike.for 5'--GCGTCCATATGTCAAGCCAGAAGGCGACCAATC-3' and betalike.rev 5'--GCAATTCTCGAGTCACCAGTCCCGCTCAG-3'. The resulting PCR product was cloned into the Nde I/Xho I sites of pET24b to construct pET24b-tPol$\lambda$. The constructed plasmid pET24b-tPol$\lambda$ was transformed into *E. coli* strain BL21(DE3) (Stratagene) to express tPol$\lambda$ fused to a C-terminal His$_6$ tag. Transformed *E. coli* cells were grown at 37 °C in the presence of 40 $\mu$g/ml kanamycin and 50 $\mu$g/ml chloramphenicol until OD$_{600}$ reached 0.5. Then the cultures were induced with 0.4 mM IPTG and incubated at 22 °C for 7 hours. Cells were harvested (4000 rpm, 15 min) and resuspended in buffer A (10 mM KHPO$_4$, pH 7.0, 0.5 M NaCl, 10 mM MgCl$_2$, 10% glycerol, 0.1% beta-mercaptoethanol, 5 mM imidazole). After the addition of 2 mM PMSF, resuspended cells were lysed by passing through a French Press cell at 16,000 psi three times and the resulting lysate was cleared by spinning in an ultracentrifuge (40,000 rpm, 40 min). After overnight incubation of cleared lysate with nickel-NTA resin (Qiagen), the tPol$\lambda$ in the
supernatant was purified through a linear gradient of 20 to 500 mM imidazole in buffer B (10 mM KHPO₄, pH 7.0, 0.35 M NaCl, 2.5 mM MgAc₂, 10% glycerol, 0.1% beta-mercaptoethanol). tPolλ-containing fractions were pooled and dialyzed against buffer C (50 mM Tris-Cl, pH 7.0, 25 mM NaCl, 10% glycerol, 2 mM EDTA, and 0.1% beta-mercaptoethanol) at 4 °C. The dialyzed protein solution was passed through 10 ml DEAE-Sepharose column (Amersham Pharmacia Biotech). The loading elute was applied to a MonoS 10/10 column (Amersham Pharmacia Biotech) and eluted using a gradient of 50 to 700 mM NaCl in buffer D (25 mM HEPES, pH 7.5, 50 mM NaCl, 10% glycerol, 1 mM EDTA, 0.1% beta-mercaptoethanol). The fractions containing tPolλ were pooled and dialyzed against buffer D. The dialyzed protein solution was loaded into a pre-packed ssDNA-cellulose column (Sigma). After washing, tPolλ was then eluted with 50 to 1000 mM NaCl gradient in buffer C. The fractions containing tPolλ were pooled and dialyzed against buffer D. The dialyzed tPolλ was passed through 10 ml DEAE-Sepharose column. The loading elute was dialyzed against buffer D and concentrated using a Centriprep YM-30 (Millipore). The concentrated protein was ultimately dialyzed against buffer E (25 mM HEPES, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 50% glycerol). tPolλ was purified to >95% purity based on SDS-PAGE analysis (Figure 8.2). The concentration of the purified tPolλ was measured spectrophotometrically at 280 nm using the calculated extinction coefficient of 39,367 M⁻¹cm⁻¹.

**Synthetic oligonucleotides.** The DNA substrates listed in Table 8.1 were purchased from either Integrated DNA Technologies or TriLink Biotechnologies, and purified by denaturing polyacrylamide gel electrophoresis (18% acrylamide, 8 M urea) and the concentration was determined by UV absorbance at 260 nm with the following extinction coefficients (M⁻¹cm⁻¹): primer 19-mer, ε = 171 000; primer 21-mer, ε = 194 100; D-1 template 41-mer, ε = 396 700; D-6 template 41-mer, ε = 394 200; D-7 template 41-mer, ε
The primer strand 21-mer was 5'-32P labeled by incubation with T4 polynucleotide kinase and [γ-32P]ATP for 1 hour at 37 °C. The unreacted [γ-32P]ATP was subsequently removed by centrifugation via a Biospin-6 column (Biorad). The 5'-32P labeled primer 21-mer was then annealed with the corresponding non-radiolabeled downstream primer 19-mer and template 41-mer at a molar ratio of 1.0:1.15:1.25 respectively, to form the 21-19/41-mer single-nucleotide gapped substrate (the top strand was composed of two oligonucleotides with a single-nucleotide gap). Mixtures to be annealed were denatured at 95 °C for 8 min, and then cooled slowly to room temperature over several hours.

*Optimized reaction buffer L*. All experiments using tPolλ, if not specified, were performed in buffer L containing 50 mM Tris-Cl (pH 8.4 at 37 °C), 5 mM MgCl2, 100 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA. All reactions were carried out at 37 °C.

*Rapid quench experiments*. Experiments were carried out in a rapid chemical quench-flow apparatus (KinTek, PA). The experiments were carried out by allowing enzyme and DNA to preincubate in buffer L. An aliquot of this solution (15 µl) was rapidly mixed with an equal volume of solution of incoming nucleotide in buffer L. The reactions were quenched with 90 µl of 0.37 M EDTA (final concentration) after time intervals ranging from milliseconds to several minutes. All concentrations reported in this paper refer to concentrations during the reaction following rapid mixing.

*Measurement of equilibrium dissociation constant of next incoming nucleotide*. A preincubated solution of tPolλ and a gapped DNA substrate at fixed concentrations was mixed at varying concentrations of Mg2+•dNTP (0.25 to 120 µM) in buffer L at 37 °C to
start the reaction. The reactions at each concentration of \( \text{Mg}^{2+}\cdot\text{dNTP} \) were terminated with 0.37 M EDTA at varying times from milliseconds to minutes. The reaction products were analyzed by sequencing gel analysis. The time course of product formation was fit to a single exponential equation (Data Analysis) for each concentration of \( \text{Mg}^{2+}\cdot\text{dNTP} \) to give the observed rate constant of nucleotide incorporation. The observed rates extracted from these time courses of product formation were plotted against the concentrations of \( \text{Mg}^{2+}\cdot\text{dNTP} \) and these data were fit via hyperbolic regression (Data Analysis) to give the equilibrium dissociation constant of dNTP, \( K_d \), and the maximum rate constant for incorporation of dNTP, \( k_p \).

**Measurement of the dNTP dissociation rate constant (\( k_{off} \)).** \( \text{tPol}\lambda \) (60 nM), unlabeled D-8 (300 nM), \([\alpha-^{32}\text{P}]\text{dGTP} \) (20 \( \mu \text{M} \)), and EDTA (0.5 mM) were preincubated in a buffer which was identical to buffer L except it lacked \( \text{Mg}^{2+} \), to form the \( \text{tPol}\lambda\cdot\text{DNA}\cdot\text{dGTP}[\alpha-^{32}\text{P}] \) ternary complex. This preincubated solution was reacted with a solution containing a large molar excess of unlabeled dGTP (2 mM) and \( \text{Mg}^{2+} \) (5.5 mM) for various time intervals prior to be quenched by 0.37 M EDTA. The reaction mixtures were analyzed as described below.

**Measurement of the dissociation rate constant of 21-19/41-mer.** A preincubated solution of \( \text{tPol}\lambda \) (60 nM), unlabeled D-8 (300 nM), \([\alpha-^{32}\text{P}]\text{dGTP} \) (20 \( \mu \text{M} \)), and 0.5 mM EDTA in a buffer which was identical to buffer L except it lacked \( \text{Mg}^{2+} \), was reacted with a solution containing \( \text{Mg}^{2+} \) (5.5 mM) with no additional dGTP for various time intervals. The reactions were subsequently stopped by the addition of EDTA (0.37 M) and analyzed as described below.
Product analysis. Reaction products were analyzed by sequencing gel electrophoresis (17% acrylamide, 8 M urea, 1x TBE buffer) and quantitated with a PhosphorImager 445 SI (Molecular Dynamics).

Data analysis. Data were fit by nonlinear regression using the program KaleidaGraph (Synergy Software). The single turnover experimental data were fit to equation 1 (single exponential), where \( A \) represents the reaction amplitude

\[
[\text{Product}] = A[1 - \exp(-k_{\text{obs}}t)]
\]  \hspace{1cm} (1)

which is equal to the initial concentration of the enzyme and DNA binary complex, and \( k_{\text{obs}} \) the observed single turnover rate constant. Data from measurement of \( K_d \) of dNTP were fit to equation 2 (hyperbola), where \( k_p \) is the

\[
k_{\text{obs}} = k_p[d\text{NTP}]/\{[d\text{NTP}] + K_d\}
\]  \hspace{1cm} (2)

maximum rate constant of dNTP incorporation. The substrate specificity and polymerase fidelity were calculated as \( (k_p/K_d) \) and \( (k_p/K_d)_{\text{incorrect}}/[(k_p/K_d)_{\text{correct}} + (k_p/K_d)_{\text{incorrect}}] \), respectively. Data from the measurement of nucleotide dissociation rate constant \( k_{\text{off}} \) were fit to equation 3:

\[
[\text{Product}-3'\text{32P}] = [k_p/(k_p + k_{\text{off}})][E\cdot\text{DNA}\cdot d\text{NTP-32P}][1 - \exp(-k_{\text{off}}t)]
\]  \hspace{1cm} (3)

Data from the \([\alpha-32\text{P}]d\text{NTP} \) incorporation experiment were fit to equation 4 (25,26), where \( k_1 \) is the DNA dissociation rate constant, \( k_{\text{obs}} \) is the observed nucleotide incorporation rate constant, and \([E_0] \) is the initial enzyme active site concentration.

\[
[\text{Product}] = \{(k_{\text{obs}}k_1/(k_{\text{obs}} + k_1)) + [k_{\text{obs}}/(k_{\text{obs}} + k_1)]^2\{1 - \exp[-(k_{\text{obs}} + k_1)t]\}\}[E_0]
\]  \hspace{1cm} (4)
8.3 Results

To kinetically define the fidelity and substrate specificity of a polymerase, we need to measure the dependence of nucleotide incorporation rates on nucleotide concentration. However, nucleotide incorporation by the full-length Polλ is inhibited when nucleotide concentration is over 1 μM (13,17). In the kinetic studies of the full-length human Polλ purified in our laboratory, we confirmed the observation of substrate inhibition (W. Abdel-Gawad and Z. Suo, unpublished results). Since the proline-rich domain is determined to be the suppressor based on protein truncation studies (17), we decided to prepare tPolλ (residues 245-575, Figure 8.1) and measure the kinetic parameters and fidelity of this fragment which does not contain the BRCT and proline-rich domains and should not exhibit substrate inhibition. Our kinetic experiments demonstrated that tPolλ was active and no substrate inhibition was observed for this truncated enzyme (see below). tPolλ and Polβ share 33% sequence identity including critical residues involved in nucleotide and DNA binding, and catalysis (10-13). Thus, tPolλ allowed us to estimate the intrinsic fidelity of the full-length Polλ.

Protein purification. The C-terminal hexahistidine-tagged tPolλ (38.2 kDa) was overexpressed in E. coli BL21(DE3) and purified to homogeneity (Figure 8.2) through Ni-affinity, MonoS cation-exchange, single-stranded DNA-cellulose affinity, and DEAE-sepharose ion-exchange columns (Materials and Methods). The yield was approximately 10 mg per liter of initial E. coli culture. The N-terminal eight amino acid residues was
confirmed by protein sequencing analysis while the C-terminal hexahistidine-tag was detected by Western blot analysis using anti-hexahistidine tag antibody (data not shown).

*Optimization of reaction conditions.* Our first objective was to optimize the reaction conditions for nucleotide incorporation by tPolλ. To optimize these conditions, we held all components constant while independently varying MgCl₂ concentration, NaCl concentration, and pH. In these assays, a solution of 30 nM 5'-[^32P]-D-1 (Table 8.1) was preincubated with a 4-fold greater tPolλ and subsequently mixed with 100 μM dTTP at 37 °C using a rapid chemical quench apparatus. The reactions were quenched by the addition of 0.37 M EDTA at various time intervals. The DNA product 22-mer and unreacted substrate 21-mer were then separated via gel electrophoresis, and subsequently quantitated using a PhosphorImager. The data were fit to equation 1 (Materials and Methods). The single turnover rate constant (*k*_obs) varied with the MgCl₂ concentration, NaCl concentration, and the buffer pH in the three aforementioned assays. Notably, in determining the optimal MgCl₂ concentration, we observed a significant decrease in reaction amplitude as we incrementally increased the concentration of MgCl₂ (Figure 8.3). The amplitude decrease was likely due to high ionic strength caused by high MgCl₂ concentration which may inhibit polymerization by either disrupting the interactions between tPolλ and its substrates (DNA and nucleotide), or by forming a non-productive ternary complex of tPolλ•DNA•dNTP (27-29). Interestingly, the observed single-turnover rate constants increased with increasing MgCl₂ concentration (Figure 8.3). The reason for this result is not clearly understood. Interestingly, the effect of MgCl₂ concentration on *k*_obs has also been found in single nucleotide incorporation catalyzed by
rat DNA polymerase beta (30) and *Sulfolobus solfataricus* DNA Polymerase IV (31). Since the increase in MgCl₂ concentration was concomitant with both an increase in the single turnover rate constant and decrease in reaction amplitude, we decided to compromise at an optimal MgCl₂ concentration of 5 mM MgCl₂ which was the highest MgCl₂ concentration that still retained full single turnover reaction amplitude (Figure 8.3).

With the exception of the time courses carried out at the non-physiological pH conditions of 6.0 and 10, we did not observe any other aberrations in regard to reaction amplitude in the optimization process for either pH (Figure 8.5) or NaCl concentration (Figure 8.4). The low reaction amplitudes observed at pH 6 and 10 indicated the ternary complex of tPolλ•DNA•dNTP became partially non-productive due to the acidic or basic reaction conditions. The results obtained from the optimization experiments for tPolλ polymerization were 5 mM MgCl₂ (Figure 8.3), 100 mM NaCl (Figure 8.4), and pH 8.4 (Figure 8.5). Thus, the optimized reaction buffer contains 50 mM Tris-Cl, pH 8.4, 5 mM MgCl₂, 100 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA (buffer L). In addition, similar single turnover rate constants and full reaction amplitudes were observed with 10-fold greater enzyme over DNA (data not shown). This indicated that the 4-fold greater enzyme over DNA was enough to ensure that almost all of the DNA molecules were bound by tPolλ which satisfied the single turnover conditions.

*Rapid equilibrium of nucleotide binding.* The kinetic mechanism of single nucleotide incorporation into a synthetic primer/template substrate catalyzed by many DNA
polymerases has been studied by employing pre-steady state kinetic methods (32-39).

The minimal mechanism shared by all DNA polymerases studied to this point, including DNA polymerase β (35), is shown in Scheme 8.1. In this scheme, the binary complex of enzyme and DNA (E•DNAn) binds an incoming nucleotide (dNTP) to form a ground-state ternary complex E•DNAn•dNTP. This ternary complex undergoes both a protein conformational change and chemical reaction to form the product-containing ternary complex, E•DNAn+1•PPP. Notably, these two steps are simplified into one step as shown in Scheme 8.1. During this step, the DNA primer is elongated by one nucleotide and pyrophosphate (PPi) is formed. For subsequent turnovers to occur in single nucleotide incorporation, the enzyme has to dissociate from the product ternary complex. Our kinetic analysis described below suggested Scheme 8.1 was the kinetic pathway for DNA polymerization catalyzed by tPolλ.

To evaluate whether the binding of an incoming nucleotide was at rapid equilibrium (k_on, k_{off} >> k_p), we carried out an experiment to assess the relative rate constants of dNTP dissociation (k_{off}) and nucleotide polymerization (k_p) (Scheme 8.1). Here, we mixed a preincubated solution of tPolλ, 5-fold unlabeled D-8, and [α-32P]dGTP (20 μM) in the optimized buffer lacking Mg^{2+} and containing EDTA (0.5 mM) with a solution containing a large molar excess of unlabeled dGTP (2 mM) and Mg^{2+} for various reaction times followed by quenching with EDTA. Although tPolλ was purified and stored in the absence of divalent metal ions, we used additional 0.5 mM EDTA present in the preincubated enzyme solution to chelate any contaminant divalent cations carried over from the protein purification. This concentration of EDTA was shown to be sufficient to
prevent any product formation in this preincubated solution (data not shown). If the
dissociation of dNTP from the E•DNA•dNTP ternary complex was much faster than the
polymerization we would expect to see very little 3’-[α-32P]-labeled product formation
due to the unfavorable kinetic partitioning and large molar excess of unlabeled dGTP that
once bound to the E•DNA binary complex, would remove the ternary complex from
observation. This assay resulted in a time course showing an insignificant amount of
radio-labeled product (approximately 1 nM at the longest reaction time in Figure 8.6).
Additionally, each product concentration in this time course was corrected for the
contribution to product formation from the incorporation of dissociated [α-32P]dGTP as
measured in a control experiment. In this control experiment, we preincubated a solution
of tPol λ (60 nM), unlabeled D-8 (300 nM), and unlabeled dGTP (20 μM) in the
optimized buffer lacking Mg2+ and containing 0.5 mM EDTA, and then mixed this
solution with another solution containing a large molar excess of unlabeled dGTP (1.98
mM), 20 μM [α-32P]dGTP, and Mg2+ (5.5 mM) for various reaction times.

The time course (●) shown in Figure 8.7 was fit to equation 3 (Materials and Methods)
to obtain the nucleotide dissociation rate constant from the E•DNA•dNTP-32P complex.
The measured $k_{off}$ and $k_p$ are $300 \pm 100$ s$^{-1}$ and $6 \pm 2$ s$^{-1}$, respectively. Thus, the nucleotide
dissociation rate constant was indeed faster than the polymerization rate constant.
Moreover, the association rate constant of the binding of dGTP to the tPolλ•D-8 binary
complex $k_{on} = k_{off}/K_d = 1.55 \times 10^8$ M$^{-1}$s$^{-1}$ was calculated from the values of $k_{off}$ (300 s$^{-1}$)
and $K_d$ (1.93 μM) which was measured below. This suggested the binding of a nucleotide
to the binary complex tPolλ•DNA was rapid and under diffusion-control. Therefore, the
binding of a nucleotide to form the ground-state ternary complex tPolλ•DNA•dNTP was at rapid equilibrium. Such a fast equilibrium has been observed with the Klenow fragment of *E. coli* DNA polymerase I (40), T7 DNA polymerase (34) and *S. solfataricus* DNA Polymerase IV (32). This rapid equilibrium of nucleotide binding which was much faster than the polymerization, allows us to measure the ground-state binding affinity of an incoming nucleotide.

*Measurement of DNA dissociation rate constant.* To ensure that the above nucleotide incorporation rate constant was not affected by preincubating the E•DNA•dNTP ternary complex in the absence of Mg$^{2+}$, we performed another control experiment. In this experiment, a preincubated solution of tPolλ (60 nM), unlabeled D-8 (300 nM), [$\alpha$-32P]dGTP (20 μM), and EDTA (0.5 mM) in the optimized buffer lacking Mg$^{2+}$ was reacted with a solution containing Mg$^{2+}$, and no additional unlabeled trap dGTP for various reaction times. This experiment yielded a time course of product formation (■) (Figure 8.7) that was very similar to that obtained when a preincubated tPolλ•DNA binary complex in the presence of Mg$^{2+}$ reacted with dNTP•Mg$^{2+}$ (data not shown). This suggested the binding of nucleotide was rapid and the nucleotide incorporation was not affected by preincubating the E•DNA•dNTP ternary complex in the absence of Mg$^{2+}$. The data (■) were fit to equation 4 (Materials and Methods) to obtain a $k_{obs}$ of 2.1 ± 0.3 s$^{-1}$ and a DNA dissociation rate constant ($k_1$ in Scheme 8.1) of 0.8 ± 0.3 s$^{-1}$. The smaller observed polymerization rate constant $k_{obs}$ in comparison to the $k_p$ values obtained above and below was due to unsaturated dNTP concentration used in this experiment. The reason why the data (■) were not fit into a burst equation was because the $k_{obs}$ value was
not significantly larger (roughly 3-fold difference) than \( k_1 \) (25,26). The comparable \( k_{obs} \) and \( k_1 \) values prevent the appearance of an obvious burst phase in the time course (■).

**Substrate specificity of correct nucleotide.** The ground-state binding affinity of dTTP \( (K_d) \) to the E•DNA binary complex (Scheme 8.1) was measured through the dTTP concentration dependence of the single turnover rate constant \( (k_{obs}) \). The single turnover method was employed because the above measured DNA dissociation rate constant \( (k_1) \) was high and the burst phase was not obvious in Figure 8.7. The experiments were performed with enzyme in molar excess over DNA substrate to allow the direct observation of nucleotide incorporation in a single pass of the reactants through the enzymatic pathway without complications resulting from the steady-state formation of products (26). A preincubated solution of 5'-radiolabeled D-1 and 4-fold tPolλ was reacted with increasing concentrations of dTTP in buffer L. The DNA product 22-mer and unextended substrate 21-mer at different time intervals were separated and quantitated as described in detail in Materials and Methods. The product concentration was plotted against reaction time intervals. The data were subsequently fit to a single exponential, equation 1 (Materials and Methods), to yield a single turnover rate constant at each concentration of dTTP (Figure 8.8). The single turnover rates were then plotted against dTTP concentrations (Figure 8.9). The data were subsequently fit to a hyperbolic equation, equation 2 (Materials and Methods), to yield a \( k_p \) of 6.0 ± 0.2 s\(^{-1}\) for the maximum dTTP incorporation rate constant, and a \( K_d \) of 2.4 ± 0.4 µM for dTTP binding. Additionally, the value of the substrate specificity \( (k_p/K_d) \) was calculated to be 2.5 µM\(^{-1}\)s\(^{-1}\) (Table 8.2).
Similar analyses implementing single turnover conditions were used to determine the kinetic parameters ($k_p$, $K_d$, and $k_p/K_d$) for the incorporations of the other three correct nucleotides, dCTP into D-6, dATP into D-7, and dGTP into D-8 (data not shown), and the parameters are listed in Table 8.2. The ground-state binding affinity and, moreover, the substrate specificity of all four correct nucleotide incorporations under single turnover conditions were similar. The binding affinity of correct nucleotides to the tPol$\lambda$•DNA binary complex is similar to that observed in Pol$\beta$ with a single-nucleotide gapped DNA (41).

**Substrate specificity of incorrect nucleotide.** Pre-steady state kinetic analysis of incorrect dGTP incorporation into D-1 was assayed by implementing similar single turnover kinetic methods as described above. tPol$\lambda$ was preincubated with 5'-radiolabeled D-1 and was then reacted with increasing concentration of dGTP in buffer L. The reactions were manually quenched with EDTA, analyzed via gel electrophoresis, and the products were quantitated using a PhosphorImager. The single turnover rate constant observed at each concentration of dGTP was obtained through the fit of the time course of product formation to equation 1 (Figure 8.10). The observed reaction rates were then plotted against the concentrations of dGTP and these data were fit by nonlinear regression into equation 2 (Materials and Methods) to obtain $k_p$, $K_d$, and substrate specificity ($k_p/K_d$) of $0.022 \pm 0.001 \text{ s}^{-1}$, $8.4 \pm 0.7 \mu$M, and $2.6 \times 10^{-3} \mu$M$^{-1}$s$^{-1}$, respectively (Table 8.2 and Figure 8.11).
With DNA substrates D-1, D-6, D-7, and D-8 the kinetic parameters for each nucleotide of the remaining eleven possible incorrect single nucleotide incorporations were determined under the same single turnover conditions (Table 8.2). On average, the incorrect nucleotide incorporations have only about 2-fold lower ground-state binding affinity ($K_d$) and one to three orders of magnitude slower incorporation rates when compared to the four correct nucleotide incorporations. Since the tight binding of mismatched nucleotides was surprising, we repeated the measurements of the misincorporations and obtained similar values to those listed in Table 8.2. In addition, we measured the kinetic parameters of dCTP incorporation into D-7 in the presence of 10-fold, rather than 4-fold (Table 8.2), excess of enzyme over DNA (data not shown). The values of $k_p$ and $K_d$ differed less than 5% of those listed in Table 8.2. This result not only confirmed the above conclusion that 4-fold excess enzyme over DNA was enough to satisfy the single-turnover conditions, but also indicated the tighter binding of an incorrect nucleotide to the tPol$\lambda$•DNA binary complex was not due to fast DNA dissociation.

Interestingly, while dGTP binding to the E•D-7 binary complex, which forms the T:dGMP wobble base pair, was characteristic of incorrect nucleotide incorporation for tPol$\lambda$, the maximum rate constant of polymerization ($k_p$) was 10- to 100-fold faster than typical incorrect incorporations, leading to its apparent higher substrate specificity. In contrast, the kinetic parameters for the dTTP incorporation into D-6 to form the G:dTMP mismatch base pair were similar to those of other mismatches. This apparent asymmetry was observed only with these two mismatches.
The values of substrate specificity and fidelity (Materials and Methods) for all twelve incorrect incorporations were calculated and listed in Table 8.2. The fidelity of tPolλ with the single-nucleotide gapped DNA is in the range of $10^{-2}$-$10^{-4}$, and it has the lowest value for the T:dGMP wobble base pair.

*Strand-displacement activity.* For the incorporation of the correct nucleotide dCTP into D-6, we observed the formation of both 22-mer and 23-mer products (Figure 8.12). The formation of 23-mer was due to a second dCTP incorporation opposite the downstream template base G which was initially base paired with the downstream primer 19-mer (Table 8.1). We also observed a small amount of 23-mer in addition to the major product 22-mer for the incorporations of mismatched nucleotides with the exception of dCTP (data not shown). These results indicated that tPolλ had a weak strand-displacement activity. This activity was not observed for the other three correct nucleotide incorporations including dTTP into D-1, dATP into D-7, and dGTP into D-8 since only 22-mer was observed in each of these three cases (data not shown). We believe the reason for these observations is due to a combination of slow rates of misincorporation on top of the downstream template G in these three cases coupled with much shorter reaction time intervals in these assays that yielded a negligible amount of product 23-mer.

Surprisingly, the misincorporation of dCTP into D-1 (Figure 8.13), D-7 (Figure 8.14), and D-8 (Figure 8.15) yielded products ranging from 22-mer to 26-mer and the major product was 23-mer. Interestingly, the next two downstream template bases of the three
DNA substrates are G followed by C. The incorporation of a second dCTP against the template G was matched and fast while the third dCTP incorporation was mismatched opposite the template C and slow, leading to the accumulation of 23-mer. The downstream template base G seemed to facilitate the strand-displacement activity of tPolλ in the multiple incorporations of dCTP. To prove this hypothesis, we changed the next template base G in D-8 to A (D-12 in Table 8.1) and the incorporation of dCTP into D-12 yielded predominantly 22-mer and a small amount of 23-mer (Figure 8.16). Moreover, the incorporation of dTTP into D-12 showed similar pattern as dCTP into D-1, D-7, and D-8: (i) three products and (ii) 23-mer as the major product (Figure 8.17). The overall substrate specificity \((k_p/K_d)\) was 5-fold lower for dCTP incorporation into D-12 than dCTP into D-8 while 2-fold higher for dTTP into D-12 than dTTP into D-8 (Table 8.2). These results suggested several misincorporations of a single nucleotide occurred without deletions or additions. In addition, these results were consistent with the hypothesis that a correct incorporation among multiple misincorporations facilitated the strand-displacement activity of tPolλ.

8.4 Discussion

Mechanistic studies of DNA polymerization. In this chapter, we detailed the overexpression and purification of the polβ-like domain of Polλ (tPolλ) (Figures 8.1 and 8.2). In our limited mechanistic studies of tPolλ, we determined the dNTP dissociation rate constant of 300 s\(^{-1}\) by performing an unlabeled dGTP trapping experiment (Figure 8.6). As far as we know, this is the first direct measurement of the dissociation rate constant of a nucleotide from the ground-state ternary complex E•DNA•dNTP. However,
this ground-state complex may not form in the absence of Mg$^{2+}$ on the basis of published crystal structures and metal ion binding studies of other polymerases (42-44). These studies have revealed a common two-metal-ion catalytic mechanism: one divalent ion is involved in both positioning the $\alpha$-phosphate of the incoming nucleotide and activation of its 3’-hydroxyl group as a nucleophile, the other metal ion both anchors the binding of the $\beta$- and $\gamma$-phosphates of the nucleotide and assists the leaving of the pyrophosphate. The small amount of radio-labeled product shown in Figure 8.6 suggested the existence of E•DNA•dNTP. However, we cannot exclude the possibility that this complex may not be the same complex in the absence and presence of Mg$^{2+}$. Furthermore, the measured dissociation rate constant should not be for the nucleotide dissociation from the tight binding ternary complex E’•DNA•dNTP which formed after initial E•DNA•dNTP complex undergoes protein conformational change. This is because the binding of the divalent metal ions induces the protein conformational change which may be a prerequisite for catalytic activity by correctly positioning the side chains of the residues located at the polymerase active site (35,45). Moreover, the measured dissociation rate constant (300 s$^{-1}$) is too high for nucleotide dissociation from tight binding ternary complex E’•DNA•dNTP. The nucleotide association rate constant (1.55x10$^8$ M$^{-1}$s$^{-1}$) calculated from this nucleotide dissociation rate constant also supports the measured dissociation rate constant is for nucleotide dissociation from the ternary complex E•DNA•dNTP since a small nucleotide molecule is expected to bind to the E•DNA binary complex at a rate constant close to the diffusion limit (1.0x10$^8$ M$^{-1}$s$^{-1}$). The large nucleotide association and dissociation rate constants indicated a rapid equilibrium of nucleotide binding relative to nucleotide incorporation.

In the absence of the unlabeled dGTP trap, the incorporation of [$\alpha^{-32}$P]dGTP allowed us to measure the DNA dissociation rate constant of 0.8 s$^{-1}$ which was less than 3-fold lower
than the nucleotide incorporation rate constant (Figure 8.7). This suggested tPolλ has a processivity ($= k_p/(k_i + k_p)$) of 0.72, which is defined as the likelihood of incorporating the next correct nucleotide following each correct nucleotide incorporation event. The processivity value of a highly processive DNA polymerase is close to 1 due to $k_p$ being much larger than $k_i$. However, the processivity of tPolλ is low which is consistent with previous reports suggesting that this enzyme is not a processive polymerase (10,13,17), and with its potential function as a polymerase which fills short-patched DNA gaps in base-excision repair pathways (see discussion below). The fast DNA dissociation rate constant has also been found with Polβ (41).

The proposed kinetic mechanism shown in Scheme 8.1 can explain all of our kinetic data at present. However, it is a minimal mechanism and requires more studies to be completed. For example, the nucleotide incorporation in the first turnover could be limited by either a putative protein conformational change as observed in replicative DNA polymerases such as T7 DNA polymerase (34) and human mitochondrial DNA polymerase (39) and in Y-family DNA polymerases such as yeast DNA polymerase η (33) and S. solfataricus DNA Polymerase IV (32) as described in Chapter 3, or the chemistry step as observed only in the X-family DNA polymerase Polβ (44,46). It would be interesting to see if the chemistry step is also rate-limiting in the first turnover of nucleotide incorporation catalyzed by another X-family polymerase tPolλ. The detailed kinetic mechanism of tPolλ is currently being studied in our laboratory.

**Fidelity of DNA polymerization.** The construct of tPolλ, unlike the full-length Polλ, did not exhibit dNTP substrate inhibition in our kinetic experiments (Figures 8.8 to 8.11) which allowed us to estimate the fidelity of the full-length Polλ using pre-steady state kinetic assays. The N-terminal BRCT and proline-rich domains (Figure 8.1) could affect
the misincorporation fidelity by the full-length Polλ. However, the fidelity of nucleotide incorporation into D-1 by the full-length Polλ was provisionally estimated to be in the range of $10^{-3}$ to $10^{-4}$ (W. Abdel-Gawad and Z. Suo, unpublished results) which is similar to the fidelity of tPolλ (Table 8.2). Thus, the N-terminal domains do not affect the intrinsic fidelity of Polλ although they either mediate the protein/protein or protein/DNA interactions through the BRCT domain, or regulate the polymerase activity of Polλ through the proline-rich domain (10-13). While this hypothesis is supported by the preliminary results that dNTP substrate inhibition was only observed with the full-length Polλ, not tPolλ, this topic will be explored further in Chapter 9.

Under single turnover conditions, the fidelity of tPolλ was estimated to be $3.2 \times 10^{-2}$ for the T:dGMP misincorporation and $10^{-3}$ to $10^{-4}$ for all other incorrect incorporations (Table 8.2). The latter value is similar to the single-base substitution error rate of $9.0 \times 10^{-4}$ when Polλ fills a 5’-nucleotide gap at a TGA codon in the lacZ gene in M13mp2 DNA (13). It is also similar to the base-substitution rate of Polλ ($9.0 \times 10^{-4}$) scored both in the lacZ α gene in M13mp2 during synthesis to fill a 407-nucleotide gap, and in the 6-nucleotide gap frameshift reversion assay (22). Interestingly, these M13mp2-based assays (13,22) also reveal the substitution error is 10-fold higher at the T:dGMP wobble base pair than other mispairs. This observation agrees with our kinetic results (Table 8.2). Therefore, the results from these M13mp2-based assays are consistent with the results generated from our pre-steady-state kinetic analysis. However, only our analysis revealed the kinetic basis for the fidelity of Polλ: (i) the selection was predominantly from the 10- to 1000-fold difference in incorporation rate constants of correct and incorrect nucleotides while the ground-state binding affinity provided only about a 2-fold contribution and (ii) the 10-fold lower fidelity at the wobble T:dGMP base pair is mainly due to the 10- to 100-fold higher misincorporation rate compared to the other mispairs.
Notably, we did not observe any frameshifts which have been found to be extremely high ($10^{-2}$ to $10^{-3}$) by the M13mp2-based assays (22). This difference can probably be attributed to different DNA gap-sizes: we used single-nucleotide gap in our studies while K. Bebenek et al. use either 6- or 407-nucleotide gaps. The downstream 5’-phosphorylated primer 19-mer in our DNA substrate 21-19/41-mer formed base pairs with the template 41-mer and prevents the formation of a misaligned frameshift DNA intermediate since only one template base was not paired. The weak strand-displacement activity of tPolλ (see below) should not affect the frameshift frequency since this activity each time unwound only one basepair formed between the 19-mer and 41-mer when a single-nucleotidne gap is filled.

The fidelity of tPolλ is about 10- to 100-fold lower than the fidelity of rat Polβ (recalculated as $10^{-4}$-$10^{-5}$ using the definition of fidelity in Table 8.2) which was measured under single turnover conditions with a single-nucleotide gapped DNA 25-19/45-mer (primer-primer/template) (41). A 5- to 10-fold higher substitution rate of Polλ over Polβ was also observed by the M13mp2-based assays (22). Interestingly, the unusually high infidelity for the wobble T:dGMP misincorporation was not observed with Polβ (41). In Polβ, the nucleotide ground-state binding affinity contributes 10- to 100-fold to the fidelity in addition to the 100- to 1000-fold contribution from the incorporation rates (41). The higher selection in nucleotide ground-state binding step thereby contributes to higher fidelity of Polβ over Polλ. The incorporation rates for both correct and incorrect nucleotides are faster with Polβ than with tPolλ. The binding affinity of correct nucleotides to these two polymerases is similarly high, but the incorrect nucleotides are more weakly bound by Polβ than by tPolλ (see below). These results suggest Polβ and Polλ are quite different enzymes although they share sequence
homology and are expected to have structural similarity (10-13). These different enzyme activities lead us to believe that these two polymerases play different biological roles.

**Wobble T:dGMP basepair.** Although the ground-state binding affinity is similar, Table 8.2 shows the misincorporation rate of dGTP opposite template base T is only about 10-fold slower than the incorporation of correct dATP and is much faster than all other misincorporations including dTTP opposite template base G. Such an asymmetry between T:dGMP and G:dTMP has been observed with DNA polymerase ι (47) and not with any other DNA polymerases including replicative polymerases. In DNA polymerase ι, the values of the substrate specificity estimated by steady-state kinetic analysis are 1.5 × 10^{-1} and 2.85 × 10^{-2} for the incorporations of T:dGMP and G:dTMP, respectively (47). The incorporation efficiency difference is even larger (11-fold) with tPolλ (Table 8.2). It would be interesting to see if such an asymmetry also occurs to the full-length Polλ. The reason for the faster formation of the T:dGMP wobble base pair over the G:dTMP is not clear and cannot be explained simply from a hydrogen-bonding perspective. It could be rationalized by consideration of increasingly favorable base-stacking interactions between dGTP, a purine, with the 3’-end base of the primer over a pyrimidine like dTTP. The more stabilized T:dGMP, when compared to G:dTMP, could align the 3’-hydroxyl moiety of the primer and the incoming nucleotide more favorably at the enzyme active site and facilitate catalysis. The reason for the faster formation of the T:dGMP wobble basepair over the other 10 mispairs is not clear either. It could be rationalized by more hydrogen-bonds between bases G and T than other mispairs. Interestingly, the asymmetry between the incorporations of T:dGMP and G:dTMP is not observed with Polβ (41), suggesting the active sites of the two X-family members are different. Further studies are needed to resolve the structural basis of this wobble base pair in terms of which amino acid residues and the type of interactions involved in the T:dGMP wobble base pairing.
Tight ground-state binding of incoming nucleotide. It is not surprising that the correct nucleotides are incorporated by tPol λ much faster than incorrect nucleotides. However, Table 8.2 shows the incoming nucleotides, regardless of whether they are correct or incorrect, have relatively high and peculiarly similar binding affinity to the binary complex of tPol λ and the single-nucleotide gapped DNA substrate. The tight binding of all incorrect nucleotides is unprecedented since most of DNA polymerases which have been studied so far generally discriminate against incorrect nucleotides by binding them weakly and incorporating them slowly (48). For example, the mismatched nucleotides have 100- to 200-fold weaker binding affinity than the matched ones to the binary complex of Pol β and a single-nucleotide gapped DNA (Table 8.3) (41). The only exception is the binding affinity of mismatched dGTP opposite a template base G in a single-nucleotide gapped DNA substrate by African swine fever virus polymerase X which is 5-fold tighter than the Watson-Crick base pair dCTP:G (49). Interestingly, this enzyme is another member of the X-family DNA polymerases. The tighter binding of mismatched dGTP over matched dCTP is probably because this X-family polymerase has higher intrinsic affinity towards deoxypurine triphosphates than deoxypyrimidine triphosphates (50-52).

The tight binding of matched nucleotides by Pol β with the gapped DNA is partly due to the contribution of its 8-kDa dRPase domain revealed by crystal structures (48). This domain does not specifically contact the DNA, nucleotide, nor the rest of Pol β in the presence of a non-gapped DNA substrate (9), but shows intense interactions with the downstream primer and the 31 kDa domain in the presence of a single-nucleotide gapped DNA (53). This leads to tighter interactions between the incoming nucleotide and surrounding amino acid residues. Although unavailable at present, modeling studies show
that the ternary structure of tPolλ is expected to be similar to the structure of Polβ and the tight binding of correct nucleotide by tPolλ is thus structurally reasonable. However, the difference in the binding affinity of incorrect nucleotides by the two homologous X-family members is significant (Table 8.3). Although the residues surrounding the incoming nucleotide in Polβ are mostly conserved in Polλ, sequence alignment analysis suggests the residues K27, R40, A185, K280, and D276 in human Polβ are changed to S268, A280, K422, R514, and A510 in human Polλ, respectively (13,22). Residue D276, which makes van der Waals interactions with the base of the incoming nucleotide in Polβ (53), weakens nucleotide binding due to its negative charge. Mutation of this residue to a neutral residue (e.g. valine and glycine) increases the correct nucleotide binding affinity by 4- to 9-fold and incorrect nucleotide binding affinity by 2.5-fold (48,54).

Consequently, its replacement with the uncharged A510 in Polλ could account for the increase in the nucleotide binding affinity, particularly for incorrect nucleotides. We are currently studying the roles of A510 and the other four residues in the binding of nucleotides by tPolλ through site-directed mutagenesis and single turnover kinetic assays.

**Weak strand-displacement activity.** Multiple product formation patterns shown in Figures 8.12 to 8.17 suggest the downstream base pairs formed between the 19-mer primer and the 41-mer template were melted during polymerization. The melting could be due to either thermal breathing of the 5’-terminus of the 19-mer or the strand-displacement activity of tPolλ or both. Interestingly, the misincorporations of dCTP into D-1, D-7, and D-8 and the misincorporation of dTTP into D-12 yielded longer products than other types of mismatches, and the values of substrate specificity were slightly higher (Table 8.2). These observations suggest the strand-displacement activity of tPolλ, rather than thermal breathing, was the major force to melt downstream base pairs during polymerization. Such an activity of tPolλ has been observed previously (14). The energy source for this
weak strand-displacement activity is probably derived from the net favorable free energy of nucleotide incorporation. One matched incorporation among multiple misincorporations (G:dCMP in D-1, D-7, and D-8; A:dTMP in D-12) facilitates the strand-displacement activity of tPolλ since the net favorable energy yielded from a correct nucleotide incorporation will be larger than from the incorporation of a mismatched nucleotide.

*Potential biological functions.* Polλ has been suggested to play a role in DNA repair synthesis associated with meiosis since it is found to be predominantly expressed in testis in stages of spermatogenesis (11). Polλ does not have proof-reading exonuclease function, but possesses a dRPase and strand-displacement activity (Figures 8.12 to 8.17). Moreover, Polλ has a BRCT domain which mediates protein/protein and protein/DNA interactions. Therefore, Polλ has been proposed to play a similar role as Polβ in BER (14). Since the insertion fidelity of full-length Polλ was as low as the fidelity of tPolλ ($10^{-2}$ - $10^{-4}$) based on our preliminary studies (W. Abdel-Gawad and Z. Suo, unpublished results), the error rate will be too high for Polλ to function as the DNA polymerase in “long patch” BER proposed previously (14,19). However, Polλ, like Polβ, could function as a DNA polymerase in “short patch” BER since the mutation possibility will be minimized if it only incorporates one nucleotide in a single-nucleotide gap. This potential function is indirectly supported by the high frameshift rate of Polλ as revealed by both the forward mutation assay (22) and the short-gap frameshift reversion assay (13). The high frameshift rate would be detrimental to genetic stability if the polymerase makes a lot of deletions when it fills a large DNA gap. In addition, the low processivity (see above discussion) and weak strand-displacement activity of Polλ also supports its role in “short patch” BER, rather than “long patch” BER.
Recent immunodepletion studies reveal that Polλ is the primary DNA polymerase to fill 1- or 2-nucleotide gaps during NHEJ in an *in vitro* system based on human nuclear extracts and that the BRCT domain of Polλ is required for this short-gap filling activity (23). The low fidelity and short-gap filling ability of Polλ revealed by our studies support this potential role. Moreover, we and others (13,17) have observed the dNTP substrate inhibition with the full-length Polλ. These observations, coupled with the tight binding of nucleotides (Table 8.2), suggest that Polλ would function as a polymerase at low cellular concentrations of dNTPs. Interestingly, cellular dNTP concentrations are highest during S and G2 phases and lowest during G0 phase (55) and the NHEJ pathway plays a dominant role in repairing gamma-radiation-induced DSBs during G0, G1 and early S phases while homologous recombination is preferentially used in late S and G2 phases (56). These results support the hypothesis that Polλ may function as a polymerase in the NHEJ pathway during the G0 phase (18,22,23). This potential function is further substantiated by the end-joining role of its close homolog in yeast, DNA polymerase IV (57,58).
8.5 Figures, Tables, Schemes

Figure 8.1. Schematic representations of the full-length and truncated DNA polymerase \( \lambda \) as well as human DNA polymerase \( \beta \). Each domain, with amino acid residue numbers indicated above, is shown as a rectangle.
Figure 8.2. Purification of the N-terminal truncated human DNA polymerase λ. SDS-PAGE analysis and subsequent Coomassie Blue staining of proteins at each step of the purification. Lane M, protein marker; lane 1, crude extracts of non-induced cells; lane 2, crude extracts of IPTG-induced cells; lane 3, clear lysate; lane 4, eluate from Ni-affinity column; lane 5, eluate from MonoS cation-exchange column; lane 6, eluate from ssDNA-cellulose column; lane 7, loading elute from DEAE-sepharose column.
Figure 8.3. Optimization of the reaction conditions for \(\text{tPol}_\lambda\). Effects of \(\text{Mg}^{2+}\) concentration on the activity of \(\text{tPol}_\lambda\) at 37 °C. A preincubated solution of 5'-\(^{32}\)P-labeled-D-1 (30 nM) and 4-fold greater \(\text{tPol}_\lambda\) (120 nM) was mixed with correct incoming nucleotide (100 μM dTTP) for various reaction times under single turnover conditions. Concentrations of all components were held constant while \(\text{Mg}^{2+}\) concentration was varied independently. Reaction rate constants \((k_{\text{obs}}, \text{s}^{-1})\) (○) and reaction amplitudes (nM) (■) were plotted simultaneously.
Figure 8.4. Optimization of the reaction conditions for tPolλ. Effects of NaCl concentration on the activity of tPolλ at 37 °C. A preincubated solution of 5'-32P-labeled-D-1 (30 nM) and 4-fold greater tPolλ (120 nM) was mixed with correct incoming nucleotide (100 μM dTTP) for various reaction times under single turnover conditions. Concentrations of all components were held constant while NaCl concentration was varied independently. Reaction rate constants ($k_{obs}$, s$^{-1}$) (●) and reaction amplitudes (nM) (■) were plotted simultaneously.
Figure 8.5. Optimization of the reaction conditions for tPolλ. Effects of pH concentration on the activity of tPolλ at 37 °C. A preincubated solution of 5’-32P-labeled-D-1 (30 nM) and 4-fold greater tPolλ (120 nM) was mixed with correct incoming nucleotide (100 μM dTTP) for various reaction times under single turnover conditions. Concentrations of all components were held constant while reaction pH was varied independently. Reaction rate constants ($k_{obs}$, 1/s) (●) and reaction amplitudes (nM) (■) were plotted simultaneously. Activity was assayed in 25 mM MES-NaOH buffer between pH 6.0 and 7.0, 25 mM Tris-Cl buffer for pH 8.0 and 8.5, and 25 mM glycine-NaOH buffer for pH 9.0 and 10.0.
Figure 8.6. Measurements of dNTP ($k_{off}$) rate constant and nucleotide incorporation rate constant ($k_p$). A preincubated solution of tPol$\lambda$ (60 nM), unlabeled D-8 (300 nM), [$\alpha$-$^{32}$P]dGTP (20 $\mu$M), and 0.5 mM EDTA in the absence of Mg$^{2+}$ was reacted with a solution containing a large molar excess of unlabeled dGTP (2 mM) and Mg$^{2+}$ (5.5 mM) for various reaction times (●). The product concentrations in this time course were corrected for the contribution to product formation from the incorporation of dissociated [$\alpha$-$^{32}$P]dGTP (see Text). The data (●) were fit to equation 3 (Materials and Methods) to obtain the nucleotide dissociation rate constant ($k_{off}$) from the E•DNA•dNTP-$^{32}$P complex. The measured $k_{off}$ and $k_p$ are $300 \pm 100$ and $6 \pm 2$ s$^{-1}$, respectively.
Figure 8.7. Measurements of DNA ($k_1$) dissociation rate constants and nucleotide polymerization rate constant ($k_p$). A preincubated solution of tPolλ (60 nM), unlabeled D-8 (300 nM), [$\alpha$-32P]dGTP (20 μM), and 0.5 mM EDTA in the absence of Mg$^{2+}$ was reacted with a solution containing Mg$^{2+}$ (5.5 mM) with no additional unlabeled dGTP for various reaction times (■). The data were fit to equation 4 (see Materials and Methods) to obtain the $k_p$ of 2.1 ± 0.3 s$^{-1}$ and a fast DNA dissociation rate constant ($k_1$) of 0.8 ± 0.3 s$^{-1}$.  

\[ \text{Product-3'} \text{-} 32\text{P (nM)} \]

\[ \text{Reaction Time (s)} \]
Figure 8.8. Concentration dependence on the pre-steady state rate constant of correct nucleotide incorporation. A preincubated solution of tPol λ (120 nM) and 5’-32P-labeled-D-1 (30 nM) was mixed with increasing concentrations of Mg2+-dTTP (0.25 μM, ○; 0.5 μM, □; 1.5 μM, ■; 6 μM, △; 12 μM, ▲; 30 μM, ▼; 60 μM, ▼; 120 μM, ▼; 200 μM, ◆) for various time intervals. The solid lines are the best fits to the single exponential equation (equation 1).
Figure 8.9. Concentration dependence on the pre-steady state rate constant of correct nucleotide incorporation. The single exponential rates obtained from the above data fitting were plotted as a function of dTTP concentration. The data (●) were then fit to the hyperbolic equation (equation 2) yielding a $k_p$ of $6.0 \pm 0.2$ s$^{-1}$ and a $K_d$ of $2.4 \pm 0.4$ μM.
Figure 8.10. Concentration dependence on the pre-steady state rate constant of incorrect dGTP incorporation. A preincubated solution of tPolλ (120 nM) and 5'-32P-labeled-D-1 (30 nM) was mixed with increasing concentrations of Mg\textsuperscript{2+}-dGTP (0.25 μM, ○; 0.5 μM, □; 1.5 μM, ▓; 6 μM, ▣; 12 μM, △; 30 μM, ▲; 60 μM, ▼; 100 μM, ▽; 120 μM, ◆) for various time intervals. The solid lines are the best fits to the single exponential equation.
Figure 8.11. Concentration dependence on the pre-steady state rate constant of incorrect dGTP incorporation. The single exponential rates obtained from the above data fitting were plotted as a function of dGTP concentration. The data (●) were then fit to the hyperbolic equation yielding a $k_p$ of $0.022 \pm 0.001$ s$^{-1}$ and a $K_d$ of $8.4 \pm 0.7$ µM.
Figure 8.12. Series of gel pictures showing progression of product formation as a function of time under conditions of 100 $\mu$M dCTP into D-6. Unreacted substrate (21-mer) is shown at the bottom of each gel picture with extended products located sequentially above corresponding unreacted substrate. Reaction time intervals (seconds) are denoted below corresponding lane.
Figure 8.13. Series of gel pictures showing progression of product formation as a function of time under conditions of 100 μM dCTP into D-1. Unreacted substrate (21-mer) is shown at the bottom of each gel picture with extended products located sequentially above corresponding unreacted substrate. Reaction time intervals (seconds) are denoted below corresponding lane.
Figure 8.14. Series of gel pictures showing progression of product formation as a function of time under conditions of 100 μM dCTP into D-7. Unreacted substrate (21-mer) is shown at the bottom of each gel picture with extended products located sequentially above corresponding unreacted substrate. Reaction time intervals (seconds) are denoted below corresponding lane.
Figure 8.15. Series of gel pictures showing progression of product formation as a function of time under conditions of 100 μM dCTP into D-8. Unreacted substrate (21-mer) is shown at the bottom of each gel picture with extended products located sequentially above corresponding unreacted substrate. Reaction time intervals (seconds) are denoted below corresponding lane.
Figure 8.16. Series of gel pictures showing progression of product formation as a function of time under conditions of 100 μM dCTP into D-12. Unreacted substrate (21-mer) is shown at the bottom of each gel picture with extended products located sequentially above corresponding unreacted substrate. Reaction time intervals (seconds) are denoted below corresponding lane.
Figure 8.17. Series of gel pictures showing progression of product formation as a function of time under conditions of 100 μM dTTP into D-12. Unreacted substrate (21-mer) is shown at the bottom of each gel picture with extended products located sequentially above corresponding unreacted substrate. Reaction time intervals (seconds) are denoted below corresponding lane.
D-1 5’-CGCAGCCGTCCAACCAAATCGTCC-3’
3’-CGTCGATCCAATGCCGTCC-5’

D-6 5’-CGCAGCCGTCCAACCAAATCGTCC-3’
3’-CGTCGATCCAATGCCGTCC-5’

D-7 5’-CGCAGCCGTCCAACCAAATCGTCC-3’
3’-CGTCGATCCAATGCCGTCC-5’

D-8 5’-CGCAGCCGTCCAACCAAATCGTCC-3’
3’-CGTCGATCCAATGCCGTCC-5’

D-12 5’-CGCAGCCGTCCAACCAAATCGTCC-3’
3’-CGTCGATCCAATGCCGTCC-5’

The downstream 19-mer primer was 5’-phosphorylated. The top strand was composed of two oligonucleotides (21-mer and 19-mer) with a single-nucleotide gap between them.

Table 8.1. DNA 21-19/41-mer substrates.
Table 8.2. Pre-steady state kinetic parameters of tPolλ.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (μM$^{-1}$s$^{-1}$)</th>
<th>Fidelity$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template A (D-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>3.6 ± 0.5</td>
<td>0.0010 ± 0.0008</td>
<td>2.8 × 10$^{-4}$</td>
<td>1.1 × 10$^{-4}$</td>
</tr>
<tr>
<td>dTTP</td>
<td>2.4 ± 0.4</td>
<td>6.0 ± 0.2</td>
<td>2.5</td>
<td>1</td>
</tr>
<tr>
<td>dGTP</td>
<td>8.4 ± 0.7</td>
<td>0.022 ± 0.001</td>
<td>2.6 × 10$^{-3}$</td>
<td>1.1 × 10$^{-3}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>5.4 ± 0.4</td>
<td>0.062 ± 0.003</td>
<td>1.1 × 10$^{-2}$</td>
<td>4.6 × 10$^{-3}$</td>
</tr>
<tr>
<td>Template G (D-6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>2.4 ± 0.7</td>
<td>0.0016 ± 0.0001</td>
<td>6.7 × 10$^{-4}$</td>
<td>2.5 × 10$^{-4}$</td>
</tr>
<tr>
<td>dTTP</td>
<td>1.8 ± 0.7</td>
<td>0.0115 ± 0.0007</td>
<td>6.3 × 10$^{-3}$</td>
<td>2.4 × 10$^{-3}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>3.2 ± 0.4</td>
<td>0.0021 ± 0.0008</td>
<td>6.5 × 10$^{-4}$</td>
<td>2.5 × 10$^{-4}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>1.1 ± 0.3</td>
<td>3.0 ± 0.2</td>
<td>2.7</td>
<td>1</td>
</tr>
<tr>
<td>Template T (D-7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>1.7 ± 0.2</td>
<td>3.87 ± 0.08</td>
<td>2.2</td>
<td>1</td>
</tr>
<tr>
<td>dTTP</td>
<td>1.9 ± 0.3</td>
<td>0.004 ± 0.001</td>
<td>2.1 × 10$^{-3}$</td>
<td>1.0 × 10$^{-3}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>2.9 ± 1.4</td>
<td>0.20 ± 0.02</td>
<td>7.1 × 10$^{-2}$</td>
<td>3.2 × 10$^{-2}$</td>
</tr>
<tr>
<td>dCTP</td>
<td>1.9 ± 0.6</td>
<td>0.015 ± 0.002</td>
<td>7.8 × 10$^{-3}$</td>
<td>3.5 × 10$^{-3}$</td>
</tr>
<tr>
<td>Template C (D-8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>1.4 ± 0.3</td>
<td>0.0046 ± 0.0001</td>
<td>3.2 × 10$^{-3}$</td>
<td>1.5 × 10$^{-3}$</td>
</tr>
<tr>
<td>dTTP</td>
<td>4.7 ± 0.5</td>
<td>0.0065 ± 0.0001</td>
<td>1.4 × 10$^{-3}$</td>
<td>6.4 × 10$^{-4}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>1.9 ± 0.4</td>
<td>4.1 ± 0.2</td>
<td>2.1</td>
<td>1</td>
</tr>
<tr>
<td>dCTP</td>
<td>1.5 ± 0.2</td>
<td>0.0098 ± 0.0002</td>
<td>6.4 × 10$^{-3}$</td>
<td>3.0 × 10$^{-3}$</td>
</tr>
<tr>
<td>Template C (D-12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>3.4 ± 0.9</td>
<td>0.0110 ± 0.0005</td>
<td>3.2 × 10$^{-3}$</td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>4.4 ± 1.4</td>
<td>0.0051 ± 0.0003</td>
<td>1.1 × 10$^{-3}$</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Calculated as $(k_p/K_d)_{\text{incorrect}}/[(k_p/K_d)_{\text{correct}} + (k_p/K_d)_{\text{incorrect}}]$

Table 8.2. Pre-steady state kinetic parameters of tPolλ.
Table 8.3. Different kinetic parameters of human DNA tPolλ and rat DNA Polβ in the presence of a single-nucleotide gapped DNA substrate.

<table>
<thead>
<tr>
<th></th>
<th>tPolλ</th>
<th>Polβ&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
<td>Incorrect</td>
</tr>
<tr>
<td>$K_d$(μM)</td>
<td>1.1-2.4</td>
<td>1.4-8.4</td>
</tr>
<tr>
<td>$k_p$ (s&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>3.0-6.0</td>
<td>0.001-0.20</td>
</tr>
<tr>
<td>$k_p/K_d$ (μM&lt;sup&gt;-1&lt;/sup&gt;s&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>2.1-2.7</td>
<td>(2.8-710)x10&lt;sup&gt;-4&lt;/sup&gt;</td>
</tr>
<tr>
<td>Fidelity&lt;sup&gt;b&lt;/sup&gt;</td>
<td>10&lt;sup&gt;-2&lt;/sup&gt; to 10&lt;sup&gt;-4&lt;/sup&gt;</td>
<td>10&lt;sup&gt;-4&lt;/sup&gt; to 10&lt;sup&gt;-5&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Ref. [24].

<sup>b</sup>Calculated as $(k_p/K_d)_{incorrect}/[(k_p/K_d)_{correct} + (k_p/K_d)_{incorrect}]$. 
Scheme 8.1. Minimal kinetic mechanism for dNTP incorporation into DNA.

\[
E \cdot DNA_n + dNTP \xrightarrow{k_{on}} E \cdot DNA_n \cdot dNTP \xrightarrow{k_p} E \cdot DNA_{n+1} \cdot PPi \xrightarrow{k_f} E + DNA_{n+1} + PPi
\]
8.6 References


A growing number of polymerases have been identified recently and many are hypothesized to function in DNA repair pathways. One of these polymerases is DNA polymerase λ (Polλ), a new member of the X-family DNA polymerases (1-3). This family of DNA polymerases is a subdivision of a larger superfamily of nucleotidyltransferases (4). Human DNA polymerase λ is encoded by a gene located in human chromosome 10 (2,3). The N-terminus of the full-length Polλ (fPolλ) is composed of a nuclear localization signal motif, a breast cancer susceptibility protein BRCA1 C-terminal (BRCT) domain, and a Proline-rich domain (Figure 9.1). BRCT domains are known to mediate protein/protein and protein/DNA interactions in DNA repair mechanisms and cell cycle check point regulation upon DNA damage (5). The Proline-rich domain, which contains multiple serine, threonine, and proline residues, is found to functionally suppress the polymerase activity of Polλ (6) and limit strand-displacement synthesis (7). The C-terminus of Polλ (tPolλ in Figure 9.1) possesses both 5’-deoxyribose-5-phosphate lyase (dRPase) and DNA polymerase activities while sharing 33% sequence identity with DNA polymerase β (Polβ) (1-3,8). Polβ, on the other hand, is also an X-family polymerase and is known to be involved in base excision repair (BER) pathways in vivo (9,10). The X-ray
crystal structures of tPol\(\lambda\) (11,12) display a high degree of similarity with the corresponding subdomains of Pol\(\beta\) (13,14), including the dRPase (also known as the 8-kDa), fingers, palm, and thumb subdomains. A comparison of the crystal structures of tPol\(\lambda\)•single-nucleotide gapped DNA, tPol\(\lambda\)•single-nucleotide gapped DNA•ddTTP, and tPol\(\lambda\)•nicked DNA•pyrophosphate suggests that no major protein domain movement occurs during catalysis (11). Pol\(\lambda\) like Pol\(\beta\) lacks 3’\(\rightarrow\)5’ exonuclease activity (1-3) and possesses low processivity when copying non-gapped DNA (8).

At present, the biological role of Pol\(\lambda\) is unknown. Pol\(\lambda\) has been suggested to play a role in DNA repair synthesis associated with meiosis (1), in “short-patch” BER (15,16), in the proliferating cell nuclear antigen-dependent BER pathway (15,17), and in the repair of double-stranded breaks through non-homologous end-joining pathways (15,18,19). All these potential physiological functions require Pol\(\lambda\) to possess gap-filling polymerase activity. The fidelity of Pol\(\lambda\) when filling medium to large gaps has been estimated to be in the range of 10\(^{-4}\) by both forward and reverse mutation assays (8,17,18). But the fidelity of the full-length Pol\(\lambda\) when filling single-nucleotide gapped DNA, which is relevant to BER, has not been determined. Previously, we measured the base substitution fidelity for filling single-nucleotide gapped DNA catalyzed by human tPol\(\lambda\) and found it to be in the range of 10\(^{-2}\) to 10\(^{-4}\), which is relatively low when compared to Pol\(\beta\) (15). However, the presence of the N-terminal BRCT and Proline-rich domains may have significant effects on the fidelity of Pol\(\lambda\). Here, we use pre-steady state kinetic methods to determine the base substitution fidelity of human fPol\(\lambda\) and its N-terminal domain truncation fragments (Figure 9.1) based on all possible dNTP incorporations into single-nucleotide gapped DNA. Any difference in the fidelity between the full-length and the truncated Pol\(\lambda\) fragments will reveal the role of the BRCT and Proline-rich domains in
polymerization efficiency and fidelity. Our kinetic results strongly support a role of Polλ as the second polymerase in BER, yet additional studies will be required to confirm this.

9.2 Materials and methods

Cloning, expression, and purification of fPolλ, dPolλ, and tPolλ. Preparation of tPolλ fused to a C-terminal His₆ tag (38.2 kDa) was described previously (15). The human genes encoding fPolλ and dPolλ (Figure 9.1) were PCR amplified, and separately inserted into the Nde I/Xho I sites of pET28b and pET24b to construct pET28b-fPolλ and pET24b-dPolλ. A stop codon was engineered before the Xho I site of pET28b-fPolλ. The constructed plasmids were individually transformed into E. coli strain BL21 CodonPlus (DE3)-RIL competent cells (Stratagene) to express fPolλ fused to both an N-terminal and a C-terminal hexahistidine tag, and dPolλ fused to a C-terminal hexahistidine tag. Both fPolλ (65.6 kDa) and dPolλ (50.1 kDa) were overexpressed in these transformed E. coli cells as was tPolλ (15). fPolλ was purified through a nickel-NTA column (Qiagen), a Heparin Sepharose Fast Flow column (GE Healthcare), a DEAE-Sepharose column (GE Healthcare), and a MonoS 10/10 column (GE Healthcare). dPolλ was purified through a nickel-NTA column (Qiagen), a Heparin Sepharose Fast Flow column, a DEAE-Sepharose column, and a HiTrap Q column (GE Healthcare). The His₆-tags of purified fPolλ and dPolλ were detected by Western blot analysis using anti-hexahistidine tag antibody (data not shown). The concentrations of the purified fPolλ and dPolλ were measured spectrophotometrically at 280 nm using the calculated extinction coefficients of 61,615 and 48,204 M⁻¹cm⁻¹, respectively.
**Synthetic oligodeoxyribonucleotides.** The oligodeoxyribonucleotides in Figure 9.2 were purchased from Integrated DNA Technologies and purified by denaturing polyacrylamide gel electrophoresis (18% acrylamide, 8 M urea). Their concentrations were determined by UV absorbance at 260 nm with calculated extinction coefficients. The 21-mer was 5’-end labeled by using Optikinase (USB) and [γ-32P]ATP (Perkin Elmer Life Sciences), and subsequently purified using a Biospin column (Bio-Rad Laboratories). Each single-nucleotide gapped DNA substrate was prepared by heating a mixture of 21-mer, 19-mer, and 41-mer, in a 1:1.25:1.15 molar ratio respectively for 8 min at 95 °C, then cooling the mixture slowly to room temperature over 3 hours as described previously (15).

**Reaction buffer L.** Buffer L contains 50 mM Tris-Cl (pH 8.4 at 37 °C), 5 mM MgCl2, 100 mM NaCl, and 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA. Buffer L was the optimized buffer used previously for studies probing the fidelity and mechanism of human tPolλ (15). All reactions reported in this paper were carried out in buffer L and at 37 °C. All concentrations refer to concentrations of components after mixing.

**Determination of the equilibrium dissociation constant (K_d) and the maximum rate constant of nucleotide incorporation (k_p) of an incoming nucleotide.** Both k_p and K_d were determined by mixing a preincubated solution of full-length Polλ and DNA at fixed concentrations with increasing concentrations of an incoming nucleotide (Gibco-BRL) in buffer L at 37 °C. These reactions were quenched at various times via the addition of 0.37 M EDTA. For the fast incorporation of a correct nucleotide, the reactions were carried out using a rapid chemical quench flow apparatus (KinTek). Aliquots of the quenched reactions were analyzed by sequencing gel analysis (17% acrylamide, 8 M urea, 1x TBE running buffer) and quantitated using a PhosphorImager 445 SI (Molecular
Dynamics). The resulting time courses of product formation were fit to a single exponential equation (equation 1) using the nonlinear regression program KaleidaGraph (Synergy Software) for each concentration of Mg$^{2+}$•dNTP to give the observed rate constant of nucleotide incorporation ($k_{obs}$). The corresponding observed incorporation rate constants were plotted against the concentrations of Mg$^{2+}$•dNTP and these data were fit to a hyperbola (equation 2) using KaleidaGraph to give $K_d$ and $k_p$.

\[
[\text{Product}] = A[1 - \exp(-k_{obs}t)] \quad \text{equation 1}
\]
\[
k_{obs} = k_p[dNTP]/([dNTP] + K_d) \quad \text{equation 2}
\]

### 9.3 Results

**Fidelity of full-length human Polλ.** Previously, we established a minimal kinetic mechanism (Scheme 9.1) for the single-nucleotide gap-filling activity of tPolλ (15). Scheme 9.1 shows that an incoming deoxyribonucleotide (dNTP) binds to the E•DNA binary complex to establish a rapid equilibrium prior to nucleotide incorporation. We expect the full-length Polλ to have similar polymerase activity as tPolλ and thus follow the same minimal mechanism shown in Scheme 9.1. This mechanism allows us to measure the apparent affinity of dNTP ($K_d$) for the fPolλ•DNA binary complex via the dNTP concentration dependence of the observed single-turnover rate constant ($k_{obs}$). Our results shown below confirm this hypothesis. The single-turnover method was employed because DNA dissociates from tPolλ with a dissociation rate constant ($k_i$) which is about 2- to 3-fold slower than the maximum nucleotide incorporation rate constant $k_p$ (15),
rendering the burst phase (amplitude = \[k_p/(k_p + k_i)\]^2) small. Thus, the experiments were performed with fPolλ in molar excess over DNA (or under single-turnover conditions) in order to avoid complication from the steady-state reaction phase (20). A preincubated solution of 5'-[\textsuperscript{32}P]-labeled D-1 (Figure 9.2) and 4-fold fPolλ was reacted with increasing concentrations of correct dTTP in buffer L. The DNA product 22-mer and remaining primer 21-mer at different time intervals were separated and quantitated. The product concentration was plotted against reaction time intervals. These data were subsequently fit to equation 1, to yield a single-turnover rate constant at each concentration of dTTP (Figure 9.3). The single-turnover rates were then plotted against dTTP concentrations (Figure 9.4). These data were subsequently fit to equation 2 (Materials and Methods), to yield a \(k_p\) of 3.9 ± 0.2 s\(^{-1}\) for the maximum dTTP incorporation rate constant, and a \(K_d\) of 2.6 ± 0.4 μM for dTTP binding. The substrate specificity \(k_p/K_d\) of dTTP incorporation into D-1 was calculated to be 1.5 x 10\(^6\) M\(^{-1}\)s\(^{-1}\) (Table 9.1).

Similar pre-steady state kinetic analyses were used to determine the kinetic parameters (Table 9.1) for the incorporation of each of the remaining three correct nucleotides (dCTP into D-6, dATP into D-7, and dGTP into D-8), and for the incorporation of each of the 12 possible incorrect incorporations into D-1, D-6, D-7, and D-8 (Figure 9.2). Notably, the ground-state binding affinity of all nucleotides to the fPolλ•DNA (Table 9.1) and the tPolλ•DNA (15) binary complexes were similar, but the maximum incorporation rate constants, particularly for incorrect nucleotides, with fPolλ (Table 9.1) were slower than the corresponding parameters for tPolλ measured previously (15). The substrate specificity \(k_p/K_d\) of all four correct nucleotides was slightly lower for fPolλ when compared to tPolλ (Figure 9.5) with a nucleotide incorporation efficiency ratio, \((k_p/K_d)_{\text{tPolλ}}/(k_p/K_d)_{\text{fPolλ}}\), less than 2 (Table 9.1). In contrast, fPolλ incorporated incorrect nucleotides with much lower substrate specificities than tPolλ (Table 9.1 and Figure 9.6).
The nucleotide incorporation efficiency ratio, \( \frac{k_p}{K_d}_{t\text{Pol}}/\frac{k_p}{K_d}_{f\text{Pol}} \), was calculated to be in the range of 10 to 100 for all twelve incorrect basepairs (Table 9.1). The variation in this ratio was due to the sequence-dependence of polymerization catalyzed by DNA polymerases. Nevertheless, the large ratios suggest that fPol\( \lambda \) achieves significantly higher fidelity than tPol\( \lambda \) by lowering the incorporation efficiency of mismatched nucleotides. To confirm this possibility, we calculated the fidelity of fPol\( \lambda \) which was sequence-dependent and in the range of 10\(^{-4}\)-10\(^{-5}\) (Table 9.1). In comparison, the fidelity of tPol\( \lambda \) with the same single-nucleotide gapped DNA is also sequence-dependent and in the range of 10\(^{-2}\)-10\(^{-4}\) (15). Therefore, the full-length Pol\( \lambda \) is 10- to 100-fold more faithful than the C-terminal fragment tPol\( \lambda \) in filling single-nucleotide gapped DNA. To ensure the fidelity difference between fPol\( \lambda \) and tPol\( \lambda \) was not due to experimental errors, we repeated above fidelity and kinetic measurements with both fPol\( \lambda \) and tPol\( \lambda \) individually, and obtained almost identical fidelity values (data not shown) to those corresponding values listed in Table 1 from our previously publication (15). The significant increase in fidelity from tPol\( \lambda \) to fPol\( \lambda \) is probably due to the presence of the BRCT domain, the Proline-rich domain, or both.

**Fidelity of dPol\( \lambda \).** To identify which of the N-terminal domain(s) up-regulates the fidelity of the full-length Pol\( \lambda \), we engineered and purified a fragment, dPol\( \lambda \) (residues 132-575, Figure 9.1), which lacks the BRCT domain. With two diagnostic substrates D-6 and D-7, the kinetic parameters and fidelity of dPol\( \lambda \) were determined in the same manner as described above and listed in Table 9.2. dPol\( \lambda \) incorporated nucleotides into D-6 and D-7 with a fidelity in the range of 10\(^{-4}\)-10\(^{-5}\), which was very similar to the fidelity of fPol\( \lambda \) (Table 9.1). The nucleotide incorporation efficiency ratios, \( \frac{k_p}{K_d}_{d\text{Pol}}/\frac{k_p}{K_d}_{f\text{Pol}} \), were calculated to be in the range of 0.5 to 2.1 (Table 9.2 and Figure 9.7). The average ratio for all eight possible nucleotide incorporations into both D-6 and D-7 was calculated to
be 1.2. These ratios indicate that the substrate specificity of nucleotides with dPolλ, 
\( \frac{k_p}{K_d} \), whether slightly larger or smaller, was within 2-fold of the substrate 
specificity of corresponding nucleotides with fPolλ, \( \frac{k_p}{K_d} \). Because 2-fold 
differences are considered kinetically insignificant (20), we concluded that dPolλ and 
fPolλ have very similar, if not identical, gap-filling efficiency. Thus, the deletion of the 
BRCT domain in dPolλ did not affect the nucleotide incorporation efficiency and the 
fidelity of fPolλ. Similar polymerization efficiency and fidelity of dPolλ and fPolλ also 
indirectly confirmed that the N-terminal hexahistidine tag only present in the 
recombinant fPolλ did not affect the fidelity of fPolλ. Comparison of the fidelity of fPolλ 
(\(10^{-4}\) to \(10^{-5}\), Table 9.1), dPolλ (\(10^{-4}\) to \(10^{-5}\), Table 9.2), and tPolλ (\(10^{-2}\) to \(10^{-4}\)) (15) 
reveals an intriguing conclusion: the non-enzymatic Proline-rich domain alone 
significantly enhanced the fidelity of human Polλ.

9.4 Discussion

DNA polymerases λ and β share high sequence homology (54%) and sequence identity 
(33%) (1-3,8). Polλ, like Polβ, possesses the two key enzymatic activities (gap-filling 
polymerase and dRPase) required by BER. The role of Polλ in DNA repair is further 
supported by the following observations: (i) like Polβ (21), Polλ is expressed at high 
levels in the developing mouse testes, suggesting a possible function of Polλ in DNA 
repair pathways associated with meiotic recombination (1); (ii) in an in vitro BER 
reconstitution assay, recombinant human Polλ and Polβ can replace each other to 
efficiently repair uracil-containing DNA in the presence of human uracil-DNA 
glycosylase, human AP endonuclease, and human DNA ligase I (22); (iii) Polλ is the
only X-family DNA polymerase found in higher plants and is induced by DNA-damaging treatments (23); (iv) mouse embryonic fibroblast cell extract contains substantial amounts of active Polλ, which contributes to uracil-initiated short-patch BER (16); (v) monoclonal antibodies against Polλ strongly reduce in vitro BER in the Polβ<sup>−/−</sup> cell extract (16); and (vi) Polλ protects mouse fibroblasts against oxidative DNA damage and is recruited to oxidative DNA damage sites (24). Thus, Polλ may compliment or support the function of Polβ in BER in vivo. But this hypothesized role was initially weakened by the observed difference in single-nucleotide gap filling fidelity between Polβ and tPolλ. The fidelity of tPolλ (10<sup>−2</sup> to 10<sup>−4</sup>) (15) is about 10- to 100-fold lower than the fidelity of rat Polβ (recalculated as 10<sup>−4</sup>-10<sup>−5</sup> using the definition of fidelity in Table 9.1) which was measured under similar single-turnover conditions with four single-nucleotide gapped DNA 25-19/45-mer (primer-primer/template) substrates (25). If the full-length Polλ had similar fidelity to tPolλ, it would make 10- to 100-fold more base substitution errors than Polβ if it participated in the repair of single-base lesions. DNA base modifications/losses are known to account for a large portion of total cellular DNA damage.

Under single-turnover conditions, the deoxyribonucleotide incorporation fidelity of human fPolλ was measured in the presence of four different single-nucleotide gapped DNA substrates shown in Figure 9.2. Surprisingly, the fidelity of fPolλ was determined to be in the range of 10<sup>−4</sup> to 10<sup>−5</sup> (Table 9.1) and was indeed significantly higher than that of tPolλ (10<sup>−2</sup> to 10<sup>−4</sup>) (15). This range for fPolλ was identical to the fidelity range determined for Polβ (26). In addition, the efficiency for correct nucleotide incorporation into single-nucleotide gapped DNA catalyzed by fPolλ (1.2 x 10<sup>6</sup> - 1.8 x 10<sup>6</sup> M<sup>−1</sup>s<sup>−1</sup>, Table 9.1) was only slightly lower than Polβ (4 x 10<sup>6</sup> - 6 x 10<sup>6</sup> M<sup>−1</sup>s<sup>−1</sup>) (26). Similar polymerization efficiency and fidelity in combination with other biological evidence discussed above strongly support the hypothesis that Polλ and Polβ both possess similar
in vivo roles, such as functioning in BER. So far, the generation of knock-out mice through deletion of exons 5 to 7 of the Polλ gene has not yet confirmed the involvement of Polλ in BER or any other biological processes (27). The published mouse Polλ knock-out experiments (27,28) are likely complicated by the existence of Polβ (28), which could fill in and compensate for the loss of functional Polλ.

The measured fidelity of fPolλ (Table 9.1) was 10- to 100-fold higher than the fidelity of tPolλ (10⁻² to 10⁻⁴), estimated by employing the same single-turnover kinetic assay with identical DNA substrates (15). This indicated that the N-terminal domains of Polλ (Figure 9.1) significantly enhance the fidelity of Polλ. Because tPolλ possesses slightly higher correct nucleotide incorporation efficiency than fPolλ (Table 9.1) and both enzymes bind to all correct and incorrect nucleotides with similarly high affinity (Table 9.1 and reference (15)), it is unlikely that the absence of the N-terminal domains will cause any type of misfolding of tPolλ and thus contribute to its low fidelity. This assumption was validated based on the following structural evidence: (i) when the high resolution tPolλ binary structure (2.1 Å) (12) and the full-length Polβ ternary structure (2.2 Å) (14) are overlaid, the α carbons superimposed very well with an rms deviation of 1.4 Å for 113 C-α atoms (12), suggesting that tPolλ folds well and similarly to Polβ in the absence of the two N-terminal domains of Polλ; (ii) no unfolded regions in tPolλ are observed in its 1.95 to 2.3 Å binary (12) and ternary crystal structures (11). Interestingly, the average ratio of \((k_p/K_d)_{dPol\lambda}/(k_p/K_d)_{fPol\lambda}\) for all eight possible nucleotide incorporation with D-6 and D-7 was 1.2 (Figure 9.7), indicating that dPolλ and fPolλ almost have equal nucleotide incorporation efficiency. As expected, these two enzymes also have very similar fidelity (Tables 9.1 and 9.2). Thus, the deletion of the BRCT domain affected neither the single-nucleotide gap filling efficiency nor the fidelity of Polλ, while differences in fidelity between dPolλ and tPolλ demonstrated that the non-enzymatic
Proline-rich domain dramatically enhanced the fidelity of Pol\(\lambda\) by up to a 100-fold. This conclusion is consistent with the proposed function of the BRCT domain which is generally known to mediate the protein/protein or protein/DNA interactions required for particular DNA repair pathways and cellular responsiveness to DNA damage (1-3,8). Our conclusion is also supported by the observation of Shimazaki et al. in which they have found that deletion of the BRCT domain of human Pol\(\lambda\) did not affect DNA synthesis with DNA substrates including poly(dA)/oligo(dT) and activated DNA (6). The Proline-rich domain has been proposed to merely couple polymerase action to the protein-protein and protein-DNA interactions required during DNA repair (1). Thus, the Proline-rich domain is generally assumed to be dispensable for the polymerase activity of Pol\(\lambda\). However, the difference in fidelity between tPol\(\lambda\) and dPol\(\lambda\) observed here suggested that the flexible Proline-rich domain actively regulates the polymerase fidelity. Enhancement of the fidelity of a polymerase by a non-enzymatic domain is both surprising and unprecedented. Therefore, the 11-kDa Proline-rich domain must interact with the subdomains of tPol\(\lambda\) and optimize the geometry of the polymerase active site to achieve higher fidelity. The fidelity modulation observed here differs dramatically from the fidelity enhancement contributed by the enzymatic activity of a 3’→5’ exonuclease domain of a replicative DNA polymerase, or by an accessory protein as in the case of the mitochondrial DNA polymerase complex (29). Amazingly, the enhancement of the Pol\(\lambda\) fidelity by the Proline-rich domain is as large as what has been contributed by the proof-reading 3’→5’ exonuclease domain to a replicative DNA polymerase (29,30). Our results further suggest that DNA polymerases, the vital enzymes that replicate and maintain genomic DNA, have evolved through different mechanisms to adjust their polymerization fidelity in order to best perform diverse physiological functions.
Interestingly, the fidelity difference between tPol\(\lambda\) and fPol\(\lambda\) resulted mainly from stronger discrimination against mismatched nucleotides by fPol\(\lambda\). The average substrate specificity of incorrect incorporations was 10- to 100-fold lower with fPol\(\lambda\) than with tPol\(\lambda\) (Figure 9.6). In comparison, fPol\(\lambda\) incorporated correct dNTPs with an efficiency \((k_p/K_d)\) of \((1.2 - 1.8) \times 10^6\text{ M}^{-1}\text{s}^{-1}\) (Table 9.1), which was only about two-fold lower than the corresponding range of \((2.1 - 2.7) \times 10^6\text{ M}^{-1}\text{s}^{-1}\) with tPol\(\lambda\) (Figure 9.5). The larger decrease in the incorporation efficiency of mismatched over matched dNTPs (Table 9.1 and Figures 9.5 and 9.6) with fPol\(\lambda\) led to the higher fidelity of Pol\(\lambda\). In comparison, the accessory subunit enhances the fidelity of mitochondrial DNA polymerase by 14-fold by increasing the incorporation efficiency of a correct A:T basepair more than a T:T mismatch (29). Furthermore, the decrease in the incorporation efficiency of both matched and mismatched dNTPs was due to slower incorporation rate constants \((k_p)\) with fPol\(\lambda\) than with tPol\(\lambda\) since the binding of all dNTPs \((K_d)\) by these two enzymes was similarly tight (Table 9.1 and reference (15)). This suggested fPol\(\lambda\) was slower yet more faithful than tPol\(\lambda\) in filling single-nucleotide gapped DNA. This correlation goes against the general trend summarized from a survey of the A-, B-, X-, and Y-families by Beard et al. that a more catalytically efficient polymerase has a lower base substitution rate (31). It is not clear what contributes to this intriguing correlation between the polymerase fidelity and nucleotide incorporation rate. We speculate that the presence of the Proline-rich domain in both fPol\(\lambda\) and dPol\(\lambda\) may either somewhat tighten the polymerase active site to achieve better geometric selection (32,33), or shield the active site of Pol\(\lambda\) from solvent which leads to greater desolvation of a nascent basepair and amplifies the free energy differences between matched and mismatched nucleotide incorporations (34). These possibilities can be evaluated by the active site structural differences between the ternary crystal structures of fPol\(\lambda\) (or dPol\(\lambda\))•single-nucleotide gapped DNA•dNTP and
tPol λ • single-nucleotide gapped DNA • dNTP (11), especially if one were to crystallize these aforementioned complexes in the presence of a mismatched incoming nucleotide.
9.5 Figures, Tables, and Schemes

Figure 9.1. Schematic representations of human fPolλ, dPolλ, tPolλ, and Polβ. Each domain, with amino acid residue numbers indicated above, is shown as a rectangle. The N-terminal thirty-five residues of fPolλ containing a nuclear localization signal motif are shown as a line.
Figure 9.2. Single-nucleotide gapped DNA substrates D-1, D-6, D-7, and D-8. Primer 21-mer was 5’-32P-labeled. The downstream primer 19-mer was 5’-phosphorylated (P).

“X” represents the unpaired base of template 41-mer.
Figure 9.3. dTTP incorporation into 5'-[\(^{32}\text{P}\)]-labeled D-1. Concentration dependence of the pre-steady state rate of correct deoxyribonucleotide incorporation. A preincubated solution of Pol\(\lambda\) (120 nM) and \(^{32}\text{P}\)-labeled-D-1 (30 nM) was rapidly mixed with increasing concentrations of Mg\(^{2+}\)-dTTP (0.1 \(\mu\)M, \(\bullet\); 0.25 \(\mu\)M, \(\bigcirc\); 0.5 \(\mu\)M, \(\blacksquare\); 1 \(\mu\)M, \(\square\); 3 \(\mu\)M, \(\blacktriangle\); 5 \(\mu\)M, \(\triangle\); 10 \(\mu\)M, \(\blacktriangledown\); 25 \(\mu\)M, \(\bigtriangledown\); 50 \(\mu\)M, \(\lozenge\)) for various time intervals. The solid lines are the best fits to the single exponential equation (equation 1).
Figure 9.4. dTTP incorporation into 5'-[\(^{32}\)P]-labeled D-1. The single exponential rates obtained from the above data fitting (Figure 9.3) were plotted as a function of dTTP concentration. The extracted rate data were then fit to the hyperbolic equation (equation 2) yielding a \(k_p\) of 3.9 ± 0.2 s\(^{-1}\) and a \(K_d\) of 2.6 ± 0.4 \(\mu\)M.
Figure 9.5. Comparison of the substrate specificity \( \frac{k_p}{K_d} \) for nucleotide incorporation catalyzed by fPol\(\lambda\) (Table 9.1) and tPol\(\lambda\) (15) separately. Four correct incorporations. The Y-axis is in the log scale.
Figure 9.6. Comparison of the substrate specificity \( \frac{k_p}{K_d} \) for nucleotide incorporation catalyzed by fPol\( \lambda \) (Table 9.1) and tPol\( \lambda \) (15) separately. Twelve incorrect incorporations. The Y-axis is in the log scale.
Figure 9.7. Both dPol\(\lambda\) (Table 9.2) and fPol\(\lambda\) (Table 9.1) incorporate nucleotides with similar efficiency with an average ratio of \(\frac{(k_p/K_d)_{dPol\lambda}}{(k_p/K_d)_{fPol\lambda}}\) equal to 1.2 (indicated by a line) for all nucleotide incorporations into D-6 and D-7.
Table 9.1. Kinetic parameters of nucleotide incorporation into single-nucleotide gapped DNA catalyzed by the full-length Polλ at 37 °C.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (M$^{-1}$s$^{-1}$)</th>
<th>Fidelity$^a$</th>
<th>Efficiency Ratio$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>2.6 ± 0.4</td>
<td>3.9 ± 0.2</td>
<td>1.5 x 10$^6$</td>
<td>1</td>
<td>1.9</td>
</tr>
<tr>
<td>dATP</td>
<td>1.86 ± 0.09</td>
<td>0.000034 ± 0.000001</td>
<td>1.8 x 10$^1$</td>
<td>1.2 x 10$^{-5}$</td>
<td>16</td>
</tr>
<tr>
<td>dCTP</td>
<td>10.4 ± 0.3</td>
<td>0.0052 ± 0.0009</td>
<td>5.0 x 10$^2$</td>
<td>3.3 x 10$^{-4}$</td>
<td>22</td>
</tr>
<tr>
<td>dGTP</td>
<td>3.2 ± 0.5</td>
<td>0.00040 ± 0.00002</td>
<td>1.3 x 10$^2$</td>
<td>8.7 x 10$^{-5}$</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>0.9 ± 0.1</td>
<td>1.57 ± 0.04</td>
<td>1.8 x 10$^6$</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>dATP</td>
<td>3 ± 1</td>
<td>0.00010 ± 0.00004</td>
<td>3.3 x 10$^1$</td>
<td>1.8 x 10$^{-5}$</td>
<td>20</td>
</tr>
<tr>
<td>dTTP</td>
<td>2.5 ± 0.3</td>
<td>0.00070 ± 0.00001</td>
<td>2.8 x 10$^2$</td>
<td>1.6 x 10$^{-4}$</td>
<td>23</td>
</tr>
<tr>
<td>dGTP</td>
<td>4 ± 1</td>
<td>0.00020 ± 0.00001</td>
<td>5.0 x 10$^1$</td>
<td>2.8 x 10$^{-5}$</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>0.9 ± 0.3</td>
<td>1.5 ± 0.1</td>
<td>1.6 x 10$^6$</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>dTTP</td>
<td>8.4 ± 0.6</td>
<td>0.0002 ± 0.0001</td>
<td>2.4 x 10$^1$</td>
<td>1.5 x 10$^{-5}$</td>
<td>88</td>
</tr>
<tr>
<td>dCTP</td>
<td>5.4 ± 0.2</td>
<td>0.004 ± 0.002</td>
<td>7.0 x 10$^2$</td>
<td>4.4 x 10$^{-4}$</td>
<td>11</td>
</tr>
<tr>
<td>dGTP</td>
<td>7 ± 4</td>
<td>0.010 ± 0.002</td>
<td>1.4 x 10$^3$</td>
<td>8.8 x 10$^{-4}$</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>2.1 ± 0.3</td>
<td>2.5 ± 0.1</td>
<td>1.2 x 10$^6$</td>
<td>1</td>
<td>1.8</td>
</tr>
<tr>
<td>dATP</td>
<td>1.5 ± 1.0</td>
<td>0.0003 ± 0.0003</td>
<td>1.7 x 10$^2$</td>
<td>1.4 x 10$^{-4}$</td>
<td>19</td>
</tr>
<tr>
<td>dCTP</td>
<td>4.7 ± 0.3</td>
<td>0.002 ± 0.002</td>
<td>3.8 x 10$^2$</td>
<td>3.2 x 10$^{-4}$</td>
<td>17</td>
</tr>
<tr>
<td>dTTP</td>
<td>3.4 ± 0.1</td>
<td>0.00047 ± 0.00001</td>
<td>1.4 x 10$^2$</td>
<td>1.2 x 10$^{-4}$</td>
<td>10</td>
</tr>
</tbody>
</table>

$^a$Calculated as $(k_p/K_d)_{Incorrect}/[(k_p/K_d)_{Correct} + (k_p/K_d)_{Incorrect}]$

$^b$Calculated as $(k_p/K_d)_{tPol\lambda}/(k_p/K_d)_{Pol\lambda}$. The $k_p/K_d$ values for tPolλ are from Table 2 of reference (15).
Table 9.2. Kinetic parameters of nucleotide incorporation into single-nucleotide gapped D-6 and D-7 catalyzed by dPolλ at 37 °C.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s⁻¹)</th>
<th>$k_p/K_d$ (M⁻¹s⁻¹)</th>
<th>Fidelitya</th>
<th>Efficiency Ratiob</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Template G (D-6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>1.1 ± 0.1</td>
<td>1.7 ± 0.2</td>
<td>1.6 x 10⁶</td>
<td>1</td>
<td>0.89</td>
</tr>
<tr>
<td>dATP</td>
<td>3.2 ± 0.9</td>
<td>0.00010 ± 0.00002</td>
<td>3.3 x 10¹</td>
<td>2.1 x 10⁻⁵</td>
<td>1.0</td>
</tr>
<tr>
<td>dTTP</td>
<td>1.8 ± 0.8</td>
<td>0.00030 ± 0.00007</td>
<td>1.7 x 10²</td>
<td>1.1 x 10⁻⁴</td>
<td>0.61</td>
</tr>
<tr>
<td>dGTP</td>
<td>3.3 ± 0.5</td>
<td>0.00022 ± 0.00002</td>
<td>6.7 x 10¹</td>
<td>4.2 x 10⁻⁵</td>
<td>1.3</td>
</tr>
<tr>
<td>Template T (D-7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>1.0 ± 0.4</td>
<td>1.8 ± 0.2</td>
<td>1.8 x 10⁶</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>dTTP</td>
<td>11 ± 6</td>
<td>0.00046 ± 0.00005</td>
<td>4.2 x 10¹</td>
<td>2.3 x 10⁻⁵</td>
<td>1.7</td>
</tr>
<tr>
<td>dCTP</td>
<td>4 ± 2</td>
<td>0.006 ± 0.001</td>
<td>1.5 x 10³</td>
<td>8.3 x 10⁻⁴</td>
<td>2.1</td>
</tr>
<tr>
<td>dGTP</td>
<td>5 ± 1</td>
<td>0.0087 ± 0.0001</td>
<td>1.7 x 10³</td>
<td>9.4 x 10⁻⁴</td>
<td>1.2</td>
</tr>
</tbody>
</table>

aCalculated as $(k_p/K_d)_{incorrect} / [(k_p/K_d)_{correct} + (k_p/K_d)_{incorrect}]$

bCalculated as $(k_p/K_d)_{dPolλ} / (k_p/K_d)_{fPolλ}$.
Scheme 9.1. Minimal mechanism for dNTP incorporation into a DNA substrate.
9.6 References


10.1 Introduction

In mammalian cells, single-base lesions are the most common form of DNA damage that arises either from exogenous DNA-damaging agents (1), or from endogenous biological processes resulting in base alkylation (2,3), base oxidation (4), spontaneous cytosine deamination (3), and hydrolytic base loss (3,5,6). Base excision repair (BER) is the major pathway to repair single-base lesions (7). Short-patch and long-patch BER are the two subpathways of BER which remove and replace one (8-10) and 2 to 11 nucleotides (11-13), respectively. Short-patch BER starts with the excision of a modified base by a DNA glycosylase, leaving a noncoding apurinic or apyrimidinic (AP) site in DNA. This lesion is further processed and repaired by a 5'-acting AP endonuclease, a DNA polymerase, a 5'-2-deoxyribose-5-phosphate lyase (dRPase), and a DNA ligase (12,14-16). It has been established that in mammalian systems, DNA polymerase β (Polβ, Figure 10.1), an X-family DNA polymerase, plays a critical role in short-patch BER (8,9). The polymerase activity of Polβ catalyzes single-nucleotide gap-filling synthesis (17) while its dRPase activity removes the 5'-terminal 2-deoxyribose-5-phosphate moiety (dRP) of a downstream strand (18). The uracil-initiated short-patch BER has been reconstituted in
vitro by using purified recombinant human enzymes (19-21) and its reaction sequences (Scheme 10.1) have been relatively well established using steady-state kinetic studies (21).

Like Polβ, the full-length DNA polymerase λ (fPolλ), a recently discovered member of the X-family DNA polymerases, also contains a dRPase (8 kDa) and a DNA polymerase domain (31 kDa) on its C-terminus (Figure 10.1) (22-25). In addition, fPolλ possesses a nuclear localization signal motif, a breast cancer susceptibility gene 1 C terminal (BRCT) domain, and a Proline-rich domain. Although the biological role of fPolλ has not been clearly identified, it is plausible that fPolλ contributes to BER since it shares 33% sequence identity to Polβ and possesses two key enzymatic activities required by BER. This hypothesis is directly supported by the following observations: (i) recombinant human fPolλ purified from *E. coli* can replace human Polβ in an in vitro reconstituted short-patch BER assay (26); (ii) mouse embryonic fibroblast Polβ−/− cell extract contains a substantial amount of active fPolλ which can also replace Polβ in a similar in vitro reconstituted BER assay, and monoclonal antibodies against fPolλ in this cell extract strongly reduce in vitro BER (27). fPolλ, like Polβ, lacks 3’→5’ exonuclease activity (22-24) and has low processivity when copying non-gapped or large-gap DNA (25). With short-gap DNA, the downstream strand, especially one with a 5’-phosphate, has been shown to increase polymerase processivity of both Polβ (28) and fPolλ (25,29).

Moreover, Polβ has been found to incorporate a single nucleotide about 10-fold more efficiently with single-nucleotide gapped DNA than non-gapped DNA (30,31), but the polymerase fidelity is not altered (31). Structural evidence suggests that the increase in polymerase processivity and efficiency is due to additional contacts that are established in a gapped DNA substrate between the dRPase domains of these two enzymes and the downstream strand (29,32). The terminal 5’-phosphate on the downstream strand is
buried in a positively charged pocket of the dRPase active site (29,32,33). So far, the effects of a downstream strand and its 5’-phosphate on the efficiency and fidelity of gap-filling synthesis catalyzed by fPol\(\lambda\) has not been quantitatively determined. In this chapter, these effects associated with human fPol\(\lambda\) will be investigated through pre-steady state kinetic studies.

Moreover, steady-state kinetic studies have identified the rate-limiting step in the reconstituted and uracil-initiated BER system (Scheme 10.1) being the removal of the dRP moiety by Pol\(\beta\) (21). fPol\(\lambda\) is estimated to possess a 4-fold slower dRPase activity than Pol\(\beta\) (26) while these two enzymes catalyze single-nucleotide gap-filling DNA synthesis with less than a 2-fold difference in catalytic efficiency (34). These kinetic data strongly indicate that 5’-dRP moiety in the downstream strand is removed after Pol\(\beta\) or fPol\(\lambda\) fills single-nucleotide gapped DNA in Scheme 10.1. However, the effect of the 5’-dRP moiety on the fidelity and efficiency of either Pol\(\beta\) or fPol\(\lambda\) has never been kinetically evaluated. In this chapter, we will use single-nucleotide gapped DNA substrates containing a dRP-mimic on the 5’-terminus of the downstream strand to determine the kinetic effect of the 5’-dRP moiety on gap-filling DNA synthesis catalyzed by human fPol\(\lambda\) through detailed pre-steady state kinetic analysis.

10.2 Materials and methods
Materials. These chemicals were purchased from the following companies: \([\gamma-^{32}\text{P}]\text{ATP},\) GE Healthcare (Piscataway, NJ); Biospin columns, Bio-Rad Laboratories (Hercules, CA); dNTPs, Gibco-BRL (Rockville, MD); T4 polynucleotide kinase, USB (Cleveland, OH). The full-length human Pol\(\lambda\) was cloned, expressed, and purified as described previously (34).

Synthetic oligonucleotides. The DNA substrates shown in Figure 10.2 were purchased from Integrated DNA Technologies (Coralville, IA), and purified by denaturing polyacrylamide gel electrophoresis (18% acrylamide, 8 M urea). Their concentrations were determined UV absorbance at 260 nm with calculated molar extinction coefficients. The primer strand 21-mer was 5'-32P labeled by incubation with T4 polynucleotide kinase and \([\gamma-^{32}\text{P}]\text{ATP}\) for 1 hour at 37 \(^\circ\)C. The unreacted \([\gamma-^{32}\text{P}]\text{ATP}\) was subsequently removed by centrifugation through a Biospin-6 column (Bio-Rad). The 5'-32P labeled primer 21-mer was then annealed with the corresponding non-radiolabeled downstream strand 19-mer to the template 41-mer at a molar ratio of 1.0:1.25:1.15 respectively, to form a 21-19/41-mer single-nucleotide gapped substrate (the top strand was composed of two oligonucleotides with a single-nucleotide gap). Mixtures to be annealed were denatured at 95 \(^\circ\)C for 8 min, and then cooled slowly to room temperature over several hours.

Optimized reaction buffer \(L\). All kinetic experiments were performed in buffer \(L\) containing 50 mM Tris-Cl (pH 8.4 at 37 \(^\circ\)C), 5 mM MgCl\(_2\), 100 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA. All reactions were carried out at 37 \(^\circ\)C.

Measurement of equilibrium dissociation constant of next incoming nucleotide. A preincubated solution of 150 nM human fPol\(\lambda\) and 30 nM \([^{32}\text{P}]\)-labeled DNA was rapidly
mixed with increasing concentrations of dNTP (0.25 to 120 μM) in buffer L in a rapid chemical quench-flow apparatus (KinTek, PA) to initiate the reaction. The reactions at each concentration of dNTP were terminated with 0.37 M EDTA at varying times from milliseconds to minutes. Reaction products were analyzed by sequencing gel electrophoresis (17% acrylamide, 8 M urea, 1x TBE buffer) and quantitated with a PhosphorImager 445 SI (Molecular Dynamics). The time course of product formation was fit to a single exponential equation (equation 1) for each concentration of dNTP to give the observed rate constant of nucleotide incorporation ($k_{\text{obs}}$). “A” represents the reaction amplitude which is equal to the initial concentration of the enzyme and DNA binary complex. The observed rate constants extracted from these time courses of product formation were then plotted against the concentrations of dNTP and these data were fit via hyperbolic regression (equation 2) to give the equilibrium dissociation constant of dNTP, $K_d$, and the maximum rate constant for incorporation of dNTP, $k_p$. The substrate specificity and polymerase fidelity were calculated as $(k_p/K_d)$ and $(k_p/K_d)_{\text{incorrect}}/[(k_p/K_d)_{\text{correct}} + (k_p/K_d)_{\text{incorrect}}]$, respectively.

\[
\text{[Product]} = A[1 - \exp(-k_{\text{obs}} t)] \quad \text{(equation 1)}
\]

\[
k_{\text{obs}} = k_p[d\text{NTP}]/\{[d\text{NTP}] + K_d\} \quad \text{(equation 2)}
\]
10.3 Results

*Model DNA substrates.* Based on steady-state kinetic results (21), the natural DNA substrate for the polymerase activity of tPolλ in BER is a dRP-DNA type substrate (Figure 10.2) which exists as a mixture of different chemical species in solution (see Discussion) (35). To simplify the kinetic analysis, we prepared four H-DNA substrates shown in Figure 10.3, whose downstream strands contained a 5’-terminal 1,2-dideoxyribose-5-phosphate moiety which mimicked 5’-dRP. The H-DNA substrates are stable and exist in one chemical form in solution due to the lack of C1-hydroxyl group in their dRP-mimic moiety. To examine the kinetic effect of the downstream strand and its 5’-phosphate, we designed gapped D-1OH (Figure 10.5). This substrate contained a regular downstream 19-mer that was not 5’-phosphorylated as in D-DNA (D-1) (Figure 10.4), and non-gapped D-1N (Figure 10.6), which lacked a downstream strand.

*Kinetic effect of a dRP-mimic.* Previously, we have established a minimal kinetic mechanism (Scheme 10.2) for dTTP incorporation onto D-DNA (D-1) (Figure 10.4) catalyzed by the C-terminal Polβ-like domain of human Polλ (tPolλ, Figure 10.1) (36). This scheme shows that an incoming dNTP binds to the tPolλ•D-1 binary complex to establish a rapid equilibrium prior to nucleotide incorporation. We have also demonstrated that human fPolλ follows the same minimal mechanism shown in Scheme 10.2 (34). This mechanism allows us to measure the apparent affinity of dNTP ($K_d$) for the Polλ•DNA binary complex via the dNTP concentration dependence of the observed single-turnover rate constant ($k_{obs}$). The polymerase efficiency and fidelity of both tPolλ (36) and the fPolλ (34) have thus been determined. Here, we utilized similar pre-steady
state kinetic methods to analyze and determine the kinetic parameters of all sixteen possible single-nucleotide incorporations onto the four DNA substrates (see Figure 10.3).

Under single-turnover conditions, a preincubated solution of 30 nM 5’-[32P]-labeled H-7 (Figure 10.2) and 150 nM human fPolλ was reacted with increasing concentration of an incoming dATP in buffer L (Materials and Methods). The single-turnover method was employed because DNA dissociates from fPolλ with a dissociation rate constant ($k_d$) which is only about 2- to 3-fold slower than the maximum nucleotide incorporation rate constant ($k_p$) (36), rendering the burst phase insignificant. Thus, the experiments were performed with fPolλ in molar excess over DNA to allow the direct observation of nucleotide incorporation in a single pass of the reactants through the enzymatic pathway without complications resulting from the steady-state formation of products (37). The DNA product 22-mer and remaining primer 21-mer were separated by gel electrophoresis and quantitated with a PhosphorImager. The product concentration was plotted against reaction time intervals. These data were subsequently fit to equation 1 (Materials and Methods), to yield an observed single-turnover rate constant ($k_{obs}$) at each concentration of dATP (Figure 10.7). The observed single-turnover rate constants were then plotted against dATP concentration (Figure 10.8). These data were subsequently fit to equation 2 (Materials and Methods), to yield a $k_p$ of $0.40 \pm 0.01$ s$^{-1}$ for the maximum dATP incorporation rate constant, and a $K_d$ of $0.78 \pm 0.08$ μM for dATP binding. The substrate specificity or polymerase efficiency ($k_p/K_d$) of dATP incorporation onto H-7 was calculated to be $0.485 \mu M^{-1}s^{-1}$ (Table 10.1).

The kinetic parameters (Table 10.1) for the incorporation of each of the remaining three correct nucleotides (dTTP onto H-1, dCTP onto H-6, and dGTP onto H-8) and for the incorporation of each of the 12 possible misincorporations onto H-1, H-6, H-7, and H-8
(Figure 10.2) were carried out in the same manner as described above. Notably, the ground-state binding affinity of all nucleotides to $\text{fPol}\lambda\cdot\text{H-DNA}$ (Table 10.1) were within 10-fold, but the $k_p$ values of correct nucleotides were three to five orders of magnitude higher than those of incorrect nucleotides. The differences in $k_p$ and $K_d$ led to significantly different nucleotide incorporation efficiencies ($k_p/K_d$) between correct and incorrect nucleotide incorporations and resulted in a fidelity of gap-filling synthesis in the range of $10^{-4}$ to $10^{-5}$ (Table 10.1).

**Kinetic effect of the 5'-phosphate of a downstream strand.** The kinetic parameters of nucleotide incorporation onto D-1OH (Figure 10.5) were measured with $\text{fPol}\lambda$ under single-turnover conditions described above and listed in Table 10.2. The apparent binding affinity of all four nucleotides was unaffected by the absence of the 5'-phosphate of the 19-mer in D-1OH, but the $k_p$ values were lowered by 10- to 100-fold (Table 10.2). Thus, the substrate specificity ($k_p/K_d$) of matched and mismatched nucleotides with D-1OH was 10-fold and 100-fold smaller, respectively, than those with D-DNA (D-1) (Table 10.2). Notably, dATP incorporation onto D-1OH was too slow to be measured. Slightly larger effect on incorrect incorporations compared to correct dTTP incorporation resulted in a slightly higher fidelity with D-1OH than with D-DNA (D-1).

**Kinetic effect of a downstream strand.** The significant kinetic effect of the 5’-phosphate moiety suggested the downstream strand itself could have dramatic influence on single-nucleotide gap-filling efficiency of $\text{fPol}\lambda$. To examine this hypothesis, we measured the pre-steady state kinetic parameters (Table 10.2) of single nucleotide incorporation onto D-1N which lacked a downstream strand (Figure 10.6) by employing single-turnover experiments as described above. The correct dTTP was incorporated at a rate constant of $0.025 \text{ s}^{-1}$ and a ground-state binding affinity of $2.6 \mu\text{M}$ (Table 10.2). In comparison to the
kinetic data with D-1OH (Table 10.2), the lack of a downstream strand further decreased the incorporation efficiency of correct dTTP and incorrect dCTP by approximately an additional 10-fold each. Incorporations of incorrect dGTP and dATP onto D-1N were too slow to be observed in several hours.

10.4 Discussion

The short-patch base excision repair pathway involves at least five enzymes which catalyze six reaction steps, including two catalyzed by Polβ (Scheme 10.1). Many lines of biochemical and in vitro evidence (22-27,38,39) indicate that the polymerase and dRPase activities of fPolλ are able to catalyze these two BER steps in vitro and may do so in vivo. Under the in vitro reconstituted BER reaction conditions, the velocity values for the uracil-DNA glycosylase, AP endonuclease, the DNA polymerase and dRPase activities of Polβ, and DNA ligase I have been measured to be 420, 100, 4.5, 0.75, 4.0 nM/s, respectively (21). The rate-limiting step is identified to be the dRP excision reaction which occurs at a similar velocity as the overall reconstituted BER system (0.6 nM/s) (21). In addition, the polymerase activity of Polβ that catalyzes the single-nucleotide gap-filling synthesis is 6-fold faster than its dRPase activity which cleaves the dRP moiety on the 5’-terminus of the downstream strand. This suggests that the natural single-nucleotide gapped DNA for the polymerase activity of Polβ should contain a dRP-moiety on its downstream strand. Thus, the best model DNA substrates are the dRP-DNA substrates (Figure 10.2), rather than the D-DNA substrates (Figure 10.4). However, because the 2-deoxyribose moiety in dRP-DNA is an equilibrium mixture of α– and β–hemiacetals (2-deoxy-D-erythro-pentofuranoses), an aldehyde, and a hydrated aldehyde in solution (35),
rigorous kinetic analysis becomes overly complicated, and as such D-DNA substrates
have been predominantly used as the gapped DNA substrates to examine the kinetics of
nucleotide incorporation catalyzed by both Polβ (30,31,40-44) and Polλ (34,36). To
overcome the instability of dRP-DNA in solution, we decided to use the H-DNA
substrates (Figure 10.3) in our pre-steady state kinetic analysis with human fPolλ.

Under single-turnover conditions, the kinetic parameters of all sixteen possible nucleotide
incorporations were measured using the H-DNA substrates (Figure 10.3), and the fidelity
of human fPolλ was determined to be in the range of $10^{-4}$ to $10^{-5}$ (Table 10.1). This range
was identical to the fidelity with D-DNA (34), suggesting the presence of the dRP-mimic
did not affect gap-filling error frequency of fPolλ. However, the substrate specificity
values (Table 10.1) are 1.6- to 6.7-fold lower than those corresponding values for
nucleotide incorporation onto D-DNA (34), resulting in an average efficiency ratio,
$\left(\frac{k_p}{K_d}\right)_{D-DNA}/\left(\frac{k_p}{K_d}\right)_{H-DNA}$, of 3.4 (Table 10.1). The difference in substrate specificity
$\left(\frac{k_p}{K_d}\right)$ can be mathematically attributed to either $k_p$, $K_d$, or both. However, the ground-
state binding affinities ($K_d$) of both correct and incorrect nucleotides with H-DNA (Table
10.1) are very similar to those corresponding $K_d$ values obtained with D-DNA (34),
indicating that the presence of the dRP-mimic on the 5’-terminus of the downstream
strand did not affect the binding affinity of an incoming nucleotide. In contrast, the
sixteen maximum incorporation rate constants ($k_p$) with H-DNA on average were 3.5-fold
lower than those corresponding $k_p$ values with D-DNA, suggesting the presence of the
dRP-mimic moderately affected catalysis during nucleotide incorporation. However,
what is the catalytic efficiency of fPolλ with the natural dRP-DNA substrates (Figure
10.2)? Due to close chemical similarity between dRP-DNA and H-DNA (Figure 10.2 and
10.3), it is reasonable to speculate that fPolλ will incorporate nucleotides onto dRP-DNA
with efficiencies closer to those observed with H-DNA as than with D-DNA. This
hypothesis goes against the qualitative experiments of Srivastava et al. (21) in which they have observed similar nucleotide incorporation efficiency with Polβ in the presence or absence of the dRP moiety. It is known that the dRP excision catalyzed by the dRPase domain occurs via a Schiff-base formation and β–elimination (45-47). If the Schiff-base formation between the C1 of the dRP group and Lys312 of fPolλ (29,48) is faster than nucleotide incorporation while the β–elimination limits both dRP excision and BER, it is possible that fPolλ will be more efficient with dRP-DNA than with H-DNA because the covalent anchoring of the downstream strand of dRP-DNA may facilitate gap-filling DNA synthesis catalyzed by fPolλ. More kinetic experiments are required to examine the aforementioned hypothesis. From a structural perspective, the effect of the dRP moiety or its mimic on DNA, dNTP binding, and the polymerase active site conformation remain unclear because all ternary crystal structures of both tPolλ (33,49) and Polβ (32,50,51) are solved in the presence of “D-DNA”, rather than “dRP-DNA” or “H-DNA”. However, these structures do reveal intimate contacts between the dRPase domains of both Polβ and tPolλ (Figure 10.1) and the downstream strand (29,32). The terminal 5’-phosphate of the downstream strand is buried in a positively charged pocket of the dRPase active site (29,32), e.g. Tyr267, Arg275, Tyr279, Lys307, and Arg308 in tPolλ (29,48). Thus, the dRP-mimic likely interacts with the dRPase domain which in turn may affect the active site conformation of fPolλ and thus the $k_p$ value of nucleotide incorporation.

Consistently, the intimate interactions between the dRPase domain, the downstream strand, and its 5’-phosphate moiety did impact catalysis significantly. Table 10.3 indicated that fPolλ incorporated a matched dTTP most efficiently with single-nucleotide gapped D-DNA (D-1) (Figure 10.4). The absence of the 5’-phosphate group in D-1OH (Figure 10.5) caused an 11-fold decrease in dTTP incorporation efficiency ($k_p/K_d$) while lack of the entire downstream strand in D-1N (Figure 10.6) led to an additional 15-fold
decrease. As a consequence, the catalytic efficiency decreased by 160-fold from D-DNA (D-1) to D-1N. Similarly, Polβ is found to have 6- to 40-fold higher nucleotide incorporation efficiency with single-nucleotide gapped DNA than with non-gapped DNA (52,53). The low substrate specificity of matched dTTP with D-1N further suggested that fPolλ is too inefficient to be a primer/template-dependant polymerase. In contrast, the dTTP incorporation efficiency of fPolλ (1.5 μM⁻¹ s⁻¹) is close to the range of Polβ (1.9-8.5 μM⁻¹ s⁻¹) in the presence of single-nucleotide gapped DNA (12,14-16). Moreover, our recent fidelity studies suggest that human fPolλ has similar single-nucleotide gap-filling fidelity (10⁻⁴-10⁻⁵) as Polβ (34). Taken together, these results strongly suggested that fPolλ like Polβ preferred short gapped DNA over non-gapped DNA and was likely a gap-filling polymerase involved in BER.

Interestingly, from D-1N to D-1OH, and to D-DNA (D-1), the change in dTTP incorporation efficiency was due to a 100-fold variation in $k_p$, while the binding of an incoming nucleotide ($K_d$) was similarly tight (Table 10.2). In contrast, the increase of catalytic efficiency from non-gapped to gapped DNA with Polβ was due to a considerable change in $K_d$, rather than $k_p$ (12,14-16). These suggested that fPolλ and Polβ achieve higher catalytic efficiency through different mechanisms, and these two enzymes have evolved divergently. Similar $K_d$ values of dTTP with three different DNA substrates (Table 10.3) indicated that the unprecedented high nucleotide binding affinity was due to the intimate interactions between an incoming nucleotide and the active site residues of fPolλ as revealed by the ternary crystal structures of tPolλ (33), rather than the presence of a downstream strand including its 5’-phosphate moiety. The significant decrease in $k_p$ from D-DNA (D-1) to D-1OH, and to D-1N (Table 10.3) suggested the intense interactions between the downstream strand including its 5’-phosphate and the dRPase domain, as demonstrated by the binary and ternary crystal structures of tPolλ (29,33), directed and
anchored the productive binding of DNA and dNTP at the active site of fPolλ. Consequently, the absence of these interactions will either improperly align the 3’-OH of the upstream primer 21-mer and the α-phosphate of the dNTP for in-line attack, or affect the local protein conformational change, leading to slower catalysis. In the meantime, the fidelity of fPolλ (Table 10.2) was slightly increased, however. This trend is consistent with what we have observed with different truncated fragments of Polλ (34), but disagrees with the general trend summarized from a survey of the A-, B-, X-, and Y-families by Beard et al. that a more catalytically efficient DNA polymerase has a higher polymerization fidelity (54).

In conclusion, our pre-steady state kinetic data have demonstrated that the downstream strand and its 5’-phosphate moiety are critical to the polymerase efficiency of fPolλ. For the first time, we have quantitatively evaluated the kinetic effect of a dRP-mimic on the 5’-terminus of a downstream strand. Because this dRP-mimic only moderately affected the incorporation efficiency of both correct and incorrect nucleotides while having insignificant effect on the fidelity of human DNA polymerase λ, gapped substrates like D-DNA in Figure 10.4, which have been predominantly used in the literature, are reasonable model substrates.
10.5 Figures, Tables, and Schemes

Figure 10.1. Schematic representation of the domain organization of human DNA polymerases $\beta$ and $\lambda$ (full-length and C-terminal fragment). Each domain, with amino acid residue numbers are indicated above, is shown as a rectangle.
Figure 10.2. dRP-DNA substrate. The 5'-terminus of the downstream strand 19-mer contains a dRP moiety. The upstream primer 21-mer is 5'-32P-labeled.
Figure 10.3. H-DNA substrate. The 5'-terminus of the downstream strand 19-mer contains a 1,2-dideoxyribose-5-phosphate moiety. X represents one of the four natural bases (adenine, cytosine, guanine, thymine). The upstream primer 21-mer is 5'-32P-labeled.
Figure 10.4. **D-DNA substrate.** The 5’-terminus of the downstream strand 19-mer contains a phosphate moiety. X represents one of the four natural bases (adenine, cytosine, guanine, thymine). The upstream primer 21-mer is 5’-\(^{32}\)P-labeled.
Figure 10.5. **D1-OH DNA substrate.** The 5'-terminus of the downstream strand 19-mer contains a hydroxyl moiety. The upstream primer 21-mer is 5'-\(^{32}\text{P}\)-labeled.

![DNA sequence](image.png)

Figure 10.6. **D-1N DNA substrate.** The upstream primer 21-mer is 5'-\(^{32}\text{P}\)-labeled.
Figure 10.7. Concentration dependence of a correct nucleotide incorporation on the rate of incorporation. A preincubated solution of fPolλ (150 nM) and 5'-32P-labeled D7 (30 nM) was rapidly mixed with increasing concentrations of Mg²⁺-dATP (0.5μM, ▲; 1 μM, □; 2 μM, X; 4 μM, ◆; 8 μM, +; 12 μM, △; 25 μM, ●; 48 μM, ■) for various time intervals. The solid lines are the best fits to equation 1.
Figure 10.8. Concentration dependence of a correct nucleotide incorporation on the rate of incorporation. The observed rate constants obtained from the above data fitting (Figure 10.7) were plotted as a function of dATP concentration. The data (●) were then fit to equation 2, yielding a $k_p$ of $0.40 \pm 0.01$ s$^{-1}$ and a $K_d$ of $0.78 \pm 0.08$ µM.
Table 10.1. Kinetic parameters of nucleotide incorporation onto single-nucleotide gapped H-DNA containing a dRP-mimic catalyzed by human fPol at 37 ºC.

<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$(k_p/K_d)_{H-DNA}$ (μM$^{-1}$s$^{-1}$)</th>
<th>Fidelity$^a$</th>
<th>$(k_p/K_d)_{D-DNA}$ (μM$^{-1}$s$^{-1}$)$^b$</th>
<th>Efficiency Ratio$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Template A (H-1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>1.7 ± 0.3</td>
<td>0.5 ± 0.3</td>
<td>0.30</td>
<td>1</td>
<td>1.5</td>
<td>5.1</td>
</tr>
<tr>
<td>dATP</td>
<td>4 ± 1</td>
<td>1.97 x 10$^{-5}$ ± 1 x 10$^{-6}$</td>
<td>4.9 x 10$^{-6}$</td>
<td>1.7 x 10$^{-5}$</td>
<td>1.8 x 10$^{-5}$</td>
<td>3.7</td>
</tr>
<tr>
<td>dCTP</td>
<td>10 ± 5</td>
<td>2.7 x 10$^{-3}$ ± 5 x 10$^{-4}$</td>
<td>2.6 x 10$^{-4}$</td>
<td>8.7 x 10$^{-4}$</td>
<td>5.0 x 10$^{-4}$</td>
<td>2.0</td>
</tr>
<tr>
<td>dGTP</td>
<td>6.4 ± 0.6</td>
<td>2.5 x 10$^{-4}$ ± 1 x 10$^{-5}$</td>
<td>3.9 x 10$^{-5}$</td>
<td>1.3 x 10$^{-4}$</td>
<td>1.3 x 10$^{-4}$</td>
<td>3.4</td>
</tr>
<tr>
<td><strong>Template G (H-6)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dCTP</td>
<td>1.6 ± 0.3</td>
<td>1.10 ± 0.04</td>
<td>0.69</td>
<td>1</td>
<td>1.8</td>
<td>2.6</td>
</tr>
<tr>
<td>dATP</td>
<td>1.6 ± 0.6</td>
<td>1.04 x 10$^{-5}$ ± 9 x 10$^{-7}$</td>
<td>6.6 x 10$^{-6}$</td>
<td>9.5 x 10$^{-6}$</td>
<td>3.3 x 10$^{-5}$</td>
<td>5.0</td>
</tr>
<tr>
<td>dTTP</td>
<td>10 ± 3</td>
<td>1.4 x 10$^{-3}$ ± 2 x 10$^{-4}$</td>
<td>1.3 x 10$^{-4}$</td>
<td>1.9 x 10$^{-4}$</td>
<td>2.8 x 10$^{-4}$</td>
<td>2.1</td>
</tr>
<tr>
<td>dGTP</td>
<td>4.2 ± 0.7</td>
<td>4.7 x 10$^{-5}$ ± 3 x 10$^{-6}$</td>
<td>1.1 x 10$^{-5}$</td>
<td>1.6 x 10$^{-5}$</td>
<td>5.0 x 10$^{-5}$</td>
<td>4.5</td>
</tr>
<tr>
<td><strong>Template T (H-7)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>0.78 ± 0.08</td>
<td>0.40 ± 0.01</td>
<td>0.49</td>
<td>1</td>
<td>1.6</td>
<td>3.3</td>
</tr>
<tr>
<td>dTTP</td>
<td>11 ± 2</td>
<td>7.1 x 10$^{-3}$ ± 6 x 10$^{-6}$</td>
<td>6.4 x 10$^{-6}$</td>
<td>1.3 x 10$^{-5}$</td>
<td>2.4 x 10$^{-5}$</td>
<td>3.7</td>
</tr>
<tr>
<td>dCTP</td>
<td>11 ± 2</td>
<td>4.8 x 10$^{-3}$ ± 4 x 10$^{-4}$</td>
<td>4.5 x 10$^{-4}$</td>
<td>9.3 x 10$^{-4}$</td>
<td>7.0 x 10$^{-4}$</td>
<td>1.6</td>
</tr>
<tr>
<td>dGTP</td>
<td>8 ± 5</td>
<td>2.2 x 10$^{-3}$ ± 5 x 10$^{-4}$</td>
<td>2.7 x 10$^{-4}$</td>
<td>5.6 x 10$^{-4}$</td>
<td>1.4 x 10$^{-3}$</td>
<td>5.1</td>
</tr>
<tr>
<td><strong>Template C (H-8)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dGTP</td>
<td>1.5 ± 0.3</td>
<td>0.76 ± 0.02</td>
<td>0.51</td>
<td>1</td>
<td>1.2</td>
<td>2.3</td>
</tr>
<tr>
<td>dATP</td>
<td>1.5 ± 0.3</td>
<td>3.8 x 10$^{-3}$ ± 2 x 10$^{-6}$</td>
<td>2.5 x 10$^{-5}$</td>
<td>4.9 x 10$^{-5}$</td>
<td>1.7 x 10$^{-4}$</td>
<td>6.7</td>
</tr>
<tr>
<td>dCTP</td>
<td>3 ± 1</td>
<td>6.2 x 10$^{-4}$ ± 6 x 10$^{-5}$</td>
<td>1.9 x 10$^{-4}$</td>
<td>3.7 x 10$^{-4}$</td>
<td>3.8 x 10$^{-4}$</td>
<td>2.0</td>
</tr>
<tr>
<td>dTTP</td>
<td>9 ± 5</td>
<td>6 x 10$^{-4}$ ± 2 x 10$^{-4}$</td>
<td>7.3 x 10$^{-5}$</td>
<td>1.4 x 10$^{-4}$</td>
<td>1.4 x 10$^{-4}$</td>
<td>1.9</td>
</tr>
</tbody>
</table>

$^a$Calculated as $(k_p/K_d)_{incorrect}/[(k_p/K_d)_{correct} + (k_p/K_d)_{incorrect}]$

$^b$The $(k_p/K_d)_{D-DNA}$ values are from Table 1 of reference (34).

$^c$Calculated as $(k_p/K_d)_{D-DNA}/(k_p/K_d)_{H-DNA}$. 

Table 10.1. Kinetic parameters of nucleotide incorporation onto single-nucleotide gapped H-DNA containing a dRP-mimic catalyzed by human fPol at 37 ºC.
<table>
<thead>
<tr>
<th>dNTP</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (μM$^{-1}$s$^{-1}$)</th>
<th>Fidelity$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-DNA (D-1) substrate$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>2.6 ± 0.4</td>
<td>3.9 ± 0.2</td>
<td>1.5</td>
<td>1</td>
</tr>
<tr>
<td>dCTP</td>
<td>14.8 ± 5.0</td>
<td>0.022 ± 0.003</td>
<td>1.5 x 10$^{-3}$</td>
<td>9.8 x 10$^{-4}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>8.2 ± 0.6</td>
<td>0.006 ± 0.003</td>
<td>7.6 x 10$^{-4}$</td>
<td>5.8 x 10$^{-4}$</td>
</tr>
<tr>
<td>dATP</td>
<td>4.1 ± 0.5</td>
<td>0.0029 ± 0.0001</td>
<td>7.1 x 10$^{-4}$</td>
<td>5.4 x 10$^{-4}$</td>
</tr>
<tr>
<td>D-1OH substrate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>4.1 ± 0.8</td>
<td>0.56 ± 0.03</td>
<td>0.14</td>
<td>1</td>
</tr>
<tr>
<td>dCTP</td>
<td>5.5 ± 1.1</td>
<td>(1.6 ± 0.1) x 10$^{-4}$</td>
<td>2.9 x 10$^{-5}$</td>
<td>2.1 x 10$^{-4}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>8.5 ± 2.7</td>
<td>(1.3 ± 0.1) x 10$^{-5}$</td>
<td>1.5 x 10$^{-6}$</td>
<td>1.1 x 10$^{-5}$</td>
</tr>
<tr>
<td>dATP</td>
<td>No observed incorporation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-1N substrate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dTTP</td>
<td>2.6 ± 0.7</td>
<td>0.025 ± 0.002</td>
<td>9.4 x 10$^{-3}$</td>
<td>1</td>
</tr>
<tr>
<td>dCTP</td>
<td>6.6 ± 2.0</td>
<td>(8.3 ± 0.6) x 10$^{-6}$</td>
<td>1.3 x 10$^{-6}$</td>
<td>1.4 x 10$^{-4}$</td>
</tr>
<tr>
<td>dGTP</td>
<td>No observed incorporation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dATP</td>
<td>No observed incorporation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Calculated as $(k_p/K_d)_{\text{incorrect}}/[(k_p/K_d)_{\text{correct}} + (k_p/K_d)_{\text{incorrect}}]$

$^b$Data are from Table 1 of reference (34).

Table 10.2. Pre-steady state kinetic parameters with D-DNA (D-1), D-1OH, and D-1N.
### Table 10.3. Pre-steady state kinetic parameters of matched dTTP incorporation.

<table>
<thead>
<tr>
<th>DNA Substrate</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (μM$^{-1}$s$^{-1}$)</th>
<th>Efficiency ratio$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-1N</td>
<td>2.6 ± 0.7</td>
<td>0.025 ± 0.002</td>
<td>9.4 x 10$^{-3}$</td>
<td>1</td>
</tr>
<tr>
<td>D-1OH</td>
<td>4.1 ± 0.8</td>
<td>0.56 ± 0.03</td>
<td>0.14</td>
<td>15</td>
</tr>
<tr>
<td>D-DNA (D-1)</td>
<td>2.6 ± 0.4</td>
<td>3.9 ± 0.2</td>
<td>1.5</td>
<td>160</td>
</tr>
</tbody>
</table>

$^a$Calculated as $(k_p/K_d)_{DNA}/(k_p/K_d)_{D-1N}$
Scheme 10.1. Reaction sequence of uracil-initiated short-patch BER.
Scheme 10.2. Minimal kinetic mechanism for dNTP incorporation onto D-DNA.
10.6 References


374
CHAPTER 11

ADDITIONAL RESULTS, CONCLUSIONS, AND FUTURE DIRECTIONS

11.1 Introduction

DNA polymerases are fascinating biological machines that possess a wide specificity for incorporating a variety of both canonical and noncanonical nucleotide substrates into DNA, yet are also capable of incorporating the “correct” nucleotide substrate with a selection preference approaching seven orders of magnitude, by changing their incorporation specificity based on the identity of the nucleotide base of the parental DNA sequence in each enzyme turnover (1). In the rare event of an incorrect nucleotide insertion, these high fidelity replicative DNA polymerases sense the mismatch and correct the mistake, 90 to 99.9% of the time, to achieve fidelities on the order of $10^{-8}$ to $10^{-9}$. Moreover, some replicative DNA polymerases can perform this selection process at rates exceeding 300 s$^{-1}$ (2). However, since all organisms are often subjected to stressful environmental conditions which can produce a myriad of potentially lethal chemical modifications to the structure of DNA, organisms have evolved additional pathways and biological machineries to balance the need for both the maintenance of faithful replication and the need for chance mutations that are responsible for sustaining life on this planet. At the core of this machinery are specialized DNA polymerases that not only function in faithful processes like DNA recombination and DNA repair, but also
participate in deliberately unfaithful processes including DNA lesion bypass and the
generation of mutations essential for the vertebrate adaptive immunity (3). The evolution
of such machinery is no more evident than the situation in humans who possess at least
sixteen distinct DNA polymerases indicating the complexity of DNA transactions in such
a highly specialized organism.

11.2 Additional Results and Future Directions for the Y-family DNA Polymerases

Lesion bypass catalyzed by the human Y-family DNA polymerases. As mentioned in the
introduction of this Chapter, sixteen DNA polymerases have been identified in humans.
Of these sixteen, four DNA polymerases are members of the Y-family. While Dpo4 has
served as an excellent model for characterizing the Y-family as a whole, it has become
apparent here and in the literature that DNA polymerases even within the same family
posses structural and functional attributes that render them distinct from one another. As
such, it has become a goal of mine to begin to elucidate the abilities of the human Y-
family polymerases in the bypass of various DNA lesions in an effort to identify the Y-
family member most capable, and therefore most likely, to bypass a particular lesion. We
have purified three of the four DNA polymerases, human DNA polymerase η, human
DNA polymerase ι, and human and mouse DNA polymerase κ. For these studies we have
chosen not to purify and assay the fourth human Y-family DNA polymerase REV1, since
it has been shown to be a dCMP transferase that indiscriminately inserts dCMP into DNA
regardless of the template sequence (4). In addition, due to the ability of REV1 to
specifically interact with DNA polymerases η, κ, and ι yet not with other DNA polymerases such as β or µ (5-7), coupled with the observation that REV1 will switch its interaction partner based on relative concentrations of the competing Y-family DNA polymerase (5), it has recently been hypothesized that REV1 may instead serve a more important role in mediating the process of polymerase switching (discussed below).

Due to the fact that the most common DNA lesion encountered in cells is an AP site (see Chapter 4), this non-coding lesion was studied first with respect to the bypass abilities of these human Y-family DNA polymerases. Using the same running start analyses that were performed for Dpo4 and SsoPolB1 in Chapter 4, we demonstrated that human DNA polymerases η, human DNA polymerase ι, and both human and mouse DNA polymerase κ were able to bypass the AP site and continue extension until the full length bypass product was generated (data not shown). We then purified and sequenced the full length bypass products synthesized by each individual polymerase using the SOSA as described in Chapter 5. Our results for human DNA polymerase η (Figure 11.1) reveal that 85% of the time polymerase η incorporates either a dAMP (19/39 colonies, 49%) or a dGMP (15/39 colonies, 38%) opposite the lesion. The remaining incorporation events opposite the AP site include the incorporation of dCMP (1/39 colonies, 3%), dTMP (2/39 colonies, 5%), and the generation of a -2 deletion (2/39 colonies, 5%). Our results from the SOSA with mouse DNA polymerase κ (Figure 11.2) reveal predominant incorporation of dAMP opposite the lesion (33/44 colonies, 75%) followed by generation of a -2 deletion (6/44 colonies, 14%), the incorporation of dTMP (3/44 colonies, 6%), and the generation of a -1 deletion (2/44 colonies, 4%). The results for human DNA polymerase κ (data not shown) were similar to the mouse protein in that the predominant incorporation was dAMP (38/50 colonies, 76%), while also preferring to incorporate dGMP opposite the lesion (6/50 colonies, 12%) followed by the rare generation of a -2 deletion (1/50
colonies, 2%). We noted various mutations that mostly occurred outside the immediate vicinity (within four to five nucleotides) of the AP lesion for both DNA polymerase η and κ (Figures 11.1 and 11.2), suggesting that they may not be caused by the lesion itself but instead arise as a consequence of the extremely low fidelity that characterize both DNA polymerases (8,9). Amazingly, SOSA analysis performed for AP bypass catalyzed by human DNA polymerase ι reveal that it generates a plethora of mutations at each position of the AP bypass product. This high error rate (synthesizing an error at each position of the sequencing window at a probability equivalent to 10% to 80% of the flux through that position) suggests that DNA polymerase ι may play an important role in somatic hypermutation due to the fact that this enzyme is equally mutagenic upstream the AP site where it is not expected to be influenced by the lesion, as it is downstream from this lesion (data not shown). Our results here indicate that human DNA polymerase η and κ generate few mutations during the bypass of the lesion, especially when compared to DNA polymerase ι, and may in fact be the polymerases to bypass this lesion in vivo. The relevant kinetic studies to determine the rate constants of AP site bypass for all of these enzymes are forthcoming and will augment of understanding of these bypass events.

Our studies reported here provide a wealth of knowledge from an in vitro perspective regarding lesion bypass catalyzed by these specialized enzymes and will help geneticists and molecular biologists design and interpret their in vivo results. However, there are still several unanswered questions that remain. In particular, what are the mechanisms that bring about the requisite polymerase switching event within the replication machinery so that the stalled replicative DNA polymerase can be replaced with a lesion bypass polymerase? After the lesion is bypassed, how does the replication machinery reinitiate DNA synthesis and how long does this process take? In addition, how does the replication machinery negotiate between the available lesion bypass polymerases upon
encountering a lesion? Together, answers to these questions will significantly advance our understanding of the biological processes of DNA replication and mutagenesis.

*Structure-function relationships in Dpo4.* The crystal structure of Dpo4 in a ternary complex with undamaged DNA and an incoming nucleotide has shown that these Y-family DNA polymerases possess the three conserved structural domains (finger, thumb, and palm) found in all structure known DNA polymerases (10-15). However, a fourth structural domain, absent in other DNA polymerase families, is present in these Y-family polymerases and is designated “little finger” in Dpo4 (16). The little finger domain has been shown to be located in the DNA major groove upstream of the active site and interacts mainly with the backbone of both strands to increase the binding affinity of Dpo4 and DNA. A recent study involving the exchange of the little finger domains of Dpo4 and another Y-family member Sulfolobus acidocaldarius Dbh, suggests the little finger domain may play a major role in the fidelity, processivity, and lesion-bypass specificities of the Y-family DNA polymerases (17). This little finger domain is tethered to the thumb domain via a fourteen amino acid residue “linker” that has numerous basic amino acid residues in Dpo4 and wraps around the DNA interacting with the negatively charged phosphosugar backbone (16). Interestingly, Dbh has been shown to incorporate fewer nucleotides per binding event compared to Dpo4 (18,19) and binds 6-fold weaker than Dpo4 to an undamaged DNA substrate (20). Sequence alignment shows that Dpo4 and Dbh have 54% overall sequence identity, but only 41% identity in the LF domains. Although their linkers are both fourteen amino acid residues in length (Figure 11.3), the linker of Dpo4 (pI = 10) consists of more basic amino acid residues than the linker of Dbh (pI = 8.6) and may thereby interact more tightly with the negatively charged DNA at physiological pH. These results suggested to us that the linker plays an important role in the overall binding of these respective enzymes to DNA and furthermore, the size and
amino acid sequence of the linker may affect DNA binding significantly. In order to test these hypotheses, we first overexpressed and purified the Dpo4 little finger domain (231-352 residues) and the catalytic “core” fragment (1-240 residues) of Dpo4 separately (data not shown). We then investigated the binding affinity of these individual domains using the gel-mobility shift assay, which was used in Chapter 4 to show that Dpo4 binds to undamaged DNA with an equilibrium dissociation constant ($K_d$) of 10 nM. The little finger domain alone was determined to bind to DNA with a $K_d$ ($K_d \approx 1/\text{affinity}$) of 3.4 μM which is over 300-fold weaker than the affinity of wild type Dpo4 for DNA (Chapters 3 and 4). To our surprise, we did not detect the core•DNA complex formation even when up to 32 μM of the core fragment was incubated with 0.1 μM DNA (data not shown). This suggests that the core fragment does not bind to DNA or at least binds to DNA with an extremely low affinity. To examine if the addition of the little finger domain could facilitate the core fragment binding to DNA, we analyzed an incubated solution of 1.5 μM little finger domain, 0.5 μM core fragment and 0.1 μM DNA yet did not observe any ternary complex formation. This suggests there was no cooperative binding between the little finger domain and the catalytic core fragment to DNA in the absence of the covalent linkage between them. It is worth noting that this failure could be due to the fact that the purified core fragment lacked full activity since it incorporated single correct dTTP into D-1 with a $10^4$ fold slower rate constant than Dpo4 (data not shown). This low activity could be due to an unfolded core fragment in the absence of the little finger domain, although alternatively, this low activity could be due to just extremely weak binding between a competently folded core fragment and DNA. We plan to perform circular dichroism studies to determine whether or not the core fragment is folded.
To examine if the linker size is critical for the DNA-binding ability of Dpo4, we engineered and purified several Dpo4 mutants containing a variety of amino acid deletions and additions. Three addition-mutants were designed to have either one, four, or six glycine residues inserted between residues I237 and R238 of the linker while three deletion-mutants contained either a one (I237), four (N234 through I237), or six (E232 through I237) amino acid deletion (Figure 11.4). Once these mutants were overexpressed and purified, gel-mobility shift assays were used to determine the equilibrium dissociation constant \( (K_d) \) for each E•DNA complex. While the wild type Dpo4 was shown to bind DNA with a \( K_d \) of 10 nM, we observed a systematic increase in the \( K_d \) (corresponding to a decrease in binding affinity) for the 1-glycine addition mutant \( (K_d = 118 \text{ nM}) \), 4-glycine mutant \( (K_d = 279 \text{ nM}) \), and 6-glycine mutant \( (K_d = 674 \text{ nM}) \). Notably we observe a similar trend for the 1-deletion mutant \( (K_d = 159 \text{ nM}) \), 4-deletion mutant \( (K_d = 304 \text{ nM}) \), and 6-deletion mutant \( (K_d = 1688 \text{ nM}) \). These results (Figure 11.4) suggest that the separation between the thumb and little finger domain as mediated by the size of the Dpo4 linker has been optimized for binding to DNA. Unfortunately, we have not completed the mutagenesis of the basic amino acid residues of the Dpo4 linker to neutral residues in order to elucidate whether or not the sequence identity of the linker region affects DNA binding. These experiments should be completed soon. It is also important to note that the incorporation efficiencies of the linker mutants assayed thus far are within 2-fold of that observed for the wild type Dpo4 (data not shown).

The sugar selectivity of Dpo4. Another issue that we wanted to address was the sugar selectivity of Dpo4. The crystal structure of the ternary complex of Dpo4 with undamaged DNA and a correct incoming dNTP, has revealed that the sugar ring of the incoming dNTP contacts the benzene ring of Y12 specifically at the position where the 2'-hydroxyl would be found for a ribonucleotide (rNTP). Therefore these authors
concluded that Y12 functions as a steric gate against the incorporation of rNTPs (16). However, to our knowledge, no additional evidence has been provided to verify this conclusion. We thus performed site-directed mutagenesis to generate the Dpo4 Y12A construct which was then purified using the protocol described in Chapter 2. We wanted to determine the quantitative difference between the incorporation efficiency of rNTPs and dNTPs into DNA substrates after replacing the large benzene group of Y12 with the significantly smaller methyl group of alanine. We first performed the single nucleotide incorporation assays as described in Chapter 2 for the incorporation of a “matched” UTP into the D-1 substrate catalyzed by wild type Dpo4 and found that the $k_p$, $K_d$, and substrate specificity ($k_p/K_d$) were $0.0023 \pm 0.0001 \text{ s}^{-1}$, $1141 \pm 235 \text{ µM}$, and $2.1 \times 10^{-6} \text{ s}^{-1}\text{µM}^{-1}$ respectively. Thus we observed a sugar selectivity ($(k_p/K_d)_{\text{correct dNTP}}/(k_p/K_d)_{\text{matched rNTP}}$) for wild type Dpo4 of roughly 19,700 (Table 11.1). We subsequently performed an analogous assay for the incorporation of dTTP into D-1 using the purified Dpo4 mutant Y12A and obtained a substrate specificity of $0.015 \text{ s}^{-1}\text{µM}^{-1}$, a value just 2.7-fold lower than that observed for the wild type enzyme, indicating that the mutation did not adversely affect the ability of the Y12A mutant to incorporate dNTPs (Table 11.1).

Subsequently, we performed three assays to determine the substrate specificities for the incorporations of three “matched” ribonucleotides UTP, CTP, and GTP into their corresponding DNA substrates D-1, D-6, and D-8 (Table 11.2). If we assume the substrate specificities for all correct dNTP incorporations catalyzed by the Dpo4 Y12A mutant have a similar substrate specificity as dTTP into D-1, we can calculate the sugar selectivity for UTP, CTP, and GTP incorporation to be 29, 24, and 17 respectively. So this single mutation lowers the sugar selectivity of the wild type Dpo4 by roughly 860-fold (Table 11.1). While we do not know what structural features account for the remaining ~20-fold difference in sugar selectivity of the Y12A mutant, our collaborator
Dr. Hong Ling (University of Western Ontario) has recently crystallized this Y12A mutant, and we hope that the subsequent structure analysis will explain this difference.

11.3 Additional Results and Future directions for the X-family DNA Polymerases

The X-family DNA polymerases consists of DNA polymerase λ, DNA polymerase β, DNA polymerase µ, terminal deoxynucleotidyl transferase (TdT), and yeast DNA polymerase IV. My work has focused almost exclusively on the kinetic mechanism of human DNA polymerase λ, but I have also performed experiments with human DNA polymerase β, DNA polymerase µ, and TdT as discussed below.

**Detailed mechanistic studies.** In an attempt to expand the minimal kinetic mechanism reported in Chapter 8, we performed additional studies to explore the nature of nucleotide incorporation catalyzed by DNA polymerase λ. A burst experiment was performed by mixing a preincubated solution of fPolλ (50 nM) and 5′-[32P]-D-1 (150 nM) with dTTP (100 µM) at 37 °C followed by quenching the reactions at times ranging from 10 ms to 1 s. While we failed to observe an obvious burst phase, we were able to fit the data to the following equation, 

\[
[\text{Product}] = [\text{fPolλ•D-1}]\left[\frac{k}{(k+k_{-1})}\right]^2(1-\exp^{-kt}) + v_{ss}t,
\]

where \(k\) is the observed first turnover, \(k_{-1}\) is the DNA dissociation rate constant (Scheme 11.1), and \(v_{ss}\) is the steady-state rate constant. After fitting the data we determined the \(k\), the \(k_{-1}\) and \(v_{ss}\) to be 3.0 s\(^{-1}\), 0.81 s\(^{-1}\), and 9.6 nM s\(^{-1}\) respectively (Figure 11.5). These results suggested that the lack of an obvious burst phase was due to the fairly insignificant three-fold difference between the rate constant for nucleotide incorporation (3.0 s\(^{-1}\)) and the DNA dissociation rate constant (0.81 s\(^{-1}\)).
In order to determine the rate-limiting step for nucleotide incorporation catalyzed by human DNA polymerase λ in the first turnover, we performed the same pulse-chase/pulse-quench experiments as described in Chapters 3 and 6. Two experiments were performed by mixing a preincubated solution of fPolλ (1 µM) and unlabeled D-7 (100 nM, Figure 11.6) with [α-32P]-dATP (10 µM) for various time intervals. In the first experiment, reactions were quenched by the addition of 1 M HCl (pulse-quench), while in the second experiment, each reaction was chased by a large excess of unlabeled dATP (1.5 mM) for an additional 30 s, followed by 1M HCl quench (pulse-chase). After phenol/chloroform extraction of fPolλ, the remaining aqueous phase from each reaction was neutralized with 1 M NaOH/0.1 M Tris (pH 7.5) and filtered through a Biospin 6 column to remove unreacted [α-32P]-dATP. These data were subsequently quantitated and fit via a single-exponential equation to yield a pulse-quench amplitude of 11.1 ± 1.2 nM and a pulse-chase amplitude of 36.4 ± 1.2 nM (Figure 11.7). The difference in reaction amplitudes (25.3 nM) was reproducible and demonstrated the slow formation of an intermediate complex E’•DNA•dNTP. For reasons explained in Chapters 3 and 6, we conclude that this accumulation of intermediate product not observed under pulse-quench conditions, is compelling evidence for the existence of a rate-limiting protein conformational change for fPolλ.

In addition, elemental effect studies were performed to further characterize the nature of incorporation catalyzed by DNA polymerase λ. As described in Chapters 3 and 6 and elemental effect of 4- to 11-fold has been taken as evidence for a rate-limiting chemistry step for nucleotide incorporation based on the cleavage of phosphate diesters as reported by Herschlag (21). Here, a preincubated solution of fPolλ (120 nM) and 5’-[32P]-D-1 (30 nM) was rapidly mixed with either 100 µM dTTP or 100 µM S_p-dTTPαS. The single-
turnover rate constants were determined to be $3.2 \pm 0.7 \text{ s}^{-1}$ and $0.63 \pm 0.06 \text{ s}^{-1}$ respectively, resulting in an elemental effect, $k_p \frac{d\text{dTP}}{d\text{Sp-dTP}}$, of 5.1 (data not shown). This elemental effect value for correct nucleotide incorporation was very similar to the observed elemental effect for correct nucleotide incorporation into single-nucleotide gapped DNA catalyzed by DNA polymerase $\beta$ (22). Likewise, we determined the elemental effect for the incorporation of an incorrect nucleotide ($d\text{GTP}$, $5.1 \times 10^{-4} \text{ s}^{-1}$) versus its analog ($\text{Sp-dGTP}$, $3.3 \times 10^{-5} \text{ s}^{-1}$) catalyzed by DNA polymerase $\lambda$ to be equivalent to 15.6 (data not shown). This value was also similar to the value of the elemental effect derived for the incorrect incorporation of $d\text{GTP}$ by DNA polymerase $\beta$ (9.0) (23). Although the elemental effect as described in Chapters 3 and 6 is not a reliable diagnostic for determining the identity of the rate-limiting step, the fact that the elemental effects for both correct and incorrect nucleotide incorporation catalyzed by DNA polymerase $\lambda$ and DNA polymerase $\beta$ are similar, suggest that they share the same kinetic mechanisms involving their rate-limiting steps.

In the DNA polymerase field, a long-standing controversy that has been debated for decades is the identification of the rate-limiting step in the first turnover of the enzyme. Larger product yields observed from the pulse-chase rather than from the pulse-quench time course coupled with a small thio elemental effect (1- to 4-fold) have been presented as evidence in support of a rate-limiting protein conformational change step during correct nucleotide incorporation catalyzed by the Klenow fragment of *E. coli* DNA polymerase I (24), T7 DNA polymerase (2), human DNA polymerase $\gamma$ (25), HIV-1 reverse transcriptase (26), DNA polymerase $\eta$ (27), and Dpo4 as presented in Chapters 3 and 6. However, the identification of this step in the mechanism of DNA polymerase $\beta$ has remained unresolved based on conflicting reports (23,28-30) and has not been published for DNA polymerase $\lambda$. Surprisingly while thio elemental effect studies have
been performed for DNA polymerase β (see above), the pulse-chase/pulse-quench experiments have yet to be reported for this enzyme. It would be worthwhile to perform this set of experiments using DNA polymerase β and compare the results to those of DNA polymerase λ to elucidate the nature of the rate-limiting step in these X-family DNA polymerases.

*The sugar selectivity of DNA polymerase λ.* In addition to exploring the kinetic mechanism catalyzed by DNA polymerase λ, we were intrigued by the high binding affinity of the fPolλ-DNA binary complex for both correct and incorrect dNTPs. Our laboratory had already determined that ddCTP, a dCTP lacking a hydroxyl group at the 3’ position, has a 4-fold tighter binding affinity to the fPolλ-DNA binary complex than dCTP, suggesting that the 3’ hydroxyl decreases nucleotide binding in the ground-state (J. Fowler and Z. Suo unpublished results). We were curious to determine if the presence of a hydroxyl group at the 2’ position would affect nucleotide binding. We thus determined the $k_p$ and $K_d$ for incorporation of the four “matched” ribonucleotides (rNTPs), UTP, rCTP, rATP, and rGTP into their respective single-nucleotide gapped DNA substrates D-1, D-6, D-7, and D-8 (Figure 11.6) using the same single turnover conditions described in Chapter 8. Interestingly, “matched” rNTPs bind with 18- to 140-fold lower affinity and are incorporated with 100- to 800-fold slower rate constants than correct dNTPs (compare Table 11.3 to Table 1 in (31)). In addition, the fidelity and sugar selectivity, defined as $[(k_p/K_d)_{correct\ dNTP}/(k_p/K_d)_{matched\ rNTP}]$ were calculated and demonstrate that fPolλ incorporated rNTPs with a fidelity matching that of incorrect nucleotides ($\sim 10^{-4}$ to $10^{-5}$) while the sugar selectivity was in the range of 4,615 to 50,000. This discrimination against the incorporation of rNTPs provides a strong indication that steric interactions at the 2’ position of the sugar have a considerable impact on nucleotide incorporation by fPolλ. In addition, since cellular rNTPs have 10- to 100-fold higher
concentrations when compared to dNTPs (32), a matched rNTP will be incorporated at least as frequently as an incorrect dNTP based on our kinetic analysis. In our single-nucleotide gapped substrates, the incorporation of an rNTP will form a nicked DNA substrate (Figure 11.6). Interestingly, studies from the laboratories of Tom Ellenberger and Dale Ramsden have shown that an rNTP incorporated into a single-nucleotide gapped DNA substrate can be ligated to the downstream strand via human DNA ligase I (33) and human XRCC4-ligase IV (34). Thus misincorporations of “matched” rNTPs should be considered when assessing the overall fidelity of fPolλ, which we determined to be on the order of $10^{-4}$ to $10^{-5}$ (Table 11.3), similar to that observed for the incorporation of incorrect dNTPs.

We also determined the incorporation efficiency of “mismatched” rNTPs into the single-nucleotide gapped D-1 substrate (Figure 11.6) with the kinetic parameters listed in Table 11.4. Based on the calculated substrate specificities ($k_p/K_d$), the incorporation of “mismatched” rNTPs is at least 10-fold less efficient than the incorporation of a “matched” rNTP (Table 11.3), and has at least a 10,000-fold lower substrate specificity than the incorporation of a correct dNTP, making incorporations of “mismatched” rNTPs an extremely rare event.

It has since become an aim of our laboratory to examine the active site of fPolλ to determine which residue(s) is (are) involved in the ability of fPolλ to discriminate so efficiently against the ribose sugar.

*) Lesion bypass catalyzed by the X-family DNA polymerases. To determine whether or not fPolλ can facilitate the bypass of DNA lesions, we performed a running start assay, similar to the assay described in Chapter 3 with Dpo4 and SsoPolB1, for a DNA substrate
containing an AP lesion. In this set of experiments, two DNA substrates, the undamaged control D-2 and the lesion containing substrate D-3 (100 nM each, Figure 11.8) were incubated with fPolλ (100 nM) and subsequently mixed with all four dNTPs (200 µM each) at 37 °C, for various times before quenching with EDTA. This D-3 substrate contained the same abasic (AP) site analog, tetrahydrofuran (X), used in Chapters 4 and 5, located in the template twenty-two nucleotides from the 3′-end. fPolλ incorporated dNTPs into the non-gapped D-2 control DNA substrate in a distributive manner, supporting one of its proposed functions as a gap-filling DNA polymerase (Figure 11.9). However, with the D-3 DNA substrate, fPolλ showed significant pausing one nucleotide preceding the AP site and directly opposite the lesion (Figure 11.9). In addition, even after a 50 minute incubation period, we failed to observe any significant elongation of the primer past the AP site by fPolλ. We further analyzed these results by quantitating the percentage of the time fPolλ encountered the AP lesion (% encountered) as well as the percentage of these encounters that lead to bypass of the lesion (% bypass). Here we defined an “encounter” as an event where the AP lesion is expected to be in the enzyme active site pocket, as evidenced by intermediate product formation, corresponding to incorporation one nucleotide before the lesion in addition to any incorporation opposite and downstream from the lesion while “bypass” events are simply all events where nucleotide incorporation occurs downstream of the lesion. We observed that after 50 minute incubation, fPolλ encountered only 46% of the available AP sites and bypassed only 5% of these encountered lesions. Besides a third pause site, similar results were observed for tPolλ, a truncated form of fPolλ lacking the N-terminal BRCT and Proline-rich domains. This enzyme encountered 34% of the total AP sites and although tPolλ was able to synthesize through 32% of these encounters, the elongation completely stalled at the third pause site and therefore tPolλ was not able to generate the full length bypass product (Figure 11.10). This suggested that the BRCT and Proline-rich domains did not
affect the poor AP lesion bypass ability of fPol\(\lambda\). However, it is premature to conclude that fPol\(\lambda\) cannot serve as a lesion bypass polymerase in the context of an AP site since it is possible that the presence of a downstream strand may facilitate such lesion bypass (see Chapter 10).

Similar assays were performed for X-family members DNA polymerase \(\mu\), TdT, and DNA polymerase \(\beta\). While DNA polymerase \(\mu\) was able to elongate through the AP site it generated four strong pause sites and was not able to synthesize the full length product even after three hours of incubation (Figure 11.11). Our results showed that DNA polymerase \(\mu\) bypassed 51% of the AP sites it encountered (91% encountered) but overall was not able to significantly extend past the fourth pause site at least partially due to its observed low processivity with the undamaged DNA control (Figure 11.11). In contrast, the nucleotide incorporation profile for TdT was drastically different than that observed for both DNA polymerases \(\lambda\) and \(\mu\). Here, we observed extensive elongation of the D3 substrate containing the AP lesion that was not perturbed by the presence of the lesion and failed to terminate after reaching what we would consider the full length bypass product, but instead continued to incorporate nucleotides to generate exceedingly long bypass products (Figure 11.12). This extensive elongation catalyzed by TdT suggested that the bypass of the lesion is in fact template independent, which is in agreement with evidence indicating that TdT lacks the ability to incorporate nucleotides in a template-dependent manner (35). Interestingly, of the X-family members assayed here, only DNA polymerase \(\beta\) was able to bypass the AP lesion and generate the full length bypass product in a template-dependent fashion (Figure 11.13). We determined that DNA polymerase \(\beta\) bypassed 67% of the AP sites it encountered (86% encountered). In order to further analyze AP bypass catalyzed by DNA polymerase \(\beta\), we used the SOSA (see Chapter 5) to quantitatively sequence these bypass events. Results from thirty-eight
independent colonies indicated that 55% of the time (21/38 colonies) DNA polymerase β generated a -1 deletion opposite the AP site, while the remaining 45% of the time DNA polymerase β inserted either a dTMP (11/38 colonies, 29%), an dAMP (5/38 colonies, 13%) or a dCMP (1/38 colonies, 3%) (Figure 11.14). Besides one -2 deletion opposite the AP site (1/38 colonies, 3%) and a triple deletion mutation following thymine incorporation opposite the lesion (1/38 colonies, 3%), the remaining incorporation events were essentially error-free both upstream and downstream from the lesion as determined from our sequencing window (see Chapter 5). In addition, the vast majority of bypass events containing a single deletion opposite the AP site, were followed by incorporation of dCMP (20/21 colonies containing a single deletion, 95%). We expect that this may be due to a primer-template slippage mechanism, as opposed to the flipping of the AP site lesion out of the active site of DNA polymerase β. However, additional studies (outlined in Chapter 4) to determine this are currently underway in our laboratory. Overall, our results here suggest that DNA polymerase β is capable of bypassing AP sites while DNA polymerases λ, μ, and TdT are very unlikely to bypass this lesion in vivo.

Analogous experiments were performed for these X-family DNA polymerases in the context of a cisplatin-modified DNA substrate D-5 and its control substrate D-4 (Figure 11.15) with the exception of TdT due to its template-independent DNA synthesis (see above). Cisplatin is an antitumor drug used to treat a variety of cancers by forming thousands of intrastrand and interstrand covalent adducts of DNA that can trigger apoptosis in severely damaged cells. Both DNA polymerases μ and β were able to bypass the lesion, but as observed above with an AP site, only DNA polymerase β was able to generate the full length bypass product (Figure 11.16). DNA polymerase μ showed significant pausing in the vicinity of the lesion and was only able to elongate the primer roughly six nucleotides downstream from the cisplatin lesion. While polymerase μ was
able to bypass 32% of the cisplatin lesions it encountered after an incubation period of three hours (Figure 11.17), polymerase β encountered 43% of the lesions and bypassed 83% of these lesions within sixty minutes. On the other hand, DNA polymerase λ was comparatively inefficient at bypassing the cisplatin lesion, generating four strong pause sites in the vicinity of the lesion and failing to incorporate nucleotides more than three nucleotides downstream of the lesion after one hour (Figure 11.18). More studies are needed to determine whether or not any of these X-family DNA polymerase members are competent to perform lesion bypass.

**11.4. Conclusions**

*Sulfolobus solfataricus* Dpo4 was used here as a model Y-family DNA polymerase to assess both the lesion bypass properties and replication fidelity inherent to the family. The studies presented here have utilized the techniques of pre-steady state kinetics to determine the fidelity of Dpo4 to be in the range of $10^{-3}$ to $10^{-4}$ with an undamaged DNA substrate. Interestingly, the ground-state binding affinity of correct nucleotides (70 to 230 µM) is 10- to 50-fold weaker than that observed for replicative DNA polymerases. Such a low affinity is consistent with the lack of interactions between Dpo4 and the bound nucleotides as revealed in the crystal structure of Dpo4, DNA, and an incoming nucleotide (see Figure 1.1, Chapter 1). The kinetic mechanism for single nucleotide incorporation into an undamaged DNA substrate was elucidated and shown to be characterized by an “induced-fit” mechanism for the selection of correct nucleotides. Three independent lines of evidence were used to show that Dpo4 undergoes a rate-limiting protein conformational change during the incorporation of a correct nucleotide.
We have also shown that the kinetic mechanism, rate-limiting step, and fidelity of Dpo4 do not change with temperature. This dissertation also explored the mechanisms used by Dpo4 to bypass DNA lesions. Our studies showed that Dpo4 can efficiently bypass an abasic (AP) site lesion although the rate constant was several orders of magnitude slower than the rate constant on an undamaged DNA control. Dpo4 predominantly uses one of two mechanisms to bypass the AP lesion: (i) the A-rule which involves incorporation of dAMP opposite the AP site or (ii) the lesion loop-out mechanism whereby Dpo4 is able to loop-out the AP site which results in the template information 5’ to the lesion to be shifted into the active site to direct nucleotide incorporation. Further studies presented here reveal that these mechanisms are sequence dependent. Our results demonstrate that the AP site affects the nucleotide incorporation efficiency of five to six downstream events. A novel sequencing assay (SOSA) was developed to determine the precise sequence of the bypass products synthesized by Dpo4. The results from the SOSA correlated extremely well with our kinetic results and results from a restriction digestion assay verifying the existence of these two mechanisms to bypass the AP site. In addition, we used the combination of pre-steady state kinetics and crystallography to determine the mechanism of template-independent nucleotide incorporation, an activity hypothesized to increase the genomic hypermutability of retroviruses including human immunodeficiency viruses, catalyzed by Dpo4. These studies indicated that the first blunt-end dATP incorporation was 80-fold more efficient than the second, and that among natural dNTPs, dATP was the preferred substrate due to its stronger intrahelical base-stacking interactions. A 2.05 Å crystal structure from our collaborators Hong Ling and Wei Yang of Dpo4•(blunt-end DNA)•ddATP revealed that the base and the sugar of the incoming ddATP, respectively, stack against the 5’-base of the opposite strand and 3’-base of the elongating strand.
Human DNA polymerase λ, a recently discovered member of the X-family DNA polymerases, has been hypothesized to be a second polymerase involved in the base-excision repair pathways. Using pre-steady state kinetics, the C-terminal DNA polymerase β-like domain, lacking the N-terminal BRCT and proline-rich domains, was shown to have a fidelity of $10^{-2}$ to $10^{-4}$. Surprisingly, the ground-state binding affinity of correct (1.1 to 2.4 μM) and incorrect nucleotides (1.4 to 8.4 μM) was very similar while correct nucleotides (3 to 6 s$^{-1}$) were incorporated much faster than incorrect nucleotides (0.001 to 0.2 s$^{-1}$). These studies also revealed that an incoming nucleotide bound to the enzyme•DNA binary complex at a rate constant of $1.6 \times 10^8$ M$^{-1}$s$^{-1}$ to form the ground-state ternary complex while the nucleotide dissociated from this complex at a rate constant of 300 s$^{-1}$. Strikingly, further analysis reveals that the full length DNA polymerase λ has almost identical fidelity to that of DNA polymerase β suggesting that the previously mentioned C-terminal DNA polymerase β-like domain of polymerase λ is dramatically, up to 100-fold lower in fidelity. The results in this work demonstrate that the non-enzymatic proline-rich domain confers the increase in fidelity by significantly lowering the incorporation rate constants of incorrect nucleotides, illustrating a novel mechanism in which the polymerase fidelity is controlled not by an accessory protein or proofreading exonuclease domain, but by an internal regulatory domain. The studies presented here also investigate the effects of different DNA substrates on DNA synthesis catalyzed by DNA polymerase λ. In these studies, a single-nucleotide gapped DNA substrate containing a 5’-terminal 2-deoxyribose-5-phosphate moiety, thought to be the in vivo substrate for DNA polymerase λ, was compared to a series of distinct DNA substrates. Our results show that the downstream strand and its 5’-phosphate were both found to increase incorporation efficiency by 15- and 11-fold respectively, reinforcing the importance of the downstream strand on catalytic efficiency. Yet our results demonstrate that while the stable mimic of the 5’-terminal 2-deoxyribose-5-phosphate moiety did not
affect the fidelity of DNA polymerase λ, it moderately decreased the polymerase efficiency by 3.4-fold, and therefore, the more commonly used DNA substrate containing only the 5’-phosphate group on the downstream strand is a reasonable model substrate to be employed for studies of this enzyme and that of X-family member DNA polymerase β.

This dissertation also demonstrated via the novel SOSA, that of the three human Y-family members assayed, both DNA polymerase η and κ appear to facilitate AP site bypass in vivo while the results presented here suggest it is much less likely that DNA polymerase ι performs this task in vivo due to its extremely mutagenic bypass of an AP lesion.

It has become apparent from the studies described in this dissertation, that DNA polymerases which have likely evolved from a common ancestor, can accomplish vastly different biological functions and specifically utilize distinct mechanisms to catalyze nucleotide incorporation in a variety of sequence contexts.
11.5 Figures, Tables, and Schemes

**Adenine Incorporation**

\[
\begin{array}{c}
G \\
C \\
A \\
G \\
\end{array}
\]  
\[
\begin{array}{c}
G \\
G \\
A \\
C \\
\end{array}
\]

CGAGCCGTCCAACCAACTCAACGTCGATCCAATGCCGTCCTGAGTCCAGG (11/39)

C

A

TG

G

(1/39)

(1/39)

(1/39)

(2/39)

(1/39)

(1/39)

**Guanine Incorporation**

\[
\begin{array}{c}
C \\
A \\
G \\
T \\
\end{array}
\]

CGAGCCGTCCAACCAACTCAACGTCGATCCAATGCCGTCCTGAGTCCAGG (11/39)

A

G

C

(1/39)

(1/39)

(1/39)

(1/39)

**Cytosine Incorporation**

CGAGCCGTCCAACCAACTCAACGTCGATCCAATGCCGTCCTGAGTCCAGG (1/39)

**Thymine Incorporation**

CGAGCCGTCCAACCAACTCATCGTCGATCCAATGCCGTCCTGAGTCCAGG (2/39)

**No incorporation**

CGAGCCGTCCAACCAACTCAACGTCGATCCAATGCCGTCCTGAGTCCAGG (0/39)

A

T

C

(1/39)

(1/39)

(1/39)

Figure 11.1. Mutation spectrum of AP site bypass catalyzed by human DNA polymerase η at 37 °C. See legend of Figure 5.2 for scheme explanation.
Figure 11.2. Mutation spectrum of AP site bypass catalyzed by mouse DNA polymerase κ at 37 ºC. See legend of Figure 5.2 for scheme explanation.
**Figure 11.3. Comparison of the linker sequences of Dpo4 and Dbh.** The sequences are shown with the N-terminal amino acid on the left.

Dpo4  RDEYNEPIRTRVRK
Dbh   RNKYSEPVENKSKI
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Linker Sequence</th>
<th>$K_d$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>RDEYNEPIRTRVRK</td>
<td>10</td>
</tr>
<tr>
<td>1-GLY addition</td>
<td>RDEYNEPIGRTRVRK</td>
<td>118</td>
</tr>
<tr>
<td>4-GLY addition</td>
<td>RDEYNEPIGGGTRVRK</td>
<td>279</td>
</tr>
<tr>
<td>6-GLY addition</td>
<td>RDEYNEPIGGGGGTRVRK</td>
<td>674</td>
</tr>
<tr>
<td>1-deletion</td>
<td>RDEYN-RTRVRK</td>
<td>159</td>
</tr>
<tr>
<td>4-deletion</td>
<td>RDEY--RTRVRK</td>
<td>304</td>
</tr>
<tr>
<td>6-deletion</td>
<td>RD------RTRVRK</td>
<td>1688</td>
</tr>
</tbody>
</table>

**Figure 11.4. DNA binding affinity for the Dpo4 linker mutants.** The glycine additions are designated as “G” and the deleted amino acid residues are designated as “-“.
Figure 11.5. dTTP incorporation into the single-nucleotide gapped D-1 DNA substrate (see Figure 11.6) under burst conditions.
Figure 11.6. Single-nucleotide gapped DNA substrates (21-19/41-mer). The upstream primer 21-mer was 5’-32P-labeled. The downstream primer 19-mer was 5’-phosphorylated. The “Y” designates A, G, T, C in DNA substrates D-1, D-6, D-7, and D-8, respectively.

5′-CGCAGCGGTCCAACCAACTCA CGTCGATCCAATGCCGTCC-3′
3′-GCCTCGGCAGGTGTTGATYGCAGCTAGTTACGCGAGG-5′
Figure 11.7. Pulse-chase and pulse-quench assays for DNA polymerase λ.
D-2 DNA Substrate

5’ - CGCAGCGTGCCAACCAA - 3’
3’ - GCGTCGCGGCAGTTGTTGAGTACGTACAGCTTTACGGCAGG - 5’

D-3 DNA Substrate

5’ - CGCAGCGTGCCAACCAA - 3’
3’ - GCGTCGCGGCAGTTGTTGAGTXGCAGCTTTACGGCAGG - 5’

Figure 11.8. Control (D-2) and AP site (D-3) containing DNA substrates for running start assays with the X-family DNA polymerases. The “X” designates an AP site analog, tetrahydrofuran, which is located 22 nucleotides from the 3’-end.
Figure 11.9. Elongation of 17/41-mer by human fPolλ: (A) D-2; (B) D-3. D-2 is an undamaged DNA control and D-3 contains an AP site as described in the legend of Figure 11.8.
Figure 11.10. Elongation of 17/41-mer by human tPolα: (A) D-2; (B) D-3. D-2 is an undamaged DNA control and D-3 contains an AP site as described in the legend of Figure 11.8.
Figure 11.11. Elongation of 17/41-mer by human DNA polymerase μ: (A) D-2; (B) D-3. D-2 is an undamaged DNA control and D-3 contains an AP site as described in the legend of Figure 11.8.
Figure 11.12. Elongation of 17/41-mer by TdT: (A) D-2; (B) D-3. D-2 is an undamaged DNA control and D-3 contains an AP site as described in the legend of Figure 11.8.
Figure 11.13. Elongation of 17/41-mer by human DNA polymerase β: (A) D-2; (B) D-3. D-2 is an undamaged DNA control and D-3 contains an AP site as described in the legend of Figure 11.8.
**No Incorporation**

```
CGCAGCCGTCAGACCAAACCACTCA-CTCGATCCAAATGGCCTCCTGAGTCCAGG
```

**Thymine Incorporation**

```
CGCAGCCGTCAGACCAAACCACTCATCGTCGATCCAAATGGCCTCCTGAGTCCAGG
```

**Adenine Incorporation**

```
CGCAGCCGTCAGACCAAACCACTCAACGTCGATCCAAATGGCCTCCTGAGTCCAGG
```

**Cytosine Incorporation**

```
CGCAGCCGTCAGACCAAACCACTCACCGTCGATCCAAATGGCCTCCTGAGTCCAGG
```

Figure 11.14. Mutation spectrum of AP site bypass catalyzed by human DNA polymerase β at 37 ºC. See legend of Figure 5.2 for scheme explanation.
D-4 DNA Substrate

5’ -GTCCCTGTTCGGGCGCC-3’
3’ -CAGGGACAAGCCCGCGGTCCTCTGGTCTCCGATCAGACACTAG-5’

D-5 DNA Substrate

5’ -GTCCCTGTTCGGGCGCC-3’
3’ -CAGGGACAAGCCCGCGGTCCTCTGGTCTCCGATCAGACACTAG-5’

Figure 11.15. Control (D-4) and cisplatin-modified (D-5) DNA substrates for running start assays with the X-family DNA polymerases. The boldface “GG” designates the location of the cisplatin modification located at position 24 and 25 on the DNA template from the 3’-end.
Figure 11.16. Elongation of 17/44-mer by human DNA polymerase β: (A) D-4; (B) D-5. D-4 substrate is an undamaged DNA control and D-5 contains a cisplatin-modification as described in the legend of Figure 11.15.
Figure 11.17. Elongation of 17/44-mer by human DNA polymerase \( \mu \): (A) D-4; (B) D-5. D-4 substrate is an undamaged DNA control and D-5 contains a cisplatin-modification as described in the legend of Figure 11.15.
Figure 11.18. Elongation of 17/44-mer by fPol: (A) D-4; (B) D-5. D-4 substrate is an undamaged DNA control and D-5 contains a cisplatin-modification as described in the legend of Figure 11.15.
Table 11.1. Kinetic parameters of matched dNTP and rNTP incorporation into DNA catalyzed by Dpo4 and the Dpo4 Y12A mutant.

<table>
<thead>
<tr>
<th>NTP</th>
<th>Template Base</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (μM$^{-1}$s$^{-1}$)</th>
<th>Wt Dpo4 Sugar Selectivity$^a$</th>
<th>Y12A Sugar Selectivity$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dTTP$^c$</td>
<td>A (D-1)</td>
<td>230</td>
<td>9.4</td>
<td>$4.1 \times 10^{-2}$</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>UTP$^c$</td>
<td>A (D-1)</td>
<td>1141</td>
<td>0.002</td>
<td>$2.1 \times 10^{-6}$</td>
<td>19,700</td>
<td>NA</td>
</tr>
<tr>
<td>dTTP$^d$</td>
<td>A (D-1)</td>
<td>283</td>
<td>4.3</td>
<td>$1.5 \times 10^{-2}$</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>UTP$^d$</td>
<td>A (D-1)</td>
<td>618</td>
<td>0.33</td>
<td>$5.3 \times 10^{-4}$</td>
<td>NA</td>
<td>29</td>
</tr>
<tr>
<td>rCTP$^d$</td>
<td>G (D-6)</td>
<td>751</td>
<td>0.67</td>
<td>$8.9 \times 10^{-4}$</td>
<td>NA</td>
<td>24</td>
</tr>
<tr>
<td>rGTP$^d$</td>
<td>C (D-8)</td>
<td>396</td>
<td>0.25</td>
<td>$6.3 \times 10^{-4}$</td>
<td>NA</td>
<td>17</td>
</tr>
</tbody>
</table>

$^a$Calculated as $(k_p/K_d)_{\text{correct dNTP}}/(k_p/K_d)_{\text{matched rNTP}}$.

$^b$Calculated as $(k_p/K_d)_{\text{correct dTTP into D-1}}/(k_p/K_d)_{\text{matched rNTP}}$.

$^c$Incorporation catalyzed by wild type Dpo4.

$^d$Incorporation catalyzed by the Dpo4 Y12A mutant.

Note – Not applicable (NA).
Table 11.2. DNA substrates for Dpo4 incorporation assays (21/41-mer). The primer 21-mer was 5'-32P-labeled. The “Y” designates A, G, T, C in DNA substrates D-1, D-6, D-7, and D-8, respectively.
Table 11.3. Kinetic parameters of matched rNTP incorporation into single-nucleotide gapped DNA catalyzed by fPolλ.

<table>
<thead>
<tr>
<th>rNTP</th>
<th>Template Base</th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s⁻¹)</th>
<th>$k_p/K_d$ (μM⁻¹s⁻¹)</th>
<th>Fidelityᵃ</th>
<th>Sugar Selectivityᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTP</td>
<td>A (D-1)</td>
<td>118</td>
<td>0.030</td>
<td>2.5 x 10⁻⁴</td>
<td>1.7 x 10⁻⁴</td>
<td>6,000</td>
</tr>
<tr>
<td>rCTP</td>
<td>G (D-6)</td>
<td>52</td>
<td>0.020</td>
<td>3.9 x 10⁻⁴</td>
<td>2.2 x 10⁻⁴</td>
<td>4,615</td>
</tr>
<tr>
<td>rATP</td>
<td>T (D-7)</td>
<td>139</td>
<td>0.0045</td>
<td>3.2 x 10⁻⁵</td>
<td>2.0 x 10⁻⁵</td>
<td>50,000</td>
</tr>
<tr>
<td>rGTP</td>
<td>C (D-8)</td>
<td>36</td>
<td>0.0031</td>
<td>8.6 x 10⁻⁵</td>
<td>7.2 x 10⁻⁵</td>
<td>13,953</td>
</tr>
</tbody>
</table>

ᵃCalculated as $(k_p/K_d)_{\text{matched rNTP}}/[(k_p/K_d)_{\text{correct dNTP}} + (k_p/K_d)_{\text{matched rNTP}}].$

ᵇCalculated as $(k_p/K_d)_{\text{correct dNTP}}/(k_p/K_d)_{\text{matched rNTP}}.$

Data of correct dNTPs are from Chapter 8.
<table>
<thead>
<tr>
<th></th>
<th>$K_d$ (μM)</th>
<th>$k_p$ (s$^{-1}$)</th>
<th>$k_p/K_d$ (μM$^{-1}$s$^{-1}$)</th>
<th>Fidelity$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dTTP$^b$</td>
<td>2.6</td>
<td>3.9</td>
<td>1.5</td>
<td>1</td>
</tr>
<tr>
<td>UTP</td>
<td>118</td>
<td>0.030</td>
<td>$2.5 \times 10^{-4}$</td>
<td>$1.7 \times 10^{-4}$</td>
</tr>
<tr>
<td>rCTP</td>
<td>309</td>
<td>0.0036</td>
<td>$1.2 \times 10^{-5}$</td>
<td>$8.0 \times 10^{-6}$</td>
</tr>
<tr>
<td>rGTP</td>
<td>313</td>
<td>0.0061</td>
<td>$2.0 \times 10^{-5}$</td>
<td>$1.3 \times 10^{-5}$</td>
</tr>
<tr>
<td>rATP</td>
<td>No observed incorporation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Calculated as $(k_p/K_d)_{NTP}/[(k_p/K_d)_{dTTP} + (k_p/K_d)_{NTP}]$.

$^b$Data from Chapter 8.

Table 11.4. Kinetic parameters of mismatched rNTP incorporation into single-nucleotide gapped D-1 catalyzed by fPol$\lambda$. 
Scheme 11.1. Detailed Kinetic Mechanism for dNTP Incorporation Catalyzed by fPolλ.
11.6 References

List of References


421


425


Lindahl, T., and Nyberg, B. (1972) Biochemistry 11, 3610-3618

Lindahl, T., and Andersson, A. (1972) Biochemistry 11, 3618-3623


Kobayashi, Y., Watanabe, M., Okada, Y., Sawa, H., Takai, H., Nakanishi, M., Kawase,

34685-34690

Lee, J. W., Blanco, L., Zhou, T., Garcia-Diaz, M., Bebenek, K., Kunkel, T. A., Wang, Z.,


11891-11900

Johnson, A. A., Tsai, Y., Graves, S. W., and Johnson, K. A. (2000) Biochemistry 39,
1702-1708

Eger, B. T., Kuchta, R. D., Carroll, S. S., Benkovic, P. A., Dahlberg, M. E., Joyce, C. M.,

331 (Pt 1), 79-87


Maciejewski, M. W., Shin, R., Pan, B., Marintchev, A., Denninger, A., Mullen, M. A.,

946


432


