2018

Development of Novel Phenanthroline and Thiazole Orange Derived G-quadruplex Ligands and Telomerase Inhibitors

Siwen Wang
University of the Pacific, wangsiwentmxk@163.com

Follow this and additional works at: https://scholarlycommons.pacific.edu/uop_etds

Part of the Chemistry Commons, and the Pharmacy and Pharmaceutical Sciences Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Graduate School at Scholarly Commons. It has been accepted for inclusion in University of the Pacific Theses and Dissertations by an authorized administrator of Scholarly Commons. For more information, please contact mgibney@pacific.edu.
DEVELOPMENT OF NOVEL PHENANTHROLINE AND THIAZOLE ORANGE DERIVED G-QUADRUPLEX LIGANDS AND TELOMERASE INHIBITORS

by

Siwen Wang

A Dissertation Submitted to the
Graduate School
In Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Thomas J. Long School of Pharmacy and Health Sciences
Pharmaceutical and Chemical Sciences

University of the Pacific
Stockton, CA
2018
DEVELOPMENT OF NOVEL PHENANTHROLINE AND THIAZOLE ORANGE DERIVED G-QUADRUPLEX LIGANDS AND TELOMERASE INHIBITORS

by

Siwen Wang

APPROVED BY:

Dissertation Advisor: Liang Xue, Ph.D

Committee Member: Andreas Franz, Ph.D.

Committee Member: Jianhua Ren, Ph.D.

Committee Member: Vyacheslav V. Samoshin, Ph.D.

Committee Member: Lisa Wrischnik, Ph.D.

Department Co-Chair: Jianhua Ren, Ph.D.

Dean of Graduate School: Thomas H Naehr, Ph.D.
DEVELOPMENT OF NOVEL PHENANTHROLINE AND THIAZOLE ORANGE DERIVED G-QUADRUPLEX LIGANDS AND TELOMERASE INHIBITORS

Copyright 2018

by

Siwen Wang
DEDICATION

This dissertation is dedicated to my beloved family, especially my parents and my husband.

for their endless love, and encouragement.
ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor Dr. Liang Xue. I would not have been able to reach this point without his guidance and encouragement; not to mention time and effort he spent on improving my dissertation.

I would like to thank Dr. Jianhua Ren; Dr. Vyacheslav Samonshin; Dr. Andy Franz and Dr. Lisa Wrischnik for nicely agreeing to serve on my dissertation committee and providing valuable feedback, direction and assistance to my dissertation writing.

Thank you to Dr. Irina Dotsenko; Dr. Dazhou Yang and Dr. Wanbo Liu for their outstanding organic synthesis work for the present project.

Thank you to Dr. William Chan for his help with cell culture and related assays.

Thank you to Ryan Hekman and Yuntao Zhang for their help with ESI-MS spectrometry.

I also would like to extend my greatest appreciation to Ms. Susan McCann, Dr. Brett William, for always taking care of me as well as all the graduate students. My doctoral study would not be this smooth without you.

Sincerest thanks go to Brock Allen for being my best friend when I have been living and studying in a foreign country. You have made the office and lab a warm place that I am enjoying every single day.
Sincerest thanks go to all the other members from Dr. Xue’s research group, Mandeep, Mingheng, Danni, and all other the undergraduate students for always willing to offer a helping hand.

I would like to especially thank my dad Yuming Wang for his unconditional support and encouragement. Thank you to my mom Aiping Zhang for her love and care. A big thank you to my husband Liang Ma. I am so grateful that I have you by my side for these past ten years. You make my life brighter every day.
Development of Novel Phenanthroline and Thiazole Orange Derived G-quadruplex Ligands and Telomerase Inhibitors.

Abstract

by Siwen Wang

University of the Pacific
2018

The end of the human chromosome is protected by telomeres which contain a special tandem guanine-rich DNA sequence, 5’-TTAGGG. The length of telomeres is shortened during cell replications, and its length limits the replication capacity of cells. Telomerase is over-expressed in 85–90% of cancer cells, responsible for extending the telomere length in cancer cells. Guanine-rich DNA sequence can self-assemble into unique G-quadruplex structures that interfere with the extension of telomeres by telomerase. Therefore, DNA G-quadruplex has recently received much attention because of its important regulatory functions in telomerase-mediated cancerization. The formation of G-quadruplex requires monovalent cations (Na⁺ and K⁺) or small molecules known as G-quadruplex ligands. In the present work, we developed a serial of G-quadruplex ligands by tethering side-chains to two core structures: 1,10-phenanthroline (Phen) and thiazole orange (TO). Biophysical studies including DNA thermal denaturation monitored
by fluorescence or circular dichroism, fluorometric titration, and ESI-MS spectrometry reveal that the binding of the synthesized ligands to G-quadruplex is side-chain dependent. The arylsulfanyl side chains significantly improve the binding affinity and selectivity of 1,10-phenanthroline towards G-quadruplex over duplex DNA. The polyamine side chains are a suitable structural motif for remarkable G-quadruplex binding affinity based on the results from both Phen and TO derivatives. These ligands greatly inhibit the telomerase activity in vitro, determined by a modified telomeric repeat amplification protocol (TRAP) assay. Amongst these promising telomerase inhibitors, a thiazole orange derivative containing a side chain of spermine shows an outstanding telomerase inhibition effect at nanomolar concentrations, which is comparable to the most effective synthetic telomerase inhibitors, BRACO-19.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................ 14

LIST OF FIGURES ..................................................................................................... 15

LIST OF ABBREVIATIONS ......................................................................................... 24

Chapter 1. Introduction to DNA ................................................................................ 26

1.1. DNA Basics........................................................................................................ 26

1.1.1. Central dogma ............................................................................................. 26

1.1.2. The history of DNA discovery ...................................................................... 28

1.2. DNA Structures and Functions ....................................................................... 30

1.2.1. DNA structure basics ................................................................................. 30

1.2.2. DNA duplex ................................................................................................. 34

1.2.3. DNA triplex .................................................................................................. 39

1.2.4. DNA G-quadruplex ...................................................................................... 42

Chapter 2. Small Molecule-DNA Non-covalent Interactions .................................. 62

2.1. Duplex Binding ................................................................................................. 64

2.1.1. Outside edge binding (nonspecific binding) .................................................. 65

2.1.2. DNA intercalation ....................................................................................... 66

2.1.3. Duplex DNA groove binding ...................................................................... 69

2.1.4. DNA intercalation vs. groove binding ......................................................... 72
Chapter 2. G-quadruplex Binding

2.2. G-quadruplex Binding

- 2.2.1. G-quadruplex end stacking ligands
- 2.2.2. G-quadruplex groove binders

2.3. 1,10-Phenanthroline Analogues Metal Complexes as G-quadruplex Ligands

- 2.3.1. 1,10-phenanthroline analogues as G-quadruplex ligands
- 2.3.2. 1,10-phenanthroline analogues metal complexes as G-quadruplex ligands

2.4. TO Derivatives as G-quadruplex Ligands

Chapter 3. 5-Substituted 1,10-Phenanthroline Ni(II) Complexes as G4 Ligands

3.1. Introduction

3.2. Experimental

- 3.2.1. Materials and general methods
- 3.2.2. Mass spectrometry analysis of complexes
- 3.2.3. Fluorescent intercalator displacement assay
- 3.2.4. Preparation of cell lysate for TRAP assay
- 3.2.5. Real-time PCR telomeric repeat amplification protocol (TRAP) assay
- 3.2.6. 32P telomeric repeat amplification protocol (TRAP) assay
- 3.2.7. MTT cell viability assay

3.3. Results and Discussion
3.3.1. 5-substituted 1,10-phenanthroline formed 1-2 and 1-3 complexes with Ni(II) ions................................................................. 100

3.3.2. 5-Substituted Phen-Ni complexes induced telomeric G-quadruplex formation under Na\(^+\) and K\(^+\) deficient conditions................................................. 109

3.3.3. 5-Substituted Phen-Ni complexes with cationic and aromatic side chains increased the thermal stability of telomeric G-quadruplex DNA.......................110

3.3.4. Aromatic side-chains enhanced the stacking affinity and selectivity of 5-substituted Phen-Ni complexes against telomeric G-quadruplex DNA. ..............116

3.3.5. 5-Substitued Phen-Ni complexes inhibited telomerase activity in vitro. ................................................................................................................................. 123

3.3.6. 5-Substitued Phen-Ni complexes inhibited telomerase positive cancer cell growth. .............................................................................................................. 141

3.4 Conclusion .................................................................................................................................................................................. 144

Chapter 4. Development of Thiazole Orange Derived G-quadruplex Ligands and Telomerase Inhibitors.............................................................................................................. 146

4.1 Introduction .................................................................................................................................................................................. 147

4.2 Experimental.................................................................................................................................................................................. 151

4.2.1 Materials and general methods. ................................................................. 151

4.2.2. Circular dichroism G-quadruplex denaturation analysis. .............. 152

4.2.3. Fluorometric scan of TO derivatives interacting with G-quadruplex. .................................................................................................................. 152
4.2.4. Fluorometric titration of TO derivatives interacting with G-quadruplex and duplex DNA. ................................................................. 153

4.2.5. Mass Spectrometry. ................................................................. 153

4.2.6. TRAP assay. ........................................................................ 154

4.2.7. MTT cell viability. ................................................................. 155

4.3 Results and Discussion ................................................................ 156

4.3.1. Thiazole orange derivatives stabilized human telomeric G-quadruplexes. ................................................................. 156

4.3.2. Thiazole orange derivatives preserve G-quadruplex light-up probe properties ........................................................................ 162

4.3.3. Thiazole orange derivatives interact with human telomeric G-quadruplex with enhanced affinities. ........................................... 166

4.3.4. Thiazole orange derivatives associate with human telomeric G-quadruplexes with various ratios. ................................................. 171

4.3.5. Thiazole orange derivatives form complexes with human telomeric G-quadruplexes with various binding modes ........................................... 184

4.3.6. Thiazole orange derivatives inhibited the telomerase activity in vitro. ...................................................................................... 195

4.3.7. Positively charged side chains decrease the cytotoxicity of thiazole orange ................................................................................ 203

4.4. Conclusion .................................................................................... 205
LIST OF TABLES

Table 1-1. Summary and comparison of double-helical DNA structure parameters. .... 39
Table 1-2. Telomerase activity detected in different type of cancer ......................... 59
Table 3-1. 5-Substituted 1,10-phenanthroline derivatives. ....................................... 95
Table 3-2. Characteristic m/z values and the relative abundance (%) in ESI-MS spectra of
the 1,10-Phen-Ni complexes ..................................................................................... 103
Table 3-3. T<sub>m</sub> values of G-quadruplex derived from fluorescent thermal denaturation in
the presence of free compound or Ni<sup>2+</sup> complexes at a 4:1 ligand/DNA ratio. ............112
Table 3-4. Approximate pKa value of conjugate acid of the designated side chains of 5-
substituted phenanthroline. ......................................................................................115
Table 4-1. T<sub>m</sub> values (°C) of G4 in the presence of TO derevetives. .........................159
Table 4-2. Fluorescence intensities (au) of TO derivative-G-quadruplex complexes. ...165
Table 4-3. The ratio of K<sub>d</sub> values between G-quadruplex and dsDNA ...................... 171
Table 4-4. IR values for TO derivatives binding with the G4 and the corresponding
complex compositions. ......................................................................................... 181
Table 4-5. The ratio of IR values between CP9 binding G-quadruplex and dsDNA .... 184
LIST OF FIGURES

Figure 1-1. The Central dogma of molecular biology.......................................................... 26
Figure 1-2. The structures of nucleobases present in human genome................................. 31
Figure 1-3. The two conformations of N-glycosidic bonds................................................. 32
Figure 1-4. Conformations of deoxyribose sugar pucker................................................... 33
Figure 1-5. An example of the primary structure of DNA.................................................. 33
Figure 1-6. Watson-Crick base pairing............................................................................... 35
Figure 1-7 DNA double helix and two grooves................................................................. 36
Figure 1-8. Three DNA helices......................................................................................... 37
Figure 1-9. Hoogsteen base pairs in DNA triplexes........................................................... 40
Figure 1-10. DNA triplex.................................................................................................... 41
Figure 1-11. The top view of a DNA G-quartet (left) and a sketch of a G-quadruplex (right).................................................................................................................. 43
Figure 1-12. Hoogsteen side and Watson-Crick side for guanine hydrogen bonding.......... 44
Figure 1-13. G-quadruplexes with coordinated cations. (A) A parallel G-quadruplex with Na^+ ions, (B) A parallel G-quadruplex with K^+ ions......................................................... 45
Figure 1-14. Illustrative examples of base stacking mode in G-quadruplexes..................... 46
Figure 1-15. Various topologies for G-quadruplexes......................................................... 47
Figure 1-16. Glycosidic bonds conformations in various G-quadruplex topologies. Syn- guanines are shown in yellow and anti-guanines are in grey........................................... 48
Figure 1-17. Structures formed upon replication and transcription of regions containing G-quadruplex motifs ................................................................. 50
Figure 1-18. The structures of telomeres and telomeric sheltering complexes. ........... 52
Figure 1-19. An illustration of DNA end replication problem ................................... 54
Figure 1-20. Telomerase components. ................................................................... 57
Figure 2-1. The interaction of histone and DNA. .................................................... 62
Figure 2-2. Mechanisms of nitrogen mustard (top) and cisplatin (bottom) DNA covalent interactions ............................................................................. 63
Figure 2-3. Structures of spermine and spermidine ................................................. 66
Figure 2-4. Structures of DNA intercalative ligands ................................................ 69
Figure 2-5. Structures of dsDNA groove binders. .................................................... 71
Figure 2-6. Molecules with two binding modes against DNA. ................................... 72
Figure 2-7. Structures of G4 ligands ....................................................................... 75
Figure 2-8. Structures of G-quadruplex groove binders ......................................... 79
Figure 2-9. Structures of 1,10-phenanthroline (left) and 2,2’-bipyridine (right) .......... 80
Figure 2-10. Structures of Phen-DC3 (left) and its cationic amino side chain derivatives (middle); 4,7-diamino-1,10-Phen derivative (right). ............................... 82
Figure 2-11. Structures of 1,10-phenanthroline derivatives ..................................... 84
Figure 2-12. Structures of thiazole orange and derivatives ....................................... 88
Figure 3-1. Job plots obtained based on continuous variation analysis for (A) 20-Ni and (B) 21-Ni ............................................................................... 101
Figure 3-2. ESI spectra of phenanthroline and Ni^{2+} solutions with a 2:1 molar ratio in three different solvent compositions ........................................... 102
Figure 3-3. CID spectrum of 20-Ni complex ion [1-2-Cl]⁺ at m/z 714.8. .......................... 104
Figure 3-4. CID spectrum of 20-Ni complex ion [1-3]²⁺ at m/z 495.4. .......................... 106
Figure 3-5. CID spectrum of 21-Ni complex ion [1-2-Cl]⁺ at m/z 739.2. .......................... 107
Figure 3-6. CID spectrum of 21-Ni complex ion [1-3]²⁺ at m/z 513.8. .......................... 108
Figure 3-7. CD spectra of PCG1 measured when titrating with A) 20-Ni; B) 20; C) 21-Ni; D) 21. The continuous changes in the CD spectra correspond to the incremental amount of Phen-Ni ranging from 0 -20 µM. DNA = 5 µM. ........................................... 110
Figure 3-8. ΔT_m (T_m increment compared to the compound free control 57.7 °C) values of G-quadruplex derived from fluorescent thermal denaturation in the presence of Phen-Ni at a 4:1 ligand/DNA ratio. π: Complexes with aromatic side chains; +: Complexes with cationic side chains, and the pKa value of conjugate acid is between 2.0 and 7.0; ++: Complexes with cationic side chains, and the pKa value of conjugate acid is higher than 7.0. .................................................................................................................. 113
Figure 3-9. The G⁴DC₅₀ values of all Phen-Ni complexes from the FID assay. ............. 118
Figure 3-10. The G⁴DC₅₀ values of free 20, 21 and 20-Ni, 21-Ni determined by FID assay. ................................................................................................................................. 120
Figure 3-11. The G⁴DC₅₀ and Δ⁴DC₅₀ values of the Phen-Ni complexes from the FID assay. ................................................................................................................................. 121
Figure 3-12. Scheme of a TRAP (Telomeric Repeat Amplification Protocol) assay for telomerase activity detection. ................................................................................. 124
Figure 3-13. The results of real-time PCR TRAP concentration-dependent inhibition profile for TMPyP4. .................................................................................................................. 126
Figure 3-14. Ct values acquired from the real-time PCR TRAP assay for twenty-one
complexes. ................................................................. 128

Figure 3-15. Ct values acquired from the real-time PCR TRAP assay for the
centration studies of effective inhibitors. ................................................................. 129

Figure 3-16. Telomerase inhibition factors calculated from the real-time PCR TRAP
assay for effective inhibitors. .................................................................................... 130

Figure 3-17. Illustration of the telomerase-extended products in a TRAP assay. ........ 132

Figure 3-18. $^{32}\text{P}$ TRAP reactions for all Phen-Ni complexes (50 µM). Reactions in the
absence or presence of Ni$^{2+}$ (50 µM), BIBR (50 µM), and TMPyP4 (100 µM) were
carried out as controls. IS: PCR product as internal standard (TSNT as template). The
reaction in the right lane has no telomerase................................................................. 134

Figure 3-19. $^{32}\text{P}$ TRAP reactions in the presence of 13 (left) at concentrations of 50, 25,
12.5, 6.25 and 3.125 µM (from left to right). $^{32}\text{P}$ TRAP reactions in the presence of 13-Ni
(right) at concentrations of 50, 40, 30, 20, 15 and 10 µM (from left to right). Reactions in
the presence of 13-Cu$^{2+}$ (50 µM), Ni$^{2+}$ (50 µM), BIBR (50 µM), and TMPyP4 (100 µM)
were carried out as controls. IS: PCR product as internal standard (TSNT as template).
............................................................................................................................................... 135

Figure 3-20. $^{32}\text{P}$ TRAP reactions in the presence of 14-Ni at concentrations of 50, 25,
12.5, 6.25 and 3.125 µM (from left to right). Reactions in the presence of 14 (50 µM), 14-
Cu$^{2+}$ (50 µM), Ni$^{2+}$ (50 µM) and BIBR (50 µM) were carried out as controls. IS: PCR
product as internal standard (TSNT as template).......................................................... 136

Figure 3-21. $^{32}\text{P}$ TRAP reactions in the presence of 8-Ni at concentrations of
50,45,40,35,30,25 and 20 µM (from left to right). Reactions in the presence of 8 (50 µM),
BIBR (50 µM) and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as
internal standard (TSNT as template). ................................................................. 137

Figure 3-22. $^{32}$P TRAP reactions in the presence of 17-Ni at concentrations of 30, 25, 20, 15 and 10 µM (from left to right). Reactions in the presence of 17 (50 µM), BIBR (50 µM) and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template). ................................................................. 138

Figure 3-23. $^{32}$P TRAP reactions in the presence of 3-Ni at concentrations of 50, 40, 30, 20, 10 and 5 µM (from left to right). Reactions in the presence of 3 (50 µM), BIBR (50 µM) and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template). ................................................................. 139

Figure 3-24. $^{32}$P TRAP reactions in the presence of 20-Ni (left) and 21-Ni (right) at concentrations of 50, 25, 12.5, 6.25 and 3.125 µM (from left to right). Reactions in the presence of 20 or 21 (50 µM), 20-Cu$^{2+}$ or 21-Cu$^{2+}$ (50 µM), Ni$^{2+}$ (50 µM), BIBR (50 µM) and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template). ................................................................. 141

Figure 3-25. Cell proliferation of two cancer cell lines in the presence of 5-substituted Phen-Ni complexes, determined by an MTT assay. ................................................................. 143

Figure 3-26. Cell proliferation of HeLa cells in the presence of free 5-substituted phenanthrolines determined by an MTT assay. The cell proliferation in the absent of ligand was normalized to 100%. ................................................................. 144

Figure 4-1. The structure of Thiazole Orange. ................................................................. 147

Figure 4-2. Structures of twelve TO derivatives and TO acid. ........................................ 149

Figure 4-3. The structure of TOTO and an illustration of bis-intercalation of TOTO. 221
Figure 4-4. CD spectra of G4 in the presence of 150 mM Na\(^+\), 100 mM K\(^+\), and in the absence of cation ........................................................................................................ 158

Figure 4-5. G-quadruplex T\(_m\) values determined by the CD thermal denaturation experiments in the presence of TO derivatives ............................................................................................................. 159

Figure 4-6. Fluorescence scan of TO derivatives in the presence of human telomeric G-quadruplex. C\(_{G\text{-quadruplex}}\)=0.25 µM; C\(_{TO}\)=0.5 µM ........................................................................................................ 163

Figure 4-7. Maximum fluorescence intensities of twelve TO derivatives in the presence of G-quadruplexes ........................................................................................................ 164

Figure 4-8. Fluorometric titration curves of TO and TO derivatives against human telomeric G-quadruplex. ----two binding site mode; —one binding site mode ........... 168

Figure 4-9. K\(_d\) values of TO and derivatives binding to G4 (solid) and ds17 (blank) determined by the fluorometric titration experiments ..................................................... 169

Figure 4-10. Fluorometric titration curves of TO and TO derivatives against ds17 ----two binding site mode; —one binding site mode ..................................................... 170

Figure 4-11. Negative ion ESI mass spectra of solutions containing CP1-3 and G4. *: Free G4, α: 1-1 complexes, β: 1-2 complexes γ: 1-3 complexes ........................................... 173

Figure 4-12. Negative ion ESI mass spectra of solutions containing CP4-6 and G4. *: Free G4, α: 1-1 complexes, β: 1-2 complexes γ: 1-3 complexes ........................................... 175

Figure 4-13. Negative ion ESI mass spectra of solutions containing CP7-8 or TO-acid and G4. *: Free G4, α: 1-1 complexes, β: 1-2 complexes, γ: 1-3 complexes ............... 177

Figure 4-14. Negative ion ESI mass spectra of solutions containing CP9-12 and PCG1. ................................................................................................................................. 178

Figure 4-15. IR values for TO derivatives binding with G4 ........................................... 180

Figure 4-17. Comparison of IR$_{G4}$ and IR$_{ds}$ values of TO and derivatives.

Figure 4-18. The proposed fragmentation pathway of G4-CP9 complexes in MS/MS.

Figure 4-19. Energy dependent product ion spectra of G4-CP9 1-3 complexes (γ2N).

From top to bottom: a) ESI-MS spectrum of a solution containing CP9 and G4; MS/MS product ion spectra of G4-CP9 [1-3] complexes at the collision energy of b) 5 au and c) 10 au. α: 1-1 complexes, β: 1-2 complexes, γ: 1-3 complexes; nN, n= the number NH$_4^+$ ions.

Figure 4-20. Energy dependent product ion spectra of G4-CP9 1-2 complexes (β2N).

From top to bottom: a) ESI-MS spectrum of a solution containing CP9 and G4; MS/MS product ion spectra of G4-CP9 [1-2] complexes at the collision energy of b) 5 au and c) 10 au. α: 1-1 complexes, β: 1-2 complexes, γ: 1-3 complexes; nN, n=the number NH$_4^+$ ions.

Figure 4-21. Energy dependent product ion spectra of G4-CP9 1-1 complexes (α2N).

From top to bottom: a) ESI-MS spectrum of a solution containing CP9 and PCG1; MS/MS product ion spectra of G4-CP9 [1-1] complexes at the collision energy of b) 5 au and c) 10 au. α: 1-1 complexes, β: 1-2 complexes, γ: 1-3 complexes; nN, n=the number NH$_4^+$ ions.

Figure 4-22. Energy dependent product ion spectra of G5-CP9 complexes. From top to bottom: a) ESI-MS spectrum of a solution containing CP9 and G5; MS/MS product ion spectra of G5-CP9 [1-3] complexes at the collision energy of b) 2 au and c) 6 au; d)
MS/MS product ion spectra of G5-CP9 [1-2] complexes at the collision energy of 10 au.

α: 1-1 complexes, β: 1-2 complexes, γ: 1-3 complexes; .......................................................... 193

Figure 4-23. $^{32}$P-TRAP reactions in the presence of TO at concentrations of 50, 25, 12.5, 6.25, 3.125, 1.56, 0.78 and 0.39 µM (from left to right). Reactions in the presence BIBR (50 µM) and TMPyP4 (50 µM) were carried out as controls. IS: Internal Standard (TSNT as template). .......................................................... 196

Figure 4-24. $^{32}$P TRAP reactions in the presence of TO derivatives at a concentration of 6.25 µM. Reactions in the presence BIBR (50 µM), TMPyP4 (50 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template). ......................... 198

Figure 4-25. $^{32}$P TRAP reactions in the presence of CP4 (left) and CP5 (right) at concentrations of 10, 5, 2.5, 1.25, and 0.625 µM (from left to right). Reactions in the presence BIBR (50 µM) and TMPyP4 (50 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).......................................................... 199

Figure 4-26. $^{32}$P TRAP reactions in the presence of CP6 at concentrations of 10, 5, 2.5, 1.25, 0.625, 0.313, 0.156, and 0.078 µM (from left to right). Reactions in the presence BIBR (50 µM), TMPyP4 (50 µM) and spermine (100 and 10 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template). ............................. 200

Figure 4-27. $^{32}$P TRAP reactions in the presence of BRACO-19 (left) and CP6 (right) at concentrations of 10, 2.5, 0.625, 0.156 and 0.039 µM (from left to right). Reactions in the presence BIBR (50 µM) and TMPyP4 (50 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).......................................................... 201

Figure 4-28. $^{32}$P TRAP reactions in the presence of CP9, CP10, CP11 and CP12 (from left to right) at concentrations of 10, 5, 2.5, 1.25, and 0.625 µM (from left to right). Reactions
in the presence PIPER (10 μM), BIBR (50 μM) and TMPyP4 (50 μM) were carried out as controls. IS: PCR product as internal standard (TSNT as template). ........................................ 202

Figure 4-29. Cell proliferation of two cancer cell lines treated by TO derivatives at 100 μM. The cell proliferation in the absence of a ligand was normalized to 100%. .......... 204

Figure A-1. MS spectrum of Phen12-Ni_50% MeOH................................................................. 223

Figure A-2. MS spectrum of Phen15-Ni_50% MeOH................................................................. 224

Figure A-3. MS spectrum of Phen18-Ni_50% MeOH................................................................. 225

Figure A-4. MS spectrum of Phen20-Ni_50% MeOH................................................................. 226

Figure A-5. MS spectrum of Phen21-Ni_50% MeOH................................................................. 227

Figure A-6. FID G4 titration course of 5-substituted Phen Ni(II) complexes. ............. 234

Figure A-7. d5DC50 plots for Phen-Ni..................................................................................... 237

Figure A-8. G4DC50 plots for 13, 20 and 21 in the absence of Ni. ................................. 238

Figure A-9. Concentration study of effective 5-substituted Phen-Ni complexes on tumor cells. ................................................................. 242

Figure A-10. CD thermal denaturation smoothed profiles of TO derivatives. .......... 244

Figure A-11. Fluorometric titration of TO derivatives against a duplex 17mer......... 251
LIST OF ABBREVIATIONS

A  Adenine
ALT Alternative lengthening telomere mechanism
C  Cytosine
CD  Circular dichroism
DNA  Deoxyribonucleic acid
ESI  Electrospray ionization
Hp  3-Hydroxypyrrole
FRET  Fluorescence resonance energy transfer
FID  Fluorescence DNA probe displacement assay
G  Guanine
G4  G-quadruplex
GMP  Guanosine Monophosphate
I  Inosine
Im  Imidazole
IS  Internal standard
MTT  3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NMR  Nuclear magnetic resonance
PCR  Polymerase chain reaction
Py  Pyrrole
Phen  Phenanthroline
RNA  Ribonucleic acid
T  Thymine
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFO</td>
<td>Triplex forming oligonucleotide</td>
</tr>
<tr>
<td>TRAP</td>
<td>Telomeric repeat amplification protocol</td>
</tr>
<tr>
<td>TO</td>
<td>Thiazole orange</td>
</tr>
<tr>
<td>T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>U</td>
<td>Uracil</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
</tbody>
</table>
Chapter 1. Introduction to DNA

1.1. DNA Basics

1.1.1. Central dogma. DNA occupies an essential position in the biological system, which is described by Francis Crick using the central dogma of molecular biology (Figure 1-1. solid and dotted arrows).\(^1\) This scheme implies the general flows of genetic information, which is transcribed from DNA into RNA and ultimately translated into protein.

![Central dogma of molecular biology](image)

Figure 1-1. The Central dogma of molecular biology.\(^2\)
Genetic characteristics are stored in DNA as linear codes composed of nucleotide monomers in the chromosome. During self-replication process, DNA polymerase is employed to synthesize daughter sequences by adding nucleotides that are complementary to the DNA template. This process yields identical copies of the DNA which are passed down into descendent cells after cell division. High fidelity of DNA self-replication is the basis of biological inheritance. Transcription and translation are spatially and temporally separated in eukaryotic cells. Transcription occurs in the nucleus, and one of the DNA strands (non-coding strand with anticodon) is used as a template to produce mRNA by RNA polymerase. The pre-mRNA is processed to produce the mature mRNA, which is transported into the cytoplasm. Because of the absence of the nucleus, prokaryotic cells immediately translate their transcripts while the mRNA is still being synthesized. In the translation process, mRNA is decoded in ribosome by inducing the binding of complementary tRNA anticodon sequences to mRNA codons. The tRNA carries specific amino acids that are coupled together into a polypeptide. The polypeptide later folds into an active protein and performs its functions.

Other than above general genetic information flows (solid arrows in Figure 1-1), Crick also included several special routes (dotted arrows in Figure 1-1) in the original central dogma. The discovery of reverse transcriptase in some retroviruses proved that RNA could make modifications to DNA.³ Telomerase is another case that RNA is used as a template to synthesize DNA which plays an important role in cell aging and cancer development.⁴ Most viruses have RNA genomes that are replicated and transcribed by RNA polymerases which provides another special case in which RNA could be synthesized from an RNA template.⁵ The direct synthesis of protein from DNA templates
has not been discovered so far in vivo, but has been observed under certain suitable conditions in vitro.⁶

After decades of studies, an additional genetic flow is characterized which was unknown when the original central dogma was conjured. The recent studies of the behaviors of prions illustrated an information feedback flow from protein to the genome (dashed arrow in Figure 1-1). As a part of the epigenetic process, environmental stress resistant phenotype was coded into DNA to become genotype through protein conformational changes. This observation has been found to be intrinsically stochastic, nevertheless, important in evolution because of the look-ahead effect.⁷-⁸

Above all, the seven general and specific genetic information flows constitute the comprehensive central dogma which captures the upstream role of DNA in molecular biology.

1.1.2. The history of DNA discovery. DNA was first discovered by Friedrich Miescher in 1868, who was dedicated to identifying the fundamental components of cells. In his experiments, Miescher isolated an unknown substance from leukocytes and sperm which was distinct from protein and lipid: It was resistant to protease, containing phosphorous but not sulphur, and it could be stained by alkali dyes. This molecule was named as nuclein because of its abundance in nuclei. It was also presumed by Miescher that nuclein might have a role in the passing of hereditary traits because it was rich in sperm cells. However, what could not be rationalized, at that time, is that how the same substance could result in the diversity of different animal species.⁹

In the 1880s, Albrecht Kossel (awarded the Nobel Prize in 1910) and Richard
Altmann developed new protocols that permitted them to isolate pure nucleic acids and enabled them to analyze its detailed chemical compositions. Kossel identified the fundamental building blocks of nuclein—the purine and pyrimidine bases, sugar, and phosphoric acid in 1901 and altered the designation "nuclein" to "nucleic acid". DNA and RNA were named as "thymus nucleic acid" and "yeast nucleic acid" after their isolation sources.

Protein but not nucleic acid was considered as the chemical basis for heritable transformations until 1943 proved by Oswald T. Avery. In his exquisite experiments, it was found that inert Streptococcus pneumoniae can only be transformed into lethal strains in the presence of DNA but not protein or carbohydrate. Avery stated that "Deoxyribonucleic acid plays a central role in determining specific characteristics in the course of reproduction". This discovery excited further research into the nature of DNA, including X-ray diffraction studies of its structure.

Almost a decade passed before, in 1952, Rosalind Franklin photographed crystallized DNA fibers. She calculated the dimensions of the strands and also deduced that the phosphates were on the outside of what was probably a seemly helical structure. In 1953, using the available X-ray data and model building, Francis Crick and James Watson (awarded the Nobel Prize in 1962) deciphered the double helical structure of DNA. They illustrated that DNA was comprised of paired sequences of complementary bases and, by the order of its bases, encodes the genes.

In the following years, the machinery of DNA self-replication, tRNA and mRNA were discovered. In 1964, Marshall Nirenberg (awarded the Nobel Prize in 1968) first
deciphered the bases in each codon which are used in protein synthesis. In 1977, Frederick Sanger (awarded the Nobel Prize in 1958 and 1980) developed rapid DNA sequencing techniques which dramatically advanced the gene mapping. In the next decades, gene started to be isolated and specific genes were found to be associated with diseases including Huntington's disease, breast and ovarian cancer.

In 1988, the National Research Council announced a program to map the human genome. The Human Genome Project was officially started in 1990 and completed in 2003. Availability of genetic information in an enlarged scale makes understanding of genetic basis of complex human traits become practical.

Throughout today, the understanding of DNA has been continuously developed. There have also been breakthroughs in the molecular biology and medical fields in which not only DNA, but also DNA-associated processes have become the targets for therapeutic applications.

1.2. DNA Structures and Functions

1.2.1. DNA structure basics. The functions of a molecule in the biological process are driven by its structures including covalent bonding and three-dimensional conformations. DNA is a polymer composed of four types of nucleotides linked by phosphodiester bonds in a specific order. DNA primary structure commonly refers to the linear nucleotides sequence, which is the most basic covalent construction of the molecule.

The monomer of a nucleic acid sequence, which is a nucleotide, consists of three components: nucleobase, sugar, and phosphate. Based on their chemical structures,
nucleobases are classified into purines (Adenine and Guanine) and pyrimidines (Cytosine, Thymine, and Uracil). A, G, C, and T are the four nucleobases that form a conventional DNA sequence and U is only present in RNA (Figure 1-2). It is logically hypothesized that DNA is originated from RNA in the course of evolution, in which, U is methylated to T. One of the main reasons for this modification is to identify and repair the deamination of cytosine into uracil in DNA, which is a common spontaneous reaction. By utilizing T instead of U, in DNA, these deaminated C (U) can be differentiated from native T and repaired by corresponding enzymes to avoid any further lethal mutations.  

In addition to these conventional nucleobases, there are some minor modified nucleobases, such as 7-methylguanine, Inosine, and 5-methylcytosine playing important roles in epigenetic genomes (Figure 1-2).

Figure 1-2. The structures of nucleobases present in human genome.

The nucleobase is connected to a pentose sugar which brings another significantly
different feature between DNA and RNA: \( \beta \)-D-ribofuranose for RNA and \( \beta \)-D-2'-deoxyribofuranose for DNA. The 2'-OH in an RNA ribose can readily attack the phosphodiester bond and lead to strand breaks. Hence, ribose is replaced by deoxyribose in the process of evolution, which makes DNA more stable than RNA. The nucleobase is anchored to the C1' position of the sugar by a \( \beta \)-N-glycosidic bond at the N9 of purine and N1 of pyrimidine. The torsional angle of the C1'-N-glycosidic bond can adopt two possible orientations (anti or syn), which are shown in Figure 1-3.

![Figure 1-3. The two conformations of N-glycosidic bonds.](image)

The sugar ring is not planar, and the sugar pucker is described as the relative positions of C2' and C3' atoms with respect to the C1'-O4'-C4' plane. The sugar can adopt either C2'-endo or C3'-endo conformation (Figure 1-4). The conformational features of base torsional angle and sugar pucker significantly contribute to the stability of different DNA secondary and tertiary structures.
A phosphate group forms two ester bonds to connect the 5'-OH and 3'-OH of the adjacent sugars. The DNA backbone is constituted by these repetitive linking units of sugar-phosphate-sugar (Figure 1-5). The terminals of 5'-OH and 3'-OH endow a polynucleotide directionality.
The linear sequence of DNA is the foundation of genetic information coding. This unique design of DNA primary structure makes compact information storage practical in the biological system. In order to carry out its other functions, including reading and replicating genetic information, DNA employs base pairings to form a duplex.

1.2.2. DNA duplex. DNA adopts double strands (duplex) under most circumstances in living cells, with two strands progressing in opposite directions. When replication or transcription occurs, two strands are separated in certain regions, and the resulted single strands have to be stabilized by particular proteins.

In 1953, James D. Watson and Francis H.C. Crick proposed a three-dimensional, right-handed double helical B-DNA model based on their analysis of a DNA fiber diffraction data. In this famous model, the bases from two single strands form specific base pairs through hydrogen bonds and form rungs of a DNA ladder. With this base-pairing rule, one DNA strand is able to define the complementary one which enables it to replicate and transcribe. Figure 1-6 shows a detailed view of the Watson-Crick base pair. Hydrogen bonds form between the aligned donors (oxygen or nitrogen bonded to hydrogen) and the acceptors (oxygen or nitrogen). The G-C and A-T base paring system is used because it maximizes the number of hydrogen bonds and minimizes the torsional strains in the backbone.

Since the phosphate groups are densely negatively-charged, the DNA backbone is very polar. To stabilize the structure, DNA adopts a conformation in which the polar phosphate backbones helically coil and shield the hydrophobic nucleobases inside. Thus, the charge distribution can be optimized to reduce the repulsion. Meanwhile, the bases
can be well stacked on each other, so the aromatic stacking force will greatly contribute to the double helix stabilization. The negative charges on the phosphate backbones are very important for DNA stabilities, protecting the ester linkage from being attacked by exogenous and endogenous nucleophiles.

Figure 1-6. Watson-Crick base pairing.

When the two strands are coiled into a double helix, the phosphates form two
ridges, and the two edges of the base pairs form two grooves back to back (Figure 1-7). The edge composed of the C1' from two sugars, the N3 of purine, and the C2 of pyrimidine is located in the minor groove, while the other edge is located in the major groove (Figure 1-6). The features of the grooves vary in depth and width depending on the types of the helix.

Figure 1-7 DNA double helix and two grooves.

More than one type of helix can be formed by DNA double strands (Figure 1-8), and their features are summarized in Table 1-1. B-form DNA helix is the predominant conformation under physiological conditions and was firstly characterized by Watson and Crick. B-form DNA adopts a right-handed (twisting direction follows the right-hand rule) helix with the C'2-endo sugar pucker and anti-N-glycosidic bonds. The minor groove is narrow in B-form DNA while the major groove wide, with nearly the same depth (Figure 1-8). These two grooves create distinguishable microenvironments for protein and small molecule recognition.
Under different environmental conditions, such as ionic strength, temperature, humidity, and nucleotide composition, DNA can adopt several distinct helical conformations. Two of the most well studied: A- and Z-form DNA helix, will be discussed here.

With relatively low hydration and low ionic strength, B-form DNA will transform into A-form. RNA-DNA hybrid and RNA double strands can also adopt A-form helical conformation. A-form DNA has the C3'-endo sugar pucker and tilted base-pairing which make it distinct from B-form. From the top view of an A-form helix, a large cavity can be observed located at the center because the coils are displaced away from the axis. As a result of the above features, the major groove in A-form DNA is narrow, and the minor groove is shallow and wide (Figure 1-8).

![Figure 1-8. Three DNA helices.](image)
Z-form DNA is usually observed in continuous pyrimidine-purine sequences in highly concentrated NaCl solutions \textit{in vitro} ((CG)n in 3-4 M NaCl). It is utilized as a rotor in DNA origami because it adopts a left-handed helix.\cite{29} Z-form DNA has more base pairs in every helical turn and is longer than the other conformations. The major groove in Z-form DNA is wide and shallow, and the minor groove is deep. Because of the torsional conformation, the sugar pucker and N-glycosidic bond conformation in a Z-form DNA keep changing alternatively along the sequence.
Table 1-1. Summary and comparison of double-helical DNA structure parameters.

<table>
<thead>
<tr>
<th>Properties</th>
<th>A-DNA</th>
<th>B-DNA</th>
<th>Z-DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix sense</td>
<td>Right-handed</td>
<td>Right-handed</td>
<td>Left-handed</td>
</tr>
<tr>
<td>Helix diameter</td>
<td>26Å</td>
<td>20Å</td>
<td>18 Å</td>
</tr>
<tr>
<td>Repeating unit</td>
<td>1bp</td>
<td>1bp</td>
<td>2bp</td>
</tr>
<tr>
<td>Base pair/turn</td>
<td>10.7</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Degrees of inclination of base pair from normal to the helical axis</td>
<td>19</td>
<td>-1</td>
<td>-9</td>
</tr>
<tr>
<td>Rise per base pair</td>
<td>2.3 Å</td>
<td>3.3-3.4 Å</td>
<td>3.8 Å</td>
</tr>
<tr>
<td>Pitch per helical turn</td>
<td>24.6</td>
<td>33.2</td>
<td>45.6</td>
</tr>
<tr>
<td>Glycosyl ring conformation</td>
<td>C3'-endo</td>
<td>O1'-endo to C2'-'endo</td>
<td>G: C2'-endo C1'-exo C3'-endo</td>
</tr>
<tr>
<td>Glycosyl angle conformation</td>
<td>Anti</td>
<td>Anti</td>
<td>Anti at cytosine, Syn at guanine</td>
</tr>
<tr>
<td>Major groove</td>
<td>Narrow and deep</td>
<td>Wide and deep</td>
<td>Flattened</td>
</tr>
<tr>
<td>Minor groove</td>
<td>Wide and shallow</td>
<td>Narrow and deep</td>
<td>Narrow and deep</td>
</tr>
</tbody>
</table>

1.2.3. DNA triplex. The conventional DNA helix is based upon Watson-Crick base pairs (G-C and A-T), which are the most thermodynamically stable base-pairing. However, several alternative base pairs have been discovered with important roles in DNA higher order structures and functions. One of them was Hoogsteen base-pairing, named after the scientist who first characterized it. In Hoogsteen base pairs, the N7 of purines can accept hydrogens and form additional hydrogen bonds. The N3 of cytosine
can also be protonated under acidic conditions to become a hydrogen bond donor (Figure 1-9). The formation of Hoogsteen base pairs constructs several alternative DNA structures including triplexes and G-quadruplexes.

**Figure 1-9. Hoogsteen base pairs in DNA triplexes.**

The three-stranded polynucleotide was first observed by using X-ray diffraction in 1957 by Felsenfeld, in which, a synthetic third strand poly(U) binds into the major groove of a poly(U)/poly(A) duplex.\(^{32}\) Two common triplets are A-T(U)-T(U) or G-C-C\(^+\) and C needs to be protonated in the latter under acidic conditions. In general, a triplex is formed
when a pyrimidine base binds onto the Hoogsteen position of an existing Watson-Crick base pair (Figure 1-10).

![Figure 1-10. DNA triplex.](image)

The Triplex Forming Oligonucleotide (TFO) sequence specifically occupies the major groove of a DNA duplex, which is also the binding site of most proteins and small molecules; therefore, the binding event could significantly affect the DNA replication and transcription process. Because of this, the formation of DNA triple has attracted many interests for potential applications in site-specific gene therapy. Since the formation of a triplet G-C-C\(^+\) requires an unusual acidic condition, the A-T(U)-T(U) motif becomes the more suitable choice for genome therapy applications.

TFOs have been successfully used for site-directed mutagenesis.\(^{33-34}\) However,
there are many difficulties in the triplex-forming strategy in clinical applications. For instance, sufficiently long polypurine sequences, which are critical for the binding of TFO, is rarely present in human genome. In addition, delivery of exogenous polynucleotides into cells is always problematic because of the cellular uptake and biological stability.\(^\text{35}\)

**1.2.4. DNA G-quadruplex. The basics of G-quadruplex structures.** In addition to triplex DNA, DNA G-quadruplex (G4) is another alternative structure that can be formed through Hoogsteen hydrogen bonds. The self-assembly of guanines was first proposed by Gellert \textit{et al.} in 1962 using X-ray diffraction from a concentrated GMP (guanosine monophosphate) solution, and guanine-rich sequences later was found in telomeric DNA and many gene promoter regions.\(^\text{36}\) Figure 1-11 shows the top view of a G-quartet, in which, four guanines self-assemble via Hoogsteen hydrogen bonds (N1-O6 and N2-N7). The planar G-quartets stack on top of each other with a torsional angle, forming a four-stranded helical structure. Naturally, G-quadruplex was proved to be stabilized by monovalent cations such as K\(^+\) and Na\(^+\) sitting at the center of the G-quartets. The important factors that contribute to the G-quadruplex formation are discussed here.
**Figure 1-11.** The top view of a DNA G-quartet (left) and a sketch of a G-quadruplex (right).

**Hydrogen bonding of G-quadruplex.** There are six electronegative atoms present in a guanine, and four of them are open to hydrogen bond formation except for the N9 (used in the N-glycosidic bond) and N3 (too close to the backbone). In canonical Watson-Crick base pairs, the O6 (carbonyl oxygen) accepts hydrogen from the N4 of cytosine, while the N1 and N2 donate hydrogens to the N3 and O6 of cytosine. Noticeably, guanine is also able to accept hydrogen bonding with its N7 on the other side. When the hydrogens from the N1 and N2 (Watson-Crick side) are accepted by the O6 and N7 (Hoogsteen side) of an adjacent guanine, Hoogsteen hydrogen bonds can form. The two sides of a guanine are shown in Figure 1-12. With four guanines, a cyclic planar structure can be constructed (Figure 1-11). It should be emphasized that the hydrogen bonds between guanines are not the most crucial factors in G-quadruplex stability, because the strand distortion greatly reduces the hydrogen bonding strength. G-quadruplex can be
formed even with less hydrogen bonds when guanine is substituted by other derivatives, while cation residing in the center channel is inevitable.\textsuperscript{39}

![Figure 1-12. Hoogsteen side and Watson-Crick side for guanine hydrogen bonding.](image)

\textit{Metal cation and guanine charge-charge interactions.} In the absence of cations, G-quadruplex does not form unless specific stabilizing ligands are present. The central cavity of the G-quartet produces a space that a cation with a suitable size can fit in. The lone pair electrons of four O6 are coordinated by the cation to reduce the charge repulsion. Na\textsuperscript{+} with an ionic radius of 0.95 Å is reported as the maximum size of a cation that can fit into the G-quartets. Larger cation such as K\textsuperscript{+}, with a radius 1.33 Å, therefore, is observed sitting between the two quartets and coordinated by eight oxygens (Figure 1-13).\textsuperscript{38} Other cations such as NH\textsubscript{4}\textsuperscript{+}, Rb\textsuperscript{+}, Ti\textsuperscript{+} and even some divalent cations such as Sr\textsuperscript{2+}, Ba\textsuperscript{2+}, and Pb\textsuperscript{2+} were also reported stabilizing G-quadruplex to some extent.\textsuperscript{40-44}

Cations also affect the G-quadruplex polymorphism. It was reported that G-rich sequences adopted different conformations in the presence of Na\textsuperscript{+}, K\textsuperscript{+} and other cations.
For instance, a telomeric repeat sequence (GGGTAA)₃GGG forms a parallel G-quadruplex in K⁺ solution and mixed parallel-antiparallel conformation in the presence of Na⁺.⁴⁵-⁴⁷ Because of this feature, G-quadruplex is also developed as a highly sensitive tool to detect alkali metal ions.⁴⁰,⁴⁸

![Figure 1-13. G-quadruplexes with coordinated cations. (A) A parallel G-quadruplex with Na⁺ ions, (B) A parallel G-quadruplex with K⁺ ions.](image)

**Base stacking force of G-quadruplex.** Like other canonical nucleic acid structures, the nucleobase π–π stacking is crucial to the formation of G-quadruplex. Individual G-quadruplex can stack on the others even without covalent bond linkage.⁴⁶,⁴⁹-⁵⁰ Three common stacking geometries found in the G-quadruplex unit are shown in Figure 1-14.⁵¹ The orientation of G-quadruplex forming strands, and the glycosidic bond angle are the major factors affecting the stacking mode.

Besides the above factors, G-quadruplex is also stabilized by super helical stress, molecular crowding, as well as particular G-quadruplex binding proteins and small molecules,⁵² which will be discussed in Chapter 2.
Figure 1-14. Illustrative examples of base stacking mode in G-quadruplexes.\textsuperscript{51}

**Topology of G-quadruplex.** G-quadruplex can be formed through one, two or four G-rich strands. Each G-rich strand has specific direction defined by sugar-phosphate backbone from 5'-3'. The segment between G-tracts will form loops hanging outside the unit. Therefore, G-quadruplex can adopt various topologies, as the consequence of a wide variety of possible combinations of strand directions and loop sequences. The general structure can be described as a central unit of at least two stacked G-quartets, in which the G-rich sections are connected by loops of intervening nucleotides that are not involved in the quartets.

According to their strand directions, the polymorphic G-quadruplexes are classified into parallel and antiparallel structures.\textsuperscript{53} When four G-rich strands are all in the same direction in a G-quadruplex unit, it is termed as a parallel G-quadruplex (Figure 1-15A). Obviously, there is no need for a loop if it is formed by four disconnected sequences. If at least two of these adjacent G-strands are from the same sequence in parallel G-quadruplex, there will be loops left outside the unit connecting the top and bottom quartets. This type of connecting sequence (Figure 1-15B) is referred as a
propeller loop which is found in both telomeric and non-telomeric G-quadruplex.\textsuperscript{50, 54} Guanine glycosidic bonds in parallel quadruplexes all adopt the anti-conformation (Figure 1-16A).

Figure 1-15. Various topologies for G-quadruplexes.

When at least one of the four strands adopt the different direction from the others, an antiparallel or mix-antiparallel G-quadruplex is formed (Figure 1-15C). Two distinct types of loops, in addition to the propeller loop, are discovered in antiparallel G-quadruplexes. When two antiparallel strands from the same sequence are adjacent to each other, a lateral loop connects them (Figure 1-15D). When the antiparallel sequences are located in diagonally, a diagonal loop connects them crossing of the G-tetrads (Figure 1-15E).\textsuperscript{50, 55-56} Antiparallel G-quadruplexes were found in both telomeric and non-
telomeric G-quadruplexes.\textsuperscript{57-59} Antiparallel G-quadruplexes require both anti- and syn-guanine glycosidic bonds to maintain the particular conformation (Figure 1-16B).

Figure 1-16. Glycosidic bonds conformations in various G-quadruplex topologies. Syn-guanines are shown in yellow and anti-guanines are in grey.

Four grooves are formed in a G-quadruplex structure defined by the phosphodiester backbones and the features are depending on the particular topology. A G-quadruplex with only lateral and diagonal loops has relatively simple groove features, while propeller loops in a parallel G-quadruplex may insert into the groove generating more complex features.

As mentioned in the previous section, alkali cations stabilize G-quadruplex in the center channel of the G-quartets and consequently neutralize the negative charges from the O6 of guanine. K\textsuperscript{+} and Na\textsuperscript{+} induce different topologies in some of G-quadruplexes, but not necessary for all of them.\textsuperscript{47, 60}
**Biological relevance of G-quadruplex.** By using sequencing methods, G-quadruplex was found extensively present in the genome frequently coupled with important regulating functions such as telomere length maintenance and oncogene expression.

More than 300,000 potential G-quadruplex forming sequences were discovered across the genome, by searching using a universal G-quadruplex forming motif represented by \( d(G_3+N_1\cdots G_3+N_1\cdots G_3+N_1\cdots G_3) \).\(^1\) Many of them are highly conserved in both prokaryotic and eukaryotic organisms. These conserved regions are mostly present in mammalian species, followed by non-mammalian, and lower organisms. This observation indicates that the natural selection pressure tends to maintain G-quadruplex functions in the evolution process.\(^2\)

Most of the potential G-quadruplex formation sequences were found in telomeric regions where up to 15,000 bps of repeated TTAGGG units were discovered.\(^3\) The G-quadruplex formation and stabilization in telomeric regions are crucial in regulation of telomere length maintenance and chromosome end protection.

G-quadruplexes are also located in many oncogene promoter regions, such as MYC, JUNB, FGF4, and TERT, and regulate their transcriptions. Figure 1-17 shows a scheme of this regulation process. In a normal transcription process, an RNA polymerase binds onto the non-coding strand (with anticodons) of the DNA, and the generated RNA will have the same sequence as the coding strand (with codons). When G-quadruplex is present in the non-coding strand (S region), polymerase binding and the downstream RNA transcription are stopped at this position. Meanwhile, an unusual G-loop will be
induced on the complementary coding strand. RNA can be generated from this loop and form a stable co-transcription RNA/DNA hybrid. Consequentially, a strand exchange (recombination) can be introduced, which significantly changes the genetic information.\(^{63-65}\)

Because of its crucial functions in the biological process, a range of G4-targeting ligands have been synthesized in order to regulate the telomere stability and oncogene expression. It was reported that the expression of c-MYC in cancer cell lines was down-regulated to 20% with a 48h G4 ligand TMPyP4 treatment \textit{in vivo}.\(^6^6\) (Design of G4 ligands will be extensively discussed in Chapter 2.)

![Diagram of replication and transcription](image)

Figure 1-17. Structures formed upon replication and transcription of regions containing G-quadruplex motifs.

It has been noticed that G-quadruplex can be both stabilized and destabilized in order to regulate the related biological process. During the replication and transcription, DNA single strands are exposed and may spontaneously form unnecessary G-
quadruplexes. In most cases, G-quadruplex helicases (including BLM, WRN, FANCJ, and others) are employed to recognize and unwind these structures to maintain genetic stability.\textsuperscript{67-69} Deficiency of these enzymes leads to certain genetic diseases.\textsuperscript{70}

\textbf{Telomere and the DNA end replication problem.} Telomeres have the highest G-quadruplex formation sequence abundance not only in the human genome but also other organisms.\textsuperscript{63} They are located at the end of chromosomes and protect them against genome instability.\textsuperscript{71} The functions and features of telomeres will be discussed in the following topics:

\textbf{Structure of telomeres.} Human telomeres are composed of 9-15 kb double-stranded TTAGGG repeats (different from TTGGGG in lower organisms) ending in a 50-300 nt long 3′ single-stranded repeats, termed as G-overhang, whose complementary C-rich sequence is digested by a 5′-3′ exonuclease.\textsuperscript{72} By folding back to the double strand region, the G-overhang at the end of a telomere can form a "T-loop" (Figure 1-18).
Functions of telomeres maintaining chromosome stability. Exposed DNA strand ends (both blunt and sticky ends) are very lethal to cells and are considered as damaged sites because they tend to fuse randomly. Therefore, a range of enzymes that detect, remove and repair the damage sites are utilized to protect the DNA. The ends of chromosomes, which are natural strand terminals, have to be protected from both the end fusion and being recognized by damage seeking enzymes. Therefore, telomeres are employed to form stable higher order structures and cap the end of chromosomes.

As discussed before, the repetitive G-rich sequence in telomeres can form G-quadruplex which will bring challenges for DNA recognition by enzymes and self-replication. To stabilize these higher order structures, protect the telomeres from enzyme attacks, and also enable its replications, a complex containing six individual proteins is evolved (Figure 1-18). In this telomeric sheltering complex, TRF 1 and 2 are bound to double strands and POT 1 to G-overhangs which are bridged by TPP1 and TIN2. These
binding motifs play important roles in the telomere replication regulation, chromosome end protection, and telomere length maintaining.\textsuperscript{73-75}

Telomeres are crucial to the chromosome protection; however, they are truncated after each single replication because of the end replication problem:

\textit{End replication problem.} In 1960s, Leonard Hayflick discovered that normal human cells have limited capacity of cell division which is termed as 'Hayflick limit'. A decade later, Alexey M. Olovnikov firstly proposed the DNA end replication problem and gave the hypothesis of the resolutions which was later confirmed by the discoveries of telomeres and telomerase.\textsuperscript{76-77}

End replication problem comes from the natural features of DNA self-replication machinery. With a template, DNA polymerase replicates DNA only in 5' to 3' direction, starting from a free 3'-OH of existing DNA or RNA fragments (Figure 1-19). As discussed earlier, a DNA double helix has two single strands that proceed in two directions. DNA replication is semi-conservative in which the two strands have to be separated and replicated as the templates, respectively (in a zipper-like replication fork). DNA polymerase can only synthesize daughter DNA in 5' to 3' starting from a small piece of RNA primer that attached to the parent DNA. This process starts in 3'\textsuperscript{-}5' parent strand (leading strand) and can proceed smoothly. However, on the other strand (lagging strand), which is in 5'\textsuperscript{-}3', replication has to be performed by small discontinuous pieces starting from RNA primers. All the RNA primers will be removed after synthesis and the gaps will be filled by a DNA ligase. The leading strand can be replicated to the extreme end, but not the lagging strand. The RNA primer removed from the 5' end of the lagging
strand will leave a gap without available 3'-OH group. In consequence, there is always an un-replicated piece left on the 5' end of the daughter sequence and the DNA gets shorter and shorter during replications.

Telomeres protect chromosomes by being buffer regions at the end. G-overhangs of telomeres are shortened during the replications, and the length is considered as a mark of DNA age. In somatic cells that without any telomere length maintenance pathway, a loss of 100-200 bp happens in one cell division. Therefore, these cells can only undergo certain times of telomere replications before they start to lose the protection functions.

Figure 1-19. An illustration of DNA end replication problem.
In yeast, short telomeres, derived from the end replication problem, directly lead to cell senescence. Mice with short telomeres were found with shorter life-span and fertility defects. In human, Dyskeratosis Congenita (DC) and Idiopathic Pulmonary Fibrosis (IPF) have been proved related to irregularly short telomeres. Even though short telomeres can lead to critical cell damages, human aging is still a complicated process that cannot be simply explained by telomere shortening as a single factor. Age limitation is a consequence of accumulated effects from many contributors including telomere length, oxidative stress, and DNA mutation damages.

**Telomerase.** Since telomeres are progressively truncated and losing the protection functions during cell replications, certain strategies have to be utilized to maintain the telomere length by those 'immortal' cells, such as stem cells, germline cells, and cancer cells. Telomerase, a reverse transcriptase, is found overexpressed in these cells. It maintains telomere length by adding TTAGGGG repeats onto the telomeric G-overhangs using its RNA element as the template. Telomerase was first isolated in 1984 by Elizabeth Blackburn from *Tetrahymena pyriformis* and later was reported to be found in most human cancer cell lines and nearly absent in normal somatic cells.

**Telomerase structures and functions.** Telomerase is a complex with RNA-dependent DNA polymerase activities (reverse transcriptase). The minimal requirement for telomerase activity *in vitro* is having two subunits: the RNA template (hTERC) and the catalytic enzyme (hTERT) (Figure 1-20).

TERC, the RNA component of a telomerase, is used as the template for telomere extension. Human TERC (hTERC) is produced by RNA polymerase II and consists of
451 nucleotides, in which, nucleotides 46-53 are used as the true template of DNA synthesis. The minimum segment of an RNA template for telomerase activity contains nucleotides 10-159. The other domains form higher order structures to bind onto TERT but have no contribution to DNA synthesis.\textsuperscript{85} A particular secondary structure of TERC is highly conserved across a variety of mammals (Figure 1-20).\textsuperscript{4,82}

TERT, the catalytic component of the telomerase, contains a reverse transcriptase motif located at the C-terminus that is highly conserved across species including human, mice, yeast, worm and even plant. A telomerase-specific T motif is located next to the reverse transcript region and a large N-terminal region containing other functional domains is present at the other end.\textsuperscript{86-89}

The detailed step by step process of telomerase activity is still unclear so far. Generally, telomerase functions in a complex process including binding, synthesizing and relocating. Many telomerase-associated proteins such as Hsp90, p23, and TEP1 are involved in to regulate the activity.\textsuperscript{90}
Telomerase functions and cancer. The activity of telomerase is strictly regulated in normal somatic cells and appears at an extremely low level during most of the time and only kept active in some particular tissues including germline cells, stem cells, and some lymphocytes. The presence of high-level telomerase activity is consistent with the requirement of high proliferation capacities of the listed cells. Dysfunctions of telomerase or its related proteins in these cells lead to certain human genome diseases.

Studies showed that introduction of active hTERT into normal somatic cells can help them bypass telomere shortening problem. These cells are enabled to unlimited capacity of replications to some extent, but cannot differentiate and functionalize in tissues, which is a remarkable feature of cancer cells. It is proposed that, when telomeres become too short, very rare telomerase-positive cells (1 in 10,000,000) can help themselves overcome the apoptosis machinery and become immortalized. This provides a possible explanation of cell cancerization, which is supported by the
observations that tumor cells have short telomere length compared to normal cells and show no net loss of telomere in cell divisions. \(^{97}\)

Today, nearly 90% of different types of tumors have been proved to have irregular telomerase activity compared to normal tissues. J.W. Shay and S. Bacchetti did a comprehensive survey on nearly all major types of cancer cells (Table 1-2). \(^{93}\) The striking connection between cancer and telomerase makes anti-telomerase a promising clinical therapeutic strategy.
Table 1-2. Telomerase activity detected in different type of cancer

<table>
<thead>
<tr>
<th>Type of cancer</th>
<th>Telomerase positive tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head and neck squamous cell carcinoma</td>
<td>86%</td>
</tr>
<tr>
<td>Non-small cell lung cancer</td>
<td>78%</td>
</tr>
<tr>
<td>Small cell lung carcinoma</td>
<td>100%</td>
</tr>
<tr>
<td>Gastric carcinoma</td>
<td>85%</td>
</tr>
<tr>
<td>Colorectal carcinoma</td>
<td>89%</td>
</tr>
<tr>
<td>Pancreatic carcinoma</td>
<td>95%</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>86%</td>
</tr>
<tr>
<td>Breast carcinoma</td>
<td>88%</td>
</tr>
<tr>
<td>Ovary (fetal)</td>
<td>100%</td>
</tr>
<tr>
<td>Leiomyosarcoma</td>
<td>100%</td>
</tr>
<tr>
<td>Endometrial adenocarcinoma</td>
<td>100%</td>
</tr>
<tr>
<td>Cervical carcinoma</td>
<td>100%</td>
</tr>
<tr>
<td>Vaginal</td>
<td>100%</td>
</tr>
<tr>
<td>Testic</td>
<td>100%</td>
</tr>
<tr>
<td>Prostate carcinoma</td>
<td>90%</td>
</tr>
<tr>
<td>Bladder carcinoma</td>
<td>92%</td>
</tr>
<tr>
<td>Renal carcinoma</td>
<td>83%</td>
</tr>
<tr>
<td>Wilms'tumor</td>
<td>100%</td>
</tr>
<tr>
<td>Oligodendroglioma</td>
<td>100%</td>
</tr>
<tr>
<td>Meningioma</td>
<td>100%</td>
</tr>
<tr>
<td>Cancer Type</td>
<td>Percentage</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Neuroblastoma</td>
<td>94%</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>83%</td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td>95%</td>
</tr>
<tr>
<td>Melanoma</td>
<td>86%</td>
</tr>
<tr>
<td>Myeloma</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Immune Response Vaccines.** One of the telomerase-related strategies is using telomerase derived peptides or oligonucleotides as vaccines that will stimulate the patients’ immune response to telomerase-positive cancer cells. Many approaches and clinical trials have been performed and showed positive anticancer responses. There is no toxic effect has been reported against non-cancer cells so far, but it was observed that some cancer cells can bypass telomerase signaling and build up resistance. Further modifications on the telomerase-based vaccine and more clinical trials are being conducted.98

**Direct telomerase inhibitors.** To today, millions of molecules have been designed and screened for directly telomerase targeting. However, there is no any sufficiently effective inhibitor getting into clinical trials. The most advanced inhibitor, BIBR 1532, only showed activity in certain model studies.

**Nucleotides based telomerase inhibitors.** The other strategy is DNA-telomerase interactions based, in which, oligonucleotide-based analogs are designed to block the access of telomerase to telomeres. One of the most advanced binders is Imetelstat which
is a lapidated N3'-P5' thio-phosphoramidate oligonucleotide that binds in complementary fashion to template RNA. The phase II research of Imetelstat showed very positive results on myeloproliferative disorder, and it is under phase III safety evaluations so far.\textsuperscript{99} However, Imetelstat only exhibits inhibition effects on limited types of cancer, which may be the consequence of known off-target effects of the lipid moiety.\textsuperscript{100}

\textbf{G-quadruplex telomere ligands.} As discussed before, telomeres tend to form G-quadruplexes under physical conditions and these higher order structures prevent the access of telomerase. Thus, a number of G-quadruplex ligands (G4 ligands) have been designed and synthesized. The most effective G4 ligand is the natural product telomestatin, which was proved to inhibit leukemia cells proliferation \textit{in vitro}. The main challenge in this strategy is to improve the selectivity of the ligands against G-quadruplex over duplex DNA, which will be extensively discussed in chapter 2.

Lastly, it is noteworthy that some cancer cell lines (~ 10\%) utilize alternative lengthening telomere mechanism (ALT) in the absence of telomerase.\textsuperscript{101} ALT happens when a DNA template is annealed onto the telomere and new telomeric sequences can be simply primed and replicated from this point. This template DNA can be another telomere region sequence or some linear fragments from the outside of the chromosomes.\textsuperscript{102} Even though the natural coexistence of telomerase and ALT in human cancer cells has not been found yet, it is possible in genome engineered cells \textit{in vitro}. Therefore, ALT has to be taken into considerations for telomerase-based therapies because of the potential resistance.\textsuperscript{103}
Chapter 2. Small Molecule-DNA Non-covalent Interactions

The structural diversity of DNA provides numerous distinguished environments for protein and small molecule recognition. A wide variety of proteins are employed to regulate DNA related cellular processes including DNA packaging, replication and transcription through protein-DNA interactions. Histone, for instance, is an alkaline protein that forms complexes with duplex DNA so that DNA can be packaged into chromatins (Figure 2-1).

Figure 2-1. The interaction of histone and DNA.
In another case, telomere sheltering complexes bind onto DNA telomeric regions and regulates their extensions and replications (Figure 1-18). Other than proteins, small molecules exist naturally in cells interacting with DNA. Spermine, a polyamine compound, binds and stabilizes DNA by neutralizing the negatively charged phosphate backbones. All these discoveries enlightened an active research area of DNA related therapeutics, in which, molecules are designed to bind with certain DNA structures and consequently regulate some major disease progressions like cancer.

DNA ligands interact with DNA through both covalent and non-covalent bonds. The first chemotherapy drug nitrogen mustard is a DNA covalent ligand that alkylates to the N-7 of guanine. The resulting interstrand cross-links inhibit DNA replications and transcriptions (Figure 2-2a). Cisplatin also uses the same mechanism of action by forming interstrand links between two guanines and thus inducing apoptosis (Figure 2-2b). These alkylating agents have been widely used against cancer till today and significantly increase the survival time of many patients with malignant tumors.

Figure 2-2. Mechanisms of nitrogen mustard (top) and cisplatin (bottom) DNA covalent interactions.
However, because of their non-selective binding, DNA alkylating ligands have severe side-effects including anemia, nausea, ear damage, hair loss and dysgenesis. Furthermore, more and more cancer patients have developed drug resistance after repeated treatments of these drugs. Therefore, the need for novel chemotherapeutic drugs with less side-effects never ends. In the recent years, small molecules that interact with DNA via non-covalent binding have attracted much attention. The characteristics of these molecules will be explicitly discussed in the following sections based on their binding target including duplex and G-quadruplex DNA. 1,10-phenanthroline, thiazole orange and their analogous as DNA ligands will also be discussed in the last two sections.

2.1. Duplex Binding

Duplex is the most prevalent DNA structure in living cells and is the presumed target of many small molecules in DNA recognition. Generally, small molecules bind to DNA via three interaction modes: Non-specific binding, base interaction, and groove binding. The binding mode is determined by many factors including the structures of small molecules, DNA sequences and the experimental conditions. The same ligand may have different binding modes in different DNA local environments, and sometimes multiple binding modes can co-exist.

Apparently, small molecules that bind to duplex DNA sequence specifically can effectively regulate the desired replication and transcription processes. Meanwhile, it is always very important to study the binding selectively between duplex DNA and higher order DNA structures such as G-quadruplex when designing DNA ligands. Improving the selectivity is always the biggest challenge in this field.
2.1.1. **Outside edge binding (nonspecific binding)**. The edges of a DNA duplex are composed of phosphate diester groups. Under physiological conditions, one negative charge is retained on each phosphate group and two are retained on the terminals. Through electrostatic interactions, monovalent and divalent cations including K⁺, Na⁺, or Mg²⁺, are condensed in both major and minor grooves under physical conditions. These ionic interactions were proved to be nonspecific and the cations are free to move along the sequences. Since these cations partially neutralize the negative charges of the phosphate groups and relieve the repulsion of the two sequences in a duplex, the melting temperature (Tₘ, the temperature at which 50% duplex DNA dissociates into random coils) increases with the increasing cation concentrations. Cation binding competes with DNA ligand binding; therefore, a condition with a high salt concentration inhibits the ligand binding because of the electrostatic repulsion. Meanwhile, in the development of DNA groove binders, the expelled cations from the minor grooves are often observed when ligands bind to duplex DNA, which is considered as an important contribution of the DNA ligand stabilization because of favorable entropic effects.

Other than cations, water molecules also condense on duplex DNA and form a spine of hydration in the minor groove. The extent of hydration plays a crucial role in different DNA conformations (discussed in chapter 1).

Some ubiquitous positively charged small molecules in eukaryotic cells also nonspecifically bind onto duplex DNA. Spermine and spermidine were originally isolated from sperm cells and were found to be abundant in ribosomes (Figure 2-3). They have various important functions and one of them is binding on duplex DNA through electrostatic interactions to reduce the charge repulsion when DNA is packaged into
chromatin.\textsuperscript{111} In biological studies, high concentration of spermine is often used to precipitate DNA \textit{in vitro}.\textsuperscript{112}

![Figure 2-3. Structures of spermine and spermidine.](image)

**2.1.2. DNA intercalation.** Many DNA ligands with fused planar aromatic structures intercalate into duplex DNA by aligning themselves between base pairs. The intercalation of ligands is predominately driven by the $\pi$-$\pi$ stacking force between the ligands and nucleobases. By adding various substituents into the ligand structures, hydrophobic interactions, hydrogen bonds, and electrostatic forces can be introduced to enhance the binding affinity and specify the binding orientation. Most of these duplex DNA intercalations are located on G-C base pairs because of the bigger stacking surface of purines.\textsuperscript{113} X-ray crystallography studies revealed that in the presence of a DNA intercalator, the base pair step is separated by several Å to accommodate the insertion. Meanwhile, the local DNA sequence is unwound and lengthened to release the steric hindrance.\textsuperscript{114} Duplex DNA intercalation follows the neighbor exclusion principle in which each ligand inserts into duplex DNA with 2 or 3 base pairs at a saturation level. This exclusion takes place due to the DNA local conformation changes after the first ligand binding and a decreased linear charge density.\textsuperscript{115-116}

Several DNA intercalators have been used as DNA dyes because the stacking of
their cyclic electron deficient surfaces on the base pairs emits fluorescence. Ethidium and thiazole orange are two of the most well studied and commonly used DNA dyes. An ethidium molecule preferably inserts into G-C base pairs, and the positively charged nitrogen interacts with the phosphate backbone.\textsuperscript{117} Thiazole orange shares a similar binding mechanism as ethidium with much less mutagenic effect and hence is a safer reagent in DNA research.\textsuperscript{118} Thiazole orange, as the archetypal structure of our novel G4 ligands in chapter 4, will be extensively discussed in section 2.5.

Many duplex DNA intercalators have been developed into anticancer drugs and their binding inhibit the DNA replication and interfere the biological functions of many DNA enzymes in cancer cells. Anthracycline antibiotics are an important class of DNA intercalators, which are used for the chemotherapies of leukemias, lymphomas, breast, stomach, uterine, ovarian, bladder cancers, and lung cancers. Daunorubicin is the first anthracycline antibiotic that was isolated from \textit{Streptomyces peucetius} in 1963 (Figure 2-4).\textsuperscript{119} Daunorubicin has been used for chemotherapy of acute myeloid leukemia, acute lymphocytic leukemia, chronic myelogenous leukemia, and Kaposi’s sarcoma for more than 40 years.\textsuperscript{120} Daunorubicin has an anthracene moiety which perpendicularly inserts into successive G-C base pairs. The methoxy group of the anthracene ring is located in the major groove and coordinates with sodium ions there. The amino sugar group lies in the minor groove which will expel metal cations from the DNA. The hydroxy group of the sugar forms hydrogen bond with the proximate nucleobases.\textsuperscript{121} The mechanisms of action of anthracycline antibiotics are multifactorial, one of which is inhibition of topoisomerase activity by stabilizing duplex DNA. Topoisomerases are essential enzymes that release duplex DNA supercoils in replication. Cell’s cycle was found arrested in
G1/G2 when they were treated by Daunorubicin.\textsuperscript{122} In addition, anthracycline antibiotics function to elevate ROS level in cells by participating some enzymatic pathways and induce DNA damages.\textsuperscript{123} Although being prominent anti-cancer drugs, acute and long-term cardiotoxicity has been discovered with anthracycline antibiotics in the 1970s. This side effect significantly depends on the dose, patient age, and physical conditions.\textsuperscript{124-126} The exact mechanism of the cardiotoxicity is remains unclear so far.

Transition metals including Ru\textsuperscript{2+}, Rh\textsuperscript{2+}, Fe\textsuperscript{2+}, Zn\textsuperscript{2+}, Cu\textsuperscript{2+}, and Co\textsuperscript{2+} can coordinate with ligands such as 1,10-phenanthroline and form a rigid complex that can partially insert into DNA base pairs. The antiviral and antibacterial activities of these complexes were firstly discovered in 1952.\textsuperscript{127} Later, high-resolution NMR and X-ray crystallography studies showed that they bind to duplex DNA with both intercalation and groove binding modes. There are several advantages of metallo-intercalators: 1) By simply switching the coordinating ligands, a library of transition metal complexes with various structures as DNA ligands can be conveniently prepared. 2) Because of the chirality of their structures, many complexes have exhibited prominent stereoselectivities against different DNA conformations. The enantiomers of tris(4,7-diphenyl-1,10-phenanthroline) ruthenium(II), [Ru(DIP)\textsubscript{3}]\textsuperscript{2+}, preferentially bound to left-handed Z-form DNA and right-handed B-form DNA, respectively. In general, the complexes with the bulkier ligands have better stereoselectivities because of a greater steric hindrance when intercalating with DNA.\textsuperscript{128} 3) Transition metal complexes have been used as DNA probes due to their unique photophysical properties. For instance, [Ru(phen)\textsubscript{2} (dppz)]\textsuperscript{2+} glows in organic solutions but not in aqueous solutions because the existing hydrogen bonds deactivate the excited state of the metal ion. When it binds to DNA, changing from hydrophilic to hydrophobic
environment makes the metal ion luminesce. These complexes can also be used as artificial nucleases for cleavage of DNA because of their photochemical properties. The binding sites of the complexes can be determined based on the location of DNA strand breaks activated by photolysis. To sum up, transition metal complexes can act as DNA intercalators and have been developed into stereoselective DNA binders, cleavage agents, and luminescent probes.

2.1.3. Duplex DNA groove binding. Phosphate diester backbones and the edges of base pairs defined two distinguished grooves in duplex DNA. The major groove of B-DNA is wide and deep and provides a suitable environment for most of the protein

![Figure 2-4. Structures of DNA intercalative ligands.](image-url)
binding. By contrast, most of the small molecules fit into the minor groove because of its deep and shallow geometry, with a few exceptions of bulky transition metal complexes.

Duplex DNA minor groove binders such as netropsin and distamycin (Figure 2-5) share common structural features of multiple small aromatic rings (pyrrole, furan, benzene, imidazole etc.) linked by torsional bonds. Because of the flexible linkage, these ligands can twist and form a crescent shape that snugly fits into the minor groove. These ligands contain positive charges at the end, which further enhance the binding through electrostatic interactions. Different from the G-C preferred intercalation, most of the minor groove ligands locate themselves into A-T tracts. The O2 of thymine and N3 of adenine in A-T tracts can form H-bonding with the ligand. The G-C tracts are sterically hindered for ligand binding. The amino group at C2 of guanine, points to the minor groove, which causes steric hindrance for ligand interactions. Generally, groove binding ligands have been found with a higher sequence specificity compared to the intercalators.

First DNA groove binder netropsin was isolated from *Streptomyces netropsis* in 1951 by Finlay and the coworkers (Figure 2-5). This pyrrole-amidine structure showed antibacterial and antiviral activities. High-resolution X-ray crystallography revealed that the pyrrole rings of netropsin are located parallelly to the walls of the minor groove and displace several H2O molecules from the groove. The binding site of this antibiotic was identified as AATT or ATAT. Another polyamide-antibiotic, distamycin, was also harvested from *Streptomyces netropsis*. Interestingly, distamycin binds strongly to A-T base pair minor groove as a dimer. DNA minor groove binders that have elevated binding to G-C tracts have been developed. When one of the pyrrole rings of distamycin is replaced by an imidazole ring, the resulting ligand known as lexitropsin showed similar
affinity against G-C and A-T. In this case, the imidazole ring can form H-bonding with the NH$_2$ group of guanine.$^{133}$

![Netropsin, Distamycin, Lexitropsin](image)

**Figure 2-5. Structures of dsDNA groove binders.**

Studies on netropsin and its analogues revealed the polyamide motif could be useful for sequence-specific recognition of DNA via groove binding. In light of this finding, Dervan and co-workers designed and synthesized an array of hairpin-like polyamides with tandem imidazole and pyrrole rings: imidazole (Im), pyrrole (Py), and 3-hydroxypyrrole (Hp). The binding of these ligands follows a base-specific recognition rule that four ring-pairings (Im/Py, Py/Im, Hp/Py, and Py/Hp) distinguish all four Watson-Crick base pairs in the minor groove of DNA.$^{134}$ By constructing a polyamide with a β-Alanine linker, Dervan and co-workers were able to specifically target a 16 nt sequence in a HIV regulatory region.$^{135}$ More Recently, Sugiyama and co-workers
reported the recognition of telomeric duplex DNA using a polyamide trimer.\textsuperscript{136}

### 2.1.4. DNA intercalation vs. groove binding

Both DNA intercalators and groove binders have aromatic rings in their structures that interact with nucleobases. Most of the intercalators employ fused ring structures to form a rigid surface that stacks on nucleobases. Groove binders tend to utilize free flexible linkers that fit the molecules into the DNA groove. It is important to note that these two binding modes are not utterly exclusive, and some structures can adopt both of them simultaneously or depending on the target sequence. For instance, DAPI (4',6-diamidino-2-phenylindole) is extensively used in fluorescence microscopy and cell flow cytometry as a DNA dye (Figure 2-6). DAPI had been considered as a classical DNA groove binder for a long period of time preferably binding to AT-tract minor grooves. Later, an NMR titration experiment revealed that DAPI also intercalates to G-C base pairs, and the two binding modes co-exist under the same conditions.\textsuperscript{137} Other DNA intercalators such thiazole orange that contain non-fused aromatic rings have also been observed with two binding modes.\textsuperscript{138}

![Figure 2-6. Molecules with two binding modes against DNA.](image-url)
2.2. G-quadruplex Binding

DNA G-quadruplex is formed by guanine-rich DNA single strands with stacking G-quartets. Because of the abundant presence of G-rich sequences in telomeres and oncogene promoter regions, G4 ligands have become a fast-developing area of research. By targeting G-quadruplexes, telomerase mediated telomeres length maintenance in most cancer cells can be inhibited and leads to cell cycle arresting. Oncogene such as c-myc can also be down-regulated after the treatment of G4 ligands. Therefore, many G4 ligands have been developed so far with various core structures and side chains. For the clinical applications of G4 ligand in vivo, precisely targeting G-quadruplex without interferences with duplex DNA is always the major challenge. Thus, both affinity and selectivity are critical for a successful G4 ligand. Because of their planar aromatic and negative charged structures, most of duplex DNA intercalators can also bind to G-quadruplexes with different affinities and selectivities. Due to the shape of the G-quartets in G-quadruplexes, ligands with larger aromatic surfaces, especially those with similar geometry as a G-quartet, are more suitable for G-quadruplex binding. Therefore, structures including porphyrin, acridine, quindoline, and phenanthroline become the core archetypes of many well-developed G4 ligands. Side chains with positive charges, flexible linkers or sugar rings that can fit into the grooves or associate with the loops are often used to enhance both affinity and selectivity. In this section, the most potent G4 ligands are discussed according to their binding mode: end stacking, intercalating and groove/loop binding.

2.3.1. G-quadruplex end stacking ligands. Most of the aromatic ligands adopt an end stacking mode at the one or two terminal G-quartets through π-π stacking force.
The most potent G4 ligand with the best selectivity so far is a natural product
telomestatin (Figure 2-7). Telomestatin was discovered when researchers were screening
for telomerase inhibitors from natural products. It is a macrocyclic compound with five
oxazole rings and isolated from Streptomyces anulatus \textit{3533-SV4}. The IC$_{50}$ of telomerase
inhibition of telomestatin is as low as 5 nM which is 1000-fold lower than one of the
strongest synthetic G4 ligands TMPyP4.\textsuperscript{139} An ESI-MS study showed that telomestatin
stacks on the two terminal G-quartets of G-quadruplexes.\textsuperscript{140} A molecular simulation study
revealed that the heterocyclic structure of telomestatin provides rich positions for
hydrogen bonds with guanines of the G-quartets. The nitrogen atoms in the center of
telomestatin interact with the carbonyl channel formed by guanines through electrostatic
force.\textsuperscript{141} More importantly, telomestatin is the best selective ligand against G-quadruplex
and has almost no binding affinity to duplex DNA. The length of telomeric G-overhang
was shortened when the cells were treated by telomestatin.\textsuperscript{142} The IC$_{50}$ value of
telomestatin against two telomerase positive cell lines SW39 and SW26 are 4.1 and 1.8
µM, respectively, without obvious cytotoxicity towards normal cells.\textsuperscript{143} The high binding
affinity and excellent selectivity of telomestatin is the consequence of its macrocyclic
structure that facilitates its stacking on G-quartets and prevents the intercalation with
duplex DNA. However, due to its structural complexity, the large-scale synthesis of
telomestatin is always problematic. The total synthesis of telomestatin gave a very low
yield, and isolation from \textit{Streptomyces anulatus} only yielded 3.2 mg from 84 L culture
fluid.\textsuperscript{144} Several attempts have been made to develop the telomestatin derivatives by
Kazuo Nagasawa and co-workers. Two promising candidates of them, 6OTD and 7OTD,
exhibited fairly good affinity and selectivity comparable to telomestatin.\textsuperscript{145}
It is noteworthy that some experiments revealed that telomestatin exhibited cell toxicities on ALT (Alternative Lengthening of Telomeres) cell lines, suggesting that telomerase-positive is not the requirement for a G4 ligand to perform its anti-cancer activities. In ALT cells, topoisomerase III binds onto telemetric region and regulates the recombination and elongation of telomeres. This process can be inhibited by stabilizing the G-quadruplexes using telomestatin.

Cationic porphyrin compound TMPyP4 was first synthesized as a photosensitizer and later was used as an intercalator for duplex DNA via a threading mode (Figure 2-7). In addition, it was found to stabilize G-quadruplexes with an end stacking mode.
due to its large aromatic surface and four N-methylated side chains.\textsuperscript{148} TMPyP4 is one of the strongest G4 ligands, but with a poor selectivity. TMPyP4 shows positive effects on down-regulating the expression of oncogene \textit{c-myc} by inducing G-quadruplexes in its promoter region. The expression of the subunit of telomerase hTERT, which is regulated by c-MYC was also reduced after the treatment of TMPyP4.\textsuperscript{149} The IC\textsubscript{50} values of TMPyP4 against tumor cell lines are in the range of 9.0-28.2 µM, and between 1.7-15.5 µM against normal cell lines, further suggesting its low selectivity.\textsuperscript{150} Although with a poor selectivity, TMPyP4 is still commonly used as a positive control since it can be conveniently purchased.

Many modifications have been attempted to improve the selectivity of TMPyP4. By coordination with the metal ion Mn\textsuperscript{3+}, its selectivity against G-quadruplex over duplex DNA was improved by 10-fold. In addition, side chains containing amides were discovered to improve the binding affinity and selectivity of TMPyP4 derivatives. Thus, a cationic porphyrin Mn\textsuperscript{3+} complex with extended amide side chains was found with 1000-fold higher selectivity compared to TMPyP4. This molecule has been used to locate G-quadruplex in living cells (Figure 2-7).\textsuperscript{151-153}

BRACO-19 is one of the best synthetic G4 ligands with a remarkable selectivity, in which an acridine core is surrounded by three protonated side chains (Figure 2-7). Molecular modeling experiments suggested that the acridine moiety in the structure stacks on the G-quartets and each of the side chains occupies a groove. The selectivity of BRACO-19 is dependent on the types and positions of the side chains.\textsuperscript{154} BRACO-19 was cytotoxic against human uterus carcinoma cell line UXF1138L with an IC\textsubscript{50} value at 2.5 µM and significant telomere shortening was observed after a long-term treatment. The G-
quadruplex-regulated expression of telomerase subunit hTERT was found significantly decreased after the treatment of BRACO-19. In another study, BRACO-19 also exhibited anti-HIV activity by stabilizing the viral G-quadruplexes. Similar to many other DNA ligands, BRACO-19 has poor pharmacological properties mainly because of its high polarity. Therefore, the pre-clinical studies of BRACO-19 is still being conducted to improve its uptake efficiency.

Similar to TMPyP4 and BRACO-19, there have been many other examples of G4 ligands equipped with positively charged side chains. PIPER, a perylene based compound, has two piperidine moieties with positive charges on the terminals and can efficiently stack with G-quadruplexes and inhibits telomerase (Figure 2-7). An interesting observation for PIPER is that their binding selectivity against G-quadruplexes over duplex DNA is correlated their aggregation states in solution. Monomeric molecules intercalate into duplex DNA and aggregated products stack on G-quartets.

From above well-developed G4 ligands, many lessons of designing highly selective G-quadruplex stacking ligands can be learned: 1) A highly extended aromatic system is the best core structure for G-quadruplex stacking and provides the best selectivity (like telomestatin); 2) For relatively small ligands with practical synthetic yields, the side chains are crucial. The structures and locations of the side chains do not only improve the binding affinity but also affect the selectivity for G-quadruplexes over duplex. 3) Potent G4 ligands are always tethered with side chains with positive charges and optimized geometry for G-quadruplex interactions.

Metallo-aromatic structures are another group of potent G-quadruplex binders in
which small molecular ligands coordinate with a metal ion and form an extended aromatic surface. Other than the planar geometry of the coordinating ligands stacking on G-quartets, the metal ions at the center also interact with the guanine channel through electrostatic force. 1,10-Phenanthroline is an example of such aromatic ligands for G-quadruplex recognition, which will be extensively discussed in section 2.4.

**2.3.2 G-quadruplex groove binders.** G-quadruplex is formed by four successive G-rich segments where adjacent sequences form loops and provide distinct recognition environments for groove binders. Pyrrole amide antibiotics were discussed in the section of duplex DNA groove binders and they can also bind to G-quadruplex grooves (Figure 2-8). An NMR study revealed that distamycin dimers can bind with two opposite grooves in a G-quadruplex through hydrogen bonds and electrostatic force. Widened grooves were observed to accommodate the distamycin molecules. In another study of distamycin derivatives, the formamide group in distamycin was replaced by a charged N-formimidoil moiety and the resulting compound decreased the duplex DNA binding (Figure 2-8). Alkaloids peimine and peiminine were also reported interact with G-quadruplex through groove binding (Figure 2-8). These two compounds showed clear preference to a parallel G-quadruplex rather than duplex DNA, which may result from the different geometry of the grooves in these structures.

Lastly, ligands that intercalate into the G-quartets have not been discovered experimentally to our best knowledge, and only a few computational studies were reported. The G-quadruplex structure is highly compact, which makes intercalation difficult.
In summary, G4 ligands have recently attracted much attention for their anticancer activities. Though natural product telomestatin exhibited excellent binding affinity and selectivity against G-quadruplexes, its clinical applications are hindered by the extremely complicated synthesis. Therefore, the development of small molecules that selectively target G-quadruplexes is highly demanded. Lessons can be learned from the above studies that both the binding affinity and selectivity are significantly affected by side chains. The numbers, orientations and positive charge densities are the factors for potent binding. In light of these findings, we selected two core structures, 1,1-phenanthroline and thiazole orange and studied the binding affinity and selectivity toward G-quadruplex DNA. A brief summary of the research on these two types of compounds will be presented below.
2.3. 1,10-Phenanthroline Analogues Metal Complexes as G-quadruplex Ligands.

1,10-Phenanthroline (1,10-Phen) is a heterocyclic organic compound that chelates a broad-spectrum of metal ions (Figure 2-9). The structure of 1,10-Phen is similar to 2,2'-bipyridine, but with a much higher coordinating efficiency. It has been used as a photometric reagent for chelating Fe$^{2+}$ ions in solution. In addition, 1,10-Phen is commonly used as an important building block for synthesis of alkaloids and their analogues.

![1,10-Phenanthroline and 2,2'-bipyridine](image)

Figure 2-9. Structures of 1,10-phenanthroline (left) and 2,2'-bipyridine (right).

When DNA G-quadruplex became a promising target for cancer therapy, 1,10-Phen has been used as a building block for developing G4 ligands because of its extended aromatic surface. In addition, the chelated metal ions may enhance the interactions by substituting the cations in the G-quadruplex. It is commercially available and can be readily synthesized. In this section, the previous development of 1,10-Phen and its analogues as G4 ligands will be reviewed and serve as the important background of our research in the following chapter.
2.4.1. 1,10-phenanthroline analogues as G-quadruplex ligands. As discussed in previous sections, the two common features of G4 ligands are extended aromatic surfaces and suitable positively charged side chains. 1,10-Phen has a well-defined crescent shape with fused aromatic rings, which is a suitable trait of G4 ligands. De Cian and co-workers designed a 1,9-disubstituted-1,10-phenanthroline compound (Phen-DC3), in which, two quinolinium groups were anchored to the 1 and 9 positions of 1,10-Phen by amide linkers (Figure 2-10a). Phen-DC3 at 1 µM elevated the melting temperature of a human telomeric G-quadruplex by 29.7 °C monitored by fluorescence. More importantly, Phen-DC3 showed a remarkable selectivity against G-quadruplexes, which is comparable to telomestatin and better than most of the known synthetic G4 ligands. An X-ray crystallography revealed that Phen-DC3 has a similar surface size compared to the G-quartet providing optimized stacking force. Later, an NMR titration experiment showed that Phen-DC3 binds to G-quadruplexes by end stacking mode with 1-1 and 1-2 ratios. The phenanthroline core structure and two quinolinium rings overlap with the four guanines of the terminal G-quartets. It is noteworthy that the two amide linkers in Phen-DC3 help to lock the structure in a planar conformation that greatly enhances its G-quadruplex binding affinity. Removing this amide connectivity will lose its G-quadruplex stabilization effects. The in vivo G-quadruplex stabilization of Phen-DC3 showed that it can efficiently inhibit G-quadruplex unwinding by helicase without affecting other regular dsDNA.

Likewise, Wang et al. developed another group of 2,9 disubstituted-1,10-Phen (Figure 2-10b). The amide linkers were preserved because of the hydrogen bond locking effects. Cationic piperidine, pyridine, and tertiary amine groups were attached as side
chains. Slightly better G-quadruplex stabilization was achieved for these structures, compared to Phen-DC3. It was also proved that these ligands preferably bind to anti-parallel G-quadruplexes. Cell viability experiments revealed their cytotoxicity was shown at approximately 2 µM.\textsuperscript{168}

![Figure 2-10. Structures of Phen-DC3 (left) and its cationic amino side chain derivatives (middle); 4,7-diamino-1,10-Phen derivative (right).](image)

Other than 2,9 disubstituted-1,10-Phen, 4,7 disubstituted derivatives were also studied. Nielsen and co-workers synthesized a range of 4,7-diamino-1,10-Phen. Many of them stabilized G-quadruplex examined by CD (Figure 2-10c). The derivatives with longer side chains bearing positively charged guanidine groups exhibited better stabilization compared to shorter and neutral side chains. It was also noteworthy that 1,10-Phen itself did not stabilize G-quadruplexes in the absence of metal ions because of its small planar structure.\textsuperscript{169}

To the end, many modifications have been made at the 2,9 and 4,7 positions of 1,10-Phen. To our best knowledge, reports on modifications made at 5 and 6 positions of
1,10-Phen for developing G4 ligands are scarce. NMR and X-ray crystallography studies of 1,10-Phen derivatives suggested that side chains at 5 and 6 positions could effectively interact with G-quadruplex grooves.164-165

2.4.2. 1,10-phenanthroline analogues metal complexes as G-quadruplex ligands. Because of the two well-organized nitrogen atoms, 1,10-Phen is an outstanding chelating agent which can form complexes with a broad range of metal ions including nickel(II), copper(II), ruthenium(II), platinum(II), manganese(II), magnesium(II), and zinc(II) with 1-1, 1-2, or 1-3 ratios. By forming a complex, the extended aromatic surface can significantly contribute to G-quadruplex stacking. Meanwhile, metal ions may sit at the center of the channel of a G-quadruplex and coordinate with the carbonyl oxygens of the guanines, further enhancing the stability. Thus, many 1,10-Phen metal complex derivatives have been synthesized.

Nickel salphen (Figure 2-11a) is the first metal complex used as a G4 ligand and synthesized by Reed and co-workers.170 Two piperidine structures were attached as the side chains. Computation studies revealed that this square planar structure overlaps with the terminal G-quartets and the Ni²⁺ ion lies above the center channel. Nickel salphen exhibited an outstanding G-quadruplex stabilization and selectivity comparable to BRACO-19.

The success of nickel salphen inspired the development of metal complexes as G4 ligands. In this scenario, the same research group designed a Pt-1,10-Phen complex with a piperidine side chain on the carbon 9 (Figure 2-11b). This compound adopts an amide linker that similar to Phen-DC3 and the nitrogen atoms coordinate with the Pt²⁺ ions. A
computational study suggested that the conformation is fixed in a plane so that the benzene ring stacks on the guanines, and the extended piperidine side chain interacts with the grooves. The Pt\(^{2+}\) ion was chosen because it forms a stable planar 1-2 complex with the ligand.\(^{171}\) The Pt\(^{2+}\) ion sits in the center of the G-quadruplex and may replace the cationic ions in the G-quadruplex channel. In a FRET thermal denaturation experiment, the complex at 1 µM increased the melting temperature of a telomeric G-quadruplex by 20 °C while showed negligible effects on duplex DNA. This stabilizing effect was highly dependent on the presence of Pt\(^{2+}\) ion and only 9 °C increment of melting temperature was observed in the absence of platinum ions.\(^{172}\)

![Figure 2-11. Structures of 1,10-phenanthroline derivatives.](image)
In another report, Pt-1,10-Phen complex (Figure 2-11c) exhibited stabilization effects on an antiparallel G-quadruplex in the absence of any side chain. In the presence of 1 µM complex, the melting temperature of the telomeric G-quadruplexes was enhanced by 21 °C. An X-ray crystallography experiment revealed that Pt-1,10-Phen complex has a dimension matching the G-quartets. Interestingly, this complex has a remarkable preference to antiparallel G-quadruplexes. By replacing 1,10-Phen with bipyridines (Figure 2-11d), the resulting complex did not stabilize the G-quadruplex probably because of its flexibility, suggesting that the rigid extended aromatic structure of 1,10-Phen is pivotal for its G-quadruplex stabilization effect.\textsuperscript{173} Ruthenium ions (Ru\textsuperscript{2+}) are also frequently used in this type of complexes. Different from Pt\textsuperscript{2+}, Ru\textsuperscript{2+} ions usually form 1-3 complexes with ligands. Many Ru-1,10-Phen complex derivatives have been developed as G4 ligands and exhibited various stabilization effects.\textsuperscript{174-178}

Like in the nickel salphen, cheaper metal ions were often utilized in the complexes to stabilize G-quadruplexes. Bianco and co-workers synthesized a bis-phenanthroline, in which, two phenanthrolines were linked by a tertiary amine (Figure 2-11e). Mn\textsuperscript{2+}, Cu\textsuperscript{2+}, Ni\textsuperscript{2+}, Mg\textsuperscript{2+}, and Zn\textsuperscript{2+} ions were all studied with this ligand and the formation of 1-1 (1 metal ion coordinated with two phenanthroline function groups) complexes was observed.\textsuperscript{179} A FRET assay revealed that Ni\textsuperscript{2+} complexes had the best G-quadruplex stabilization effect amongst the above metal ions, followed by Cu\textsuperscript{2+} and Zn\textsuperscript{2+} complexes. Mn\textsuperscript{2+} and Mg\textsuperscript{2+} ions had a very low affinity with the bis-phenanthroline ligand, so the complexes could not be formed readily in the solution. Compared to Pt\textsuperscript{2+} and Ru\textsuperscript{2+} complexes, the Ni\textsuperscript{2+} complex is more cost-friendly and, more importantly, can be readily formed in solution without special reaction conditions. For Cu\textsuperscript{2+} complexes,
DNA cleavage effect was frequently observed in the presence of a reducing reagent since 
$\text{Cu}^{2+}$ ion can undergo a reduction reaction to produce reactive species for DNA damage. 
$\text{Cu}^{2+}$-Bis-Phen complexes showed less cleavage of DNA because the generation of reactive species requires a perpendicular arrangement of two aromatic functional groups which is unlikely to happen for a bis-phenanthroline ligand.

Musetti and co-workers synthesized 2,9-disubstituted (K35, Figure 2-11f) and 7-monosubstituted (K34, Figure 2-11g) 1,10-Phen. The $\text{Ni}^{2+}$ and $\text{Cu}^{2+}$ complexes were investigated for their G-quadruplex stabilization. K34 formed a 1-2 complex with $\text{Ni}^{2+}$ and K35 only formed a 1-1 complex, because the bulky groups at 2,9 positions hindered the coordination of metal ions with the phenanthroline ligands. As a consequence, the G-quadruplex stabilization effect of K35 was not as good as K34, because K35 had a much smaller stacking surface. K34 could only form a 1-1 complex with $\text{Cu}^{2+}$, which in turn had a lower G-quadruplex binding affinity compared to the $\text{Ni}^{2+}$-K34 complex. All these observations suggested that forming 1-2 complexes between metal ions and 1,10-Phen is critical for the G-quadruplex stabilization effects. Modifications on 2 or 9 positions of 1,10-Phen may affect the coordination of ligands with metal ions because of the steric hindrance of the substituent(s). This was also confirmed by another 9-substituted-1,10-Phen derivative P120 (Figure 2-11h), which only formed a 1-1 Ni complex with a moderate G-quadruplex binding affinity.

Many 5 and 6 substituted 1,10-Phen compounds have been developed mainly for photophysical detection of metal ions. Reports on the use of them as G4 ligands are limited. Fang and co-workers synthesized a $\text{Ni}^{2+}$ complex NPH (Figure 2-11i), by coordinating a $\text{Ni}^{2+}$ ion with two 1,10-Phen ligands and one p-idip (p-idip=2-(4-indole)-
imidazo-[4,5-f][1,10]phenanthroline). This 1-3 complex could stack on the G-quartets with a higher affinity than Ni\textsuperscript{2+}-1,10-Phen complexes. A FRET assay showed that a DNA G-quadruplex had a 14 °C increment of T\textsubscript{m} in the presence of 1 µM of NPH.\textsuperscript{184}

To that end, 1,10-Phen derivatives have been developed into a class of G4 ligands and several of them exhibited very prominent binding affinity and selectivity. Side chains have been attached onto different locations of 1,10-Phen. Several 2,9 substituted 1,10-Phen derivatives are much better G4 ligands compared to 1,10-Phen. However, the substituents at 2 and/or 9 positions could hinder the coordination with metal ions. Ni\textsuperscript{2+} was the most cost-friendly and efficient metal to form 1,10-Phen-based G4 ligands. To improve the coordination efficiency with metal ions, side chains should be attached to the outer edge (4, 5, 6, or 7 positions) of 1,10-Phen if we recognize it as a crescent shape.

2.4. TO Derivatives as G-quadruplex Ligands

Thiazole orange (TO) is a cyanine dye that was firstly synthesized by Linda G. Lee for the staining of reticulocytes RNA in flow cytometry analysis.\textsuperscript{185} Most of this type of DNA dyes contain two aromatic rings that skew from each other when they are free in solution. They bind to DNA via intercalation by stacking between two nucleobase pairs, and the fluorescence arises when the two aromatic rings are constrained on the same plane.\textsuperscript{186} TO has a N-methylated benzothiazole ring linked to a quinolinium by a methine bridge (Figure 2-12a). It intercalates with duplex DNA with a stoichiometric ratio of 1 dye per 2 base pairs. TO also binds to DNA G-quadruplexes via end-stacking. The binding affinities of TO with G-quadruplex and duplex DNA are approximately the same. As a DNA dye, TO has many advantages compared to its predecessors such as high quantum yield and good cell membrane permeability. When it binds to DNA, it shows
more than 1000 times fluorescence increment.\textsuperscript{187} In addition, TO is permeable to cell membrane which makes it highly useful for staining of nucleic acid in living cells.

Figure 2-12. Structures of thiazole orange and derivatives.

Utilizing TO as a DNA dye, David Monchaud \textit{et al.} developed a fluorescent intercalator displacement (FID) assay which has been extensively used in G4 ligands research.\textsuperscript{188} The design rationale of the FID assay is as follows. Firstly, the binding of TO to DNA is reversible; therefore, the bound TO could be displaced out of DNA by other strong DNA ligands in solution. Because of the drastic difference in fluorescence between free TO and bound TO, one can monitor the fluorescence change to determine the binding affinity of the ligand of interest. The lower the concentration (DC\textsubscript{50} value) of the ligand is needed to displace 50\% TO, the stronger the ligand binds to DNA. Secondly, because thiazole orange binds to G-quadruplex and duplex DNA with almost the same affinities,
the FID results can be used to evaluate the selectivity of the ligand for G-quadruplex over duplex DNA.

Due to its distinct property of fluorescence, TO has been used as a structural motif for developing G4 ligands. For instance, some known G-quadruplex selective moieties were fused onto the quinolinium ring of TO such as benzofuroquinolinium (d), isaindigotone (c), and pyridodicarboxamide bisquinolinium (b) (Figure 2-12). These extended aromatic structures all exhibited improvements on G4 selectivity over duplex compared to thiazole orange. These TO based derivatives could be used as good G-quadruplex probes but their binding affinities to G-quadruplex DNA are moderately increased compared to TO. In addition, the synthetic approach is not convenient to generate a large library of TO with various side chains.

Interestingly, a recent report suggested that TO can bind to a string of consecutive G-quadruplexes with high selectivity over duplex DNA. In the same report, the authors even detected the strong binding of TO to the G-quadruplex string in the absence of Na$^+$ and K$^+$. In addition, the binding of TO with other noncanonical DNA structures was observed. Given the fact that consecutive G-quadruplexes are most likely formed in vivo, this report have aroused more interest to use TO as a template to develop novel G4 ligands
Chapter 3. 5-Substituted 1,10-Phenanthroline Ni(II) Complexes as G4 Ligands

G-quadruplex has received much attention because of its important regulatory functions in telomerase-mediated cancerization and oncogene expression. Phenanthroline is a type of G4 ligand scaffold that coordinates with metal ions, and the consequent complex has an extended aromatic surface that can aptly stack on G-quadruplexes. In this project, a series of 5-substituted phenanthroline-based Ni(II) complexes were investigated for their ability to bind to telomeric G-quadruplex DNA and inhibit telomerase. Biophysical studies including Förster Resonance Energy Transfer (FRET) G-quadruplex thermal denaturation, Circular Dichroism (CD) titration, and Fluorescence Intercalator Displacement (FID) indicated that complexes with either cationic side chains or aromatic side chains bound to and stabilized telomere G-quadruplex DNA with significant affinities. More importantly, complexes with arylsulfanyl side chains at the 5-position exhibited improved selectivity against G-quadruplex over duplex DNA, which can be used as a promising motif for G4 ligand development in the future. Results from a Telomeric Repeat Amplification Protocol (TRAP) assay suggested that these complexes can inhibit telomerase activity in vitro with micromolar concentrations. Cytotoxicity was also observed for these complexes against telomerase positive cancer cell lines.
3.1. Introduction

Cancer is the second leading cause of death in the United States and was responsible for 22.5 percent of deaths over the last five years.\textsuperscript{193} More than 100 types of cancer have been discovered so far while finding the efficient therapy with low side-effects is still a huge challenge. The most common treatments for cancer are surgery and chemotherapy, and obviously, surgery is only feasible for solid tumors, not metastatic cancers. The very first chemotherapy drug nitrogen mustard was investigated in 1942, which had been used as a chemical weapon in World War I. Since then, many other cytotoxic compounds have been developed as anti-cancer drugs. Most of these compounds were designated to target DNA and subsequently lead to cell cycle arrest by forming stable adducts or cross-links in DNA. Regression of tumors was observed after the treatment of these drugs; however, the acute and long-term toxicity becomes a huge obstacle in their applications for cancer chemotherapy. In the 1980s, thanks to the advanced knowledge and technologies of cell biology, a network of cell signaling pathways was deciphered, and some biomarkers of cancer were discovered and became the targets of receptor inhibitors. Glivec (a kinase inhibitor), Gefitinib (an EGFR inhibitor), and Avastin (a VEGF antibody) are the benchmarks of these targeting drugs and exhibit outstanding effects on several major types of cancer. However, the over-expression of the corresponding biological targets is patient-dependent, and hence only individuals with a certain genotype, determined by genomic tests or proteomics, are sensitive to specific drugs. Therefore, in spite of acute side-effects, cytotoxic drugs are still playing pivotal roles in cancer therapy and usually are combined with the targeting inhibitors, nowadays.\textsuperscript{194} The development of new broad-spectrum chemotherapy drugs
with low side-effects is still a pressing task for researchers.

DNA is the target for most of the chemotherapy drugs. Some of them like nitrogen mustard and cisplatin form stable DNA adducts or cross-links, and thus inhibit the DNA replication. However, these drugs are non-selective and lead to acute side effects. As mentioned in the previous chapters, G-quadruplex is a secondary DNA structure that is formed by successive G-rich sequences through Hoogsteen hydrogen bonds. G-quadruplex is prevalently present in oncogene promoter regions and telomeric regions. Its formation and stability are closely correlated to cancer cell proliferation. Oncogene such as c-myc can be down-regulated in the presence of stabilized G-quadruplexes. Meanwhile, G-quadruplexes in telomeric regions hinder telomerase binding and inhibit the consequent telomere length maintenance, which is only present in cancer cells. In light of these observations, many strategies have been developed to use G-quadruplexes as a target for potential anti-cancer therapies.

Like duplex DNA ligands, some small molecules have been found interact with G-quadruplexes, and consequently increase their stability. There are two principles of successful G4 ligands: 1) high binding affinity with G-quadruplex DNA; 2) sufficient selectivity against G-quadruplex over duplex DNA. Therefore, most of the outstanding G4 ligands were designed with two types of functional motifs: 1) positively charged side chains that interact with DNA backbones to enhance the binding affinity; 2) an extended planar aromatic system that can stack with the G-quartets to enhance the selectivity. A natural product, telomestatin is one of the most potent G4 ligands with remarkable selectivity. However, its structural complexity became a huge obstacle for its efficient synthesis and production as a drug. Some synthesized small molecules have been
developed targeting G-quadruplex utilizing fused aromatic ring structures including acridine, perylene, naphthalene, phenanthroline, etc. Many of these compounds have high binding affinities against G-quadruplexes; however, very few of them have sufficient selectivity. Several phenanthroline derivatives exhibited remarkable affinity and selectivity against G-quadruplex DNA. In addition to having a rigid aromatic crescent structure and various side chain attaching positions, phenanthroline can also coordinate with metal ions and form complexes with a large stacking surface. The metal ions sit in the center of the planar surface, which may replace the original cations of G-quadruplexes and coordinate with the guanines.

Many of these phenanthroline derivatives have been synthesized, and their G-quadruplex stabilization effects were evaluated in the presence and absence of metal ions. A majority of developed derivatives were modified by tethering cationic moieties such as bisquinolinium, piperidine, pyridine, and amino groups at the 2 and 9 positions of phenanthroline and exhibited remarkable G-quadruplex binding affinities. The modifications at the 4 and 7 positions of phenanthroline has also been reported, and the resulting derivatives showed limited selectivity for G-quadruplex over duplex DNA. To the best of our knowledge, the modification at 5 or 6 positions of phenanthroline for G-quadruplex stabilization has not been extensively studied. NMR and X-ray crystallography studies revealed that when phenanthroline derivatives stack at the end of G-quadruplexes, the 5 and 6 positions of the phenanthroline ring are closely located at the G-quadruplex grooves. Hence, the interactions between side chains at these two positions and the G-quadruplex grooves could exist. A Ni complex of phenanthroline derivative with 2-(4-indole)-imidazole at the 5 and 6 positions exhibited
better G-quadruplex affinity and selectivity compared to the parent complex, suggesting that the 5 and 6 positions are attractive modifying sites for developing novel G4 ligands\(^\text{184}\).

Many phenanthroline derivative metal complexes have been synthesized with Pt\(^{2+}\) or Ru\(^{2+}\) ions, and showed significant G-quadruplex stabilization effects\(^\text{173,177}\). Complexes with similar structures were also prepared with cheaper metal ions including Cu\(^{2+}\), Ni\(^{2+}\), Mn\(^{2+}\), Mg\(^{2+}\), and Zn\(^{2+}\), in which, Ni\(^{2+}\) and Cu\(^{2+}\) ions formed more stable complexes with phenanthroline derivatives than the others\(^\text{179}\). Other than the lower cost, Ni\(^{2+}\) ions have another advantage over Pt\(^{2+}\) and Ru\(^{2+}\) ions that they can be conveniently coordinated by mixing the metal salts and ligands in water solutions without any special treatment\(^\text{195-196}\).

A facile synthesis of 5-substituted phenanthroline derivatives was reported by Samoshin and his co-workers. They kindly provided us nineteen 5-substituted phenanthrolines for evaluation\(^\text{197}\). Using their method, our group also synthesized an additional derivative (21) with a piperidine side chain, which is a known moiety enhancing G-quadruplex binding. In the previous work, Ni\(^{2+}\) complexes of these twenty derivatives were prepared by simply mixing NiCl\(_2\) and the corresponding compounds with 1-2 ratio in water solutions. The G-quadruplex stabilization was also evaluated by FRET and CD thermal denaturation.
In the present research, the stoichiometry of the $\text{Ni}^{2+}$ coordination was first
confirmed by ESI mass spectrometry. FID was used to determine the G-quadruplex stacking effects of the complexes and the results were compared with other biophysical studies. *In vitro* telomerase inhibition and cytotoxicity were investigated using a modified TRAP assay and MTT experiments.

3.2. Experimental

3.2.1. Materials and general methods. Unless otherwise specified, all chemicals for synthesis were purchased from Sigma-Aldrich or Fisher Scientific and used without further purification. Oligonucleotides were synthesized on an Applied Biosciences 392 DNA/RNA synthesizer using standard phosphoramidite protocols and purified using polyacrylamide gel electrophoresis. All the chemicals for DNA synthesis were purchased from Glen Research (Sterling, VA). The concentrations of oligonucleotide solutions were determined by a Varian Cary 100 Bio UV-Vis spectrophotometer (Walnut Creek, CA), using the molar extinction coefficients (in units of mol of nucleotide/L·cm⁻¹) obtained from Oligo Analyzer 3.1 (www.idtdna.com). T4 polynucleotide kinase was obtained from New England Biolabs. [γ-³²P]-ATP was purchased from PerkinElmer (Waltham, MA). Quantification of 5'³²P-labeled oligonucleotides was carried out using a Storm 860 phosphorimager and ImageQuant 5.1 software (Molecular Dynamics). DNA labeling was performed by incubating [γ-³²P]-ATP (300 μCi) and T4 polynucleotide kinase (20 units) in the presence of an oligonucleotide at 37 °C for 45 min. Unreacted [γ-³²P]-ATP was removed using a MicroSpin G-25 column (GE Healthcare). Cancer cell lines were a generous gift from Dr. William K. Chan. Cell culture media and supplements (fetal bovine serum, L-glutamine, and antibiotic-antimycotic) were from Invitrogen. MTT (Thiazolyl blue tetrazolium bromide) was purchased from Fisher Scientific, and the
absorbance were read on a TriStar LB 941 multimode microplate reader.

### 3.2.2. Mass spectrometry analysis of complexes

A Varian 320L triple quadrupole mass spectrometer coupled to an electrospray ionization (ESI) source (Agilent Technologies, Santa Clara, CA) was used to analyze the complexes in positive ion mode. The sample solutions were prepared by dissolving 5-substituted phenanthroline derivatives and NiCl₂ in a 2:1 ratio with 50% methanol-water (unless otherwise specified), and the final concentration of complexes was 200 μM. An aliquot of the sample solution (20 μL) was introduced into the ESI source via a six-port valve. The following instrument parameters were employed: 4.5 kV for the ESI needle onset voltage, 30 V for the capillary voltage, 30 psi for the nebulizing nitrogen gas, and 10 psi and 150°C for the drying gas. The designated ions were isolated for the collision-induced dissociation (CID) measurements which were performed at collision energies of -40-- 50 eV.

### 3.2.3. Fluorescent intercalator displacement assay

A stock solution containing 100 μM human telomeric G-quadruplex sequence (PCG1) (5’-AGG GTT AGG GTT AGG GTT AGG GTG-3’) or duplex (D1) (5’-AGC GAT GAG AGG CCA CGA GGA ATC GCT GGT ACC G-3’/ 5’-CGG TAC CAG CGA TTC CTC GTG GCC TCT CAT CGC T-3’), 10 mM lithium cacodylate buffer (pH=7.0) and 50 mM KCl was prepared and incubated at 90 °C for 5 min, and slowly cooled to room temperature. All solutions were stored at 4 °C before further use.

The reaction solution containing 0.25 μM DNA, 10 mM lithium cacodylate buffer (pH=7.0) and 50 mM KCl was prepared from the stock solution before use. This solution
was incubated with 0.5 μM thiazole orange (TO) for 30 min in the dark. Aliquots of ligands were added into this mixture solution and equilibrated for 5 min in the dark before measurement. The fluorescence spectra (λ excitation: 501 nm, slit width: 5.0 nm, scan speed 120 nm/min) were recorded from 510-650 nm after each addition. The percent fluorescence displacement (%FI displacement) was calculated at λ max by using the formula %FI displacement=100-[(FA/FA0)×100], where FA0 is the initial fluorescence intensity in the absence of ligand, and FA is the fluorescence intensity upon addition of ligand.

3.2.4. Preparation of cell lysate for TRAP assay. Cell lysates containing telomerase were extracted from cultured HeLa cells. The cells from a 90% confluent 225 cm² culture flask were washed once with cold phosphate-buffered saline (PBS) and scraped off from the culture flask in 1 ml cold PBS. The cells were pelleted in a 1.5 ml centrifuge tube by spinning for 5 min at 1000 rpm and suspended in 1 ml cold TRAPEze® 1× CHAPS Lysis Buffer (Millipore, Sigma). The suspension was incubated on ice for 30 min and centrifuged at 12,000 X g for 10 min at 4°C. The supernatant was extracted and stored at -80°C.

3.2.5. Real-time PCR telomeric repeat amplification protocol (TRAP) assay.

The reaction mixture (30 µl) was prepared containing 1×SYBR Green Mix (Thermo Fisher Scientific, polymerase included), EGTA (1 mM), TS primer (2 μM, extension substrate of telomerase, 5’-AATCCGTCGAGCAGAGTT-3’), RP (1 μM, reverse primer for telomerase extended products, 5’-GCGCGGCTTACCCTACCCTAACC-3’), cell lysis (2.5 µl, lysis buffer was used in control samples) and ligand (various concentrations). Samples were incubated at 30°C for 30 min in a 96-well plate for telomerase extension followed by 95°C for 10 min to deactivate telomerase and activate
polymerase. A two-step PCR thermal cycle (95 °C for 15 s, 60 °C for 60 s) was performed for 41 times using an MJ research option-2 real-time PCR machine (St. Bruno, Canada) and fluorescence was recorded after each cycle using the SYBR Green setting mode of the instrument. The data was plotted as the logarithm of fluorescence signal versus the cycle number, and the threshold was set to the value of 10 times standard deviation of the background.

3.2.6. **32P telomeric repeat amplification protocol (TRAP) assay.** The 5′-[^32P] end-labeled TS primer (0.1 μg, 5′-AATCCGTCGAGCAGAGTT-3′) was incubated in a reaction mixture (50 μL) [buffer: Tris-HCl (20 mM, pH 8.3), MgCl₂ (1.5 mM), KCl (63 mM), 0.05% Tween 20, EGTA (1 mM)] with cell lysate (1 μL), RP (0.1 μg, reverse primer for telomerase extended products, 5′-GCGCGGCTTACCTACCCTACCCTAACCC-3′), NT (0.1 μg, primer for polymerase as internal standard, 5′-ATCGCTTCTCGGCCTTTT-3′), TSNT (0.01 amol, template for NT primer, 5′-AATCCGTCGAGCAGGTTAAGGGCCGAAGCGAT-3′), 2 units of Taq polymerase, dNTP mix (50 μM), and ligand (3.125-100 μM) at 30°C for 30 min and 94°C for 2 min to deactivate telomerase. The reaction mixture was then incubated using a three-step PCR method (94 °C for 20 s, 59 °C for 30 s, and 72 °C for 30 s) with a total of 25 cycles. The reaction solutions were mixed with 10 μL 10×loading dye and resolved with 10% denaturing polyacrylamide gel [5% crosslink, 45% urea w/w].

3.2.7. **MTT cell viability assay.** MCF-7 cells (human breast adenocarcinoma cells) and HeLa cells (cervical cancer cells) (a generous gift from Dr. William Chan, UOP) were maintained in advanced DMEM medium supplemented with 5% fetal bovine serum, L-glutamine and antimycotic (antibiotic) at 37 °C in a humid atmosphere
containing 5% CO₂. The compounds were pre-dissolved in sterilized water. The cells (3000 cells/well) were seeded in 96-well plate. After 24h incubation, the supernatant was removed followed by adding 200 µl FBS free media containing the compound. After 72h incubation, 20 µl MTT (3-(4,5-dimethylthiazole-2-yl)-2,5-di-phenyl tetrazolium bromide) (0.5 mg/ml) was added for further another 4h incubation. Formazan product was dissolved in 150 µl of DMSO, and the absorbance was measured at 590 nm using a TriStar LB 941 Multimode Microplate Reader. Untreated cells and media without cells were used as controls. To calculate the viability (%), the following equation was used:

\[
\text{Viability (\%) = } \frac{(A_{\text{treated}}) - (A_{\text{media}})}{(A_{\text{untreated}}) - (A_{\text{media}})} \times 100\%
\]

\(A_{\text{treated}}\): The absorbance of the solution containing treated cells

\(A_{\text{media}}\): The absorbance of the media

\(A_{\text{untreated}}\): The absorbance of the solution containing untreated cells

3.3. Results and Discussion

3.3.1. 5-substituted 1,10-phenanthroline formed 1-2 and 1-3 complexes with Ni(II) ions. Ni²⁺ ions were reported forming 1-1, 1-2 and 1-3 (Ni²⁺/Phen) complexes with different 1,10-Phen derivatives. In the previous work of our group, a continuous variation analysis (Job plot) was performed and the stoichiometry of the complexes 20-Ni and 21-Ni was determined. In these experiments, Ni²⁺ ions were titrated against compound 20 and 21, and the UV absorbance was measured after each addition. The molar extinction coefficients at 315 nm were plotted against the metal ion molar fraction [metal ion molar fraction = metal ion concentration / (metal ion concentration + compound]
concentration). The cross points were determined at 0.34 and 0.30 for 20-Ni and 21-Ni, respectively (Figure 3-1). This observation revealed the dominance of forming a 1-2 complex between Ni$^{2+}$ and these two compounds in solution.

![Job plots](image)

Figure 3-1. Job plots obtained based on continuous variation analysis for (A) 20-Ni and (B) 21-Ni.

To further confirm the stoichiometry, we used a triple quadrupole mass spectrometer coupled to an electrospray ionization (ESI) source to analyze the representative complexes that have a good solubility in methanol (1-Ni, 12-Ni, 15-Ni, 18-Ni, 20-Ni and 21-Ni). All spectra (Figure A-1 to Figure A-5) are shown in relative abundances with the most abundant ion in each spectrum corresponding to 100% ion intensity (base peak).

To study the effect of solvent on complexation, the spectra of Ni-1,10-Phen complex (1-Ni) were prepared in three different solution compositions with 0%, 50% and 90% methanol, respectively. Both of the complex ion $[1-2]^{2+}$, $m/z$ 209.02, and complex ion $[1-3]^{2+}$, $m/z$ 299.12 were identified (Figure 3-2). The $[1-3]^{2+}$ ion was the more abundant ion under all three conditions.
The relative abundance of the 1-2 complex increased from 49.21% to 82.20% with the increasing methanol ratio from 0 to 90%. This trend can be a result of the different dissociation rate of water molecules from Ni(II) chloride in different solvent compositions.199

Figure 3-2. ESI spectra of phenanthroline and Ni$^{2+}$ solutions with a 2:1 molar ratio in three different solvent compositions.

Then, five other derivative complexes (12-Ni, 15-Ni, 18-Ni, 20-Ni and 21-Ni) were analyzed in 50% MeOH solutions, and both 1-2 and 1-3 complexes were observed
for all of them (Figure A-1 to Figure A-5). Different from 1-Ni, all five derivative complexes had 1-2 complexes as the more abundant ions compared to the corresponding 1-3 complexes (Table 3-2). There was no 1-1 complex present in any of the samples. It was reported that the mechanism of the kinetics of complexation between Ni$^{2+}$ and phenanthroline derivatives are side-chain dependent. The formation of the complexes can be significantly affected by the solvation properties of the ligand with different side chains.\textsuperscript{199} Although the accurate explanation of the side-chain effect remains unclear, the five derivative complexes exhibited a notable preference for 1-2 complexation when compared to the parent compound 1-Ni. Interestingly, chloride ions were observed in the 1-2 complexes for 20-Ni and 21-Ni, but not for their 1-3 complexes. It is possible that the chloride ions coordinated with the Ni$^{2+}$ ion instead of interacting with the side chains; otherwise, it should also be found for 1-3 complexes and has the same number as the compound.

Table 3-2. Characteristic m/z values and the relative abundance (%) in ESI-MS spectra of the 1,10-Phen-Ni complexes.

<table>
<thead>
<tr>
<th>12-Ni</th>
<th>15-Ni</th>
<th>18-Ni</th>
<th>20-Ni</th>
<th>21-Ni</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2 complexes m/z</td>
<td>367.11$^a$</td>
<td>353.22$^a$</td>
<td>305.99$^a$</td>
<td>340.161$^a$</td>
</tr>
<tr>
<td>Relative abundance (%)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>134.28$^d$</td>
</tr>
<tr>
<td>1-3 complexes m/z</td>
<td>536.22</td>
<td>515.33</td>
<td>448.107</td>
<td>495.81</td>
</tr>
<tr>
<td>Relative abundance (%)</td>
<td>47.00</td>
<td>71.02</td>
<td>65.85</td>
<td>100</td>
</tr>
</tbody>
</table>

The collision-induced dissociation (CID) experiments were used to confirm the structures of \textbf{20-Ni} and \textbf{21-Ni}. Fragmentation of \textbf{20-Ni} complex \([1-2-\text{Cl}]^+\), \(m/z\) 714.8 produced a complex ion \([1-1-\text{Cl}]^+\), \(m/z\) 403.8 (Figure 3-3). The ion at \(m/z\) 99.9 representing a side-chain fragment ion (N,N-diethylethenamine) was observed and the corresponding mechanisms are assigned (Scheme 3-1).

![Figure 3-3. CID spectrum of 20-Ni complex ion [1-2-Cl]$^+$ at $m/z$ 714.8.](image)
Fragmentation mechanisms of Phen-Ni complexes in a CID experiment.

Fragmentation of complex ion $[1-3]^{2+}$, $m/z$ 495.4 produced a complex ion $[1-1]^+$, $m/z$ 369.1 and $[1-2]^+$, $m/z$ 679.8 (Figure 3-4). The ion at $m/z$ 579.7 corresponded to a fragment ion losing a side chain. Interestingly, the 1-2 complex ion at $m/z$ 679.8 and 1-1 complex ion at $m/z$ 369.1 in this spectrum were present with a reduced charge state (single charge) which was not observed for the fragment ions of 1-2 complexes. This unexpected reduction of metal ions was also reported by another group when they conducted electrospray mass spectrometry in positive mode, which was attributed to a charge transferring.200 The reduction was only found in the product ions of 1-3 complexes, indicating that the charge transferring can only take place when Ni$^{2+}$ is coordinated with three phenanthroline derivatives in the parent ions.
Similar to 20-Ni, 1-3 and 1-2 complexes of 21-Ni were also fragmented under the same conditions. Similar fragmentation patterns were observed.
A side chain fragment ion, $m/z$ 111.9 and a complex ion $[1\text{-}1\text{-Cl}]^+$, $m/z$ 415.8 was produced by the fragmentation of the complex ion $[1\text{-}2\text{-Cl}]^{2+}$, $m/z$ 739.2 (Figure 3-5). Fragmentation of the complex ion $[1\text{-}3]^{2+}$, $m/z$ 513.8 produced a complex ion $[1\text{-}1]^+$, $m/z$ 380.9 and a complex ion $[1\text{-}2]^+$, $m/z$ 703.6 (Figure 3-6). The ions at $m/z$ 591.6 was corresponding to a fragment $[1\text{-}1]^+$ complex ion losing one of the side chains. A fragment $[1\text{-}1]^+$ ion was also observed at $m/z$ 269.8. Similar to 20-Ni, reduced charge state was observed to the product 1-1 and 1-2 complex ions.
In summary, the observations from both continuous variation analysis and ESI MS spectrometry measurements suggested that the 5-substituted phenanthroline derivatives predominately form a 1-2 complex with Ni$^{2+}$ ions, which can potentially bind to G-quadruplexes with its enlarged aromatic surface. The 1-2 phenanthroline derived metal complexes gave better G-quadruplex stabilization effect than the other
complexation ratios due to their square planar geometry.\textsuperscript{180-181}

\subsection*{3.3.2. 5-Substituted Phen-Ni complexes induced telomeric G-quadruplex formation under Na\textsuperscript{+} and K\textsuperscript{+} deficient conditions.}

Circular Dichroism (CD) spectrometry is frequently used to analyze G-quadruplex conformations and the CD spectral features can be empirically correlated to the conformational structures of G-quadruplexes. In the present work, we used CD spectroscopy to confirm that 20-Ni and 21-Ni can induce the formation of G-quadruplexes in the absence of cations, which indicated their strong G-quadruplex binding affinities.\textsuperscript{198} In these experiments, a CD spectrum was first recorded for a solution containing the human telomeric DNA sequence PCG1 (G4, 5'-AGGGTTAGGGTTAGGGTTAGG GT-3') without Na\textsuperscript{+} or K\textsuperscript{+}. It exhibited a negative peak at 238 nm, a major positive peak at 257 nm and a minor positive peak at 295 nm, representing a typical random coil. Then, aliquots of 20-Ni and 21-Ni were gradually added into the solution respectively, and the CD spectra were measured after each addition. The change of the molar ellipticity revealed a transition of the conformation from a random coil to an antiparallel G-quadruplex conformation (Figure 3-7). With the concentration of the complexes increasing from 0 to 20 µM, the peaks at 238 and 257 nm decreased, a positive peak at 295 nm and a negative peak at 260 nm appeared. This observation is consistent with a previous report.\textsuperscript{169} In order to confirm that the G-quadruplex was induced by the complexes but not the free 5-substituted phenanthroline derivatives, 20 and 21 were also titrated against the human telomeric DNA sequence under the same condition. Slightly conformational change was observed for 20, but this change was significantly smaller compared to the one observed for 20-Ni. Addition of 21 gave nearly no change of the molar ellipticity (Figure 3-7). These
observations suggested that **20-Ni** and **21-Ni** are strong G4 ligands and promoted us to carry out further studies to achieve a broader picture for the G-quadruplex binding properties of the 5-substituted phenanthroline complexes.

Figure 3-7. CD spectra of PCG1 measured when titrating with A) 20-Ni; B) 20; C) 21-Ni; D) 21. The continuous changes in the CD spectra correspond to the incremental amount of Phen-Ni ranging from 0-20 μM. DNA = 5 μM.

**3.3.3. 5-Substituted Phen-Ni complexes with cationic and aromatic side**
chains increased the thermal stability of telomeric G-quadruplex DNA. In previous studies, a high-throughput thermal denaturation experiment of G-quadruplex DNA monitored by Förster Resonance Energy Transfer (FRET), was conducted to investigate the stabilization effects of all twenty-one complexes. In this experiment, a fluorescence (fluorescein) and a quencher (DABSYL) molecule were linked to the two ends of the same human telomeric G-quadruplex sequence (5’-Fluorescence AGGGTTAGGGTTAGGGTTAGGGT-3’-DABSYL) and form the G-quadruplex (G2). When the G-quadruplex is folded, the fluorescence signal can be quenched by the proximate quencher. When the G-quadruplex is dissociated at a temperature above the melting temperature (T\textsubscript{m}), the fluorescence signal can be restored because unfolding the G-quadruplex into a random coil makes fluorescein and DABSYL quencher apart. The T\textsubscript{m} is defined as the temperature, at which 50% of the G-quadruplexes dissociate into random coils. It can be extrapolated by the first derivatives from the plots of the fluorescence intensities against the temperatures. The G-quadruplex melting experiments were conducted in the presence of complexes Phen-Ni or their metal free compounds Phen with a 4:1 (ligand: DNA) ratio on a 96 well plate using a real-time PCR machine, and the derived ΔT\textsubscript{m} values (T\textsubscript{m} increment compared to the compound free control 57.7 °C) are shown in Table 3-3.
Table 3-3. T_m values of G-quadruplex derived from fluorescent thermal denaturation in the presence of free compound or Ni^{2+} complexes at a 4:1 ligand/DNA ratio.

<table>
<thead>
<tr>
<th>Compound</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Group</td>
<td>No ligand</td>
<td>H</td>
<td>Phen</td>
<td>Phen-Ni</td>
<td>Phen</td>
<td>Phen-Ni</td>
<td>Phen</td>
</tr>
<tr>
<td></td>
<td>57.5</td>
<td>57.3</td>
<td>58.4</td>
<td>58.4</td>
<td>57.8</td>
<td>58.4</td>
<td>56.6</td>
</tr>
<tr>
<td></td>
<td>57.5</td>
<td>62.3</td>
<td>64.4</td>
<td>60.0</td>
<td>67.7</td>
<td>65.0</td>
<td>66.5</td>
</tr>
<tr>
<td></td>
<td>57.2</td>
<td>57.0</td>
<td>57.0</td>
<td>57.0</td>
<td>57.0</td>
<td>57.2</td>
<td>56.0</td>
</tr>
<tr>
<td></td>
<td>58.2</td>
<td>58.4</td>
<td>68.0</td>
<td>60.2</td>
<td>57.6</td>
<td>63.2</td>
<td>60.0</td>
</tr>
<tr>
<td></td>
<td>58.4</td>
<td>56.4</td>
<td>56.4</td>
<td>57.0</td>
<td>59.2</td>
<td>57.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>59.0</td>
<td>64.2</td>
<td>62.0</td>
<td>60.2</td>
<td>69.6</td>
<td>68.4</td>
<td></td>
</tr>
</tbody>
</table>
observation was consistent with previously reported studies from other groups.\textsuperscript{169, 180-181}

Stabilization effects were observed for eighteen Ni-containing complexes except \textbf{12-Ni} and \textbf{15-Ni}. The $T_m$ value in the presence of \textbf{15-Ni}, with an octyl side chain, was 56.0 °C which was lower than the control (57.5 °C), suggesting that octyl group significantly affected the G-tetrad stacking ability of the ligand. In principle, it is well acknowledged that a long hydrophobic alkyl chain without positive charges cannot efficiently interact with DNA phosphate backbones.

Figure 3-8. $\Delta T_m$ ($T_m$ increment compared to the compound free control 57.7 °C) values of G-quadruplex derived from fluorescent thermal denaturation in the presence of Phen-Ni at a 4:1 ligand/DNA ratio. $\pi$: Complexes with aromatic side chains; $+$: Complexes with cationic side chains, and the pKa value of conjugate acid is between 2.0 and 7.0; $++$: Complexes with cationic side chains, and the pKa value of conjugate acid is higher than 7.0.
Nine (20-Ni, 21-Ni, 10-Ni, 4-Ni, 6-Ni, 5-Ni, 2-Ni, 17-Ni and 13-Ni) out of eighteen G-quadruplex stabilizing complexes have better stabilizing effects according to the $T_m$ values, compared to 1-Ni. In these nine compounds, five of them (20-Ni, 21-Ni, 10-Ni, 6-Ni, 5-Ni and 9-Ni) have side chains that can be positively charged (Figure 3-8).

To understand the significance of the positive charges in G4 ligands, we compared the pKa values of the positively charged sites in these compounds (Table 3-4). Complexes 5-Ni, 20-Ni, and 21-Ni all contain amines that can be readily protonated and form positive charges under physiological conditions (pKa value of conjugate acid > 7.0). Indeed, these three compounds were all found with stabilization effects for G-quadruplexes. Complexes containing nitrogen atoms with lower pKa values are unlikely to be protonated under the physical condition. Most of them were found with much less significant stabilization effects on the G-quadruplex such as 9-Ni, 14-Ni and 17-Ni. This observation is consistent with previously reported studies that cationic side chains can enhance G-quadruplex binding affinity through electrostatic interactions with DNA grooves.\textsuperscript{201}
Table 3-4. Approximate pKa value of conjugate acid of the designated side chains of 5-substituted phenanthroline.

<table>
<thead>
<tr>
<th>#</th>
<th>Name of the cationic motif</th>
<th>Structure of the side chain</th>
<th>pKa*</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Amine</td>
<td><img src="image" alt="Amine structure" /></td>
<td>11.0&lt;sup&gt;202&lt;/sup&gt;</td>
</tr>
<tr>
<td>6</td>
<td>Pyridine</td>
<td><img src="image" alt="Pyridine structure" /></td>
<td>5.2&lt;sup&gt;202&lt;/sup&gt;</td>
</tr>
<tr>
<td>9</td>
<td>Benzimidazole</td>
<td><img src="image" alt="Benzimidazole structure" /></td>
<td>5-6&lt;sup&gt;203&lt;/sup&gt;</td>
</tr>
<tr>
<td>10</td>
<td>Triazole</td>
<td><img src="image" alt="Triazole structure" /></td>
<td>NA</td>
</tr>
<tr>
<td>14</td>
<td>Benzothiazole</td>
<td><img src="image" alt="Benzothiazole structure" /></td>
<td>1.84&lt;sup&gt;204&lt;/sup&gt;</td>
</tr>
<tr>
<td>17</td>
<td>Acetanilide</td>
<td><img src="image" alt="Acetanilide structure" /></td>
<td>&lt;2&lt;sup&gt;205&lt;/sup&gt;</td>
</tr>
<tr>
<td>20</td>
<td>Tertiary amine</td>
<td><img src="image" alt="Tertiary amine structure" /></td>
<td>10.6&lt;sup&gt;202&lt;/sup&gt;</td>
</tr>
<tr>
<td>21</td>
<td>Piperidine</td>
<td><img src="image" alt="Piperidine structure" /></td>
<td>11.1&lt;sup&gt;202&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

*: pKa values were determined experimentally in the publications for a same functional group.

Except for positive charges, aromatic stacking is the other critical contributing factor for G-quadruplex stabilization. Out of the nine complexes that efficiently stabilize G-quadruplex here, four of them (10-Ni, 6-Ni, 17-Ni and 13-Ni) have aromatic side chains. Although, overall, the aromatic side chain tethered complexes exhibited less stabilization effect compared to those with cationic side chains in this assay, it is important to note that aromatic motifs may provide better selectivity for the G-quadruplex over duplex DNA because the former contains a more extended aromatic surface for stacking.

In addition to the positive charge and stacking surface, there are other factors
enhancing the G-quadruplex binding. 2-Ni and 17-Ni have a hydroxy or carbonyl group, which may form hydrogen bonds with nucleobases or DNA backbones as a hydrogen donor or acceptor.\textsuperscript{162-163} Hence, due to the complexity of G-quadruplex stabilizing factors, an accurate side-chain rule for Phen-Ni as G4 ligands cannot be successfully acquired. However, a general trend can still be concluded. Based on the results from the FRET assay, both cationic side chains and aromatic side chains are important factors for G-quadruplex stabilization.

In general, these observations were consistent with the previously reported G4 ligands that cationic motifs can significantly enhance the G-quadruplex binding affinity. However, it is important to notice that, although FRET is a frequently used method for screening G4 ligands in a high-throughput fashion, this method has its disadvantages. The interference from the fluorophore and quencher molecules exists.\textsuperscript{206} The add-on fluorophore and quencher molecules could affect the stability and conformation of the G-quadruplex. Meanwhile, the fluorescence signal is also temperature and pH dependent which may introduce some bias to the data. Additionally, the interaction between the fluorophore/quencher molecules and ligands may exist. Hence, it is necessary to cross-inspect the FRET assay results using other methods as described below.

### 3.3.4. Aromatic side-chains enhanced the stacking affinity and selectivity of 5-substituted Phen-Ni complexes against telomeric G-quadruplex DNA.

The binding of Phen-Ni to the telomeric G-quadruplex was further investigated by the fluorescent intercalator displacement (FID) assay. Meanwhile, we also investigated the binding of Phen-Ni to duplex DNA using the same method. Molecules as such with a large stacking surface, in general, could also bind to duplex DNA. Because duplex DNA is more
abundant than G-quadruplex DNA in vivo, it is important to study the selective binding of Phen-Ni to G-quadruplex DNA over duplex DNA, which is crucial to develop telomerase-specific inhibitors. In our experiments, thiazole orange (TO) was first bound to G-quadruplex DNA (G1) or duplex DNA (D1), and the pre-bound TO was successively displaced out of DNA by the addition of a ligand. The binding of TO is known to be end stacking and such binding is reversible. Upon binding to DNA, TO fluorescence increases 1000-fold compared to free TO in solution that has a negligible quantum yield. When bound TO is displaced out of DNA by a DNA ligand, a dramatic decrease of fluorescence is detected. The amount of ligand needed to displace 50% of pre-bound TO is defined as \(^{G4}\text{DC}_{50}\) for G-quadruplex and \(^{ds}\text{DC}_{50}\) for duplex DNA. Both values can be used to compare the relative binding affinities of the ligand to G-quadruplex and duplex DNA.\(^{186}\)

11-Ni and 7-Ni exhibited higher \(^{G4}\text{DC}_{50}\) values compared to 1-Ni. 11-Ni gave the worst \(^{G4}\text{DC}_{50}\) value (46.1 \(\mu\text{M}\)) amongst all Phen-Ni compounds probably due to the repulsion between the negatively charged amine oxide and DNA backbones. Glucose is not an effective side chain for G4 ligands as the \(^{G4}\text{DC}_{50}\) value for 7-Ni was also amongst the worst, 29.9 \(\mu\text{M}\). Though the sugar moiety could interact with DNA grooves via H-bonding,\(^{121}\) the linker of 7-Ni may not be long and flexible enough to properly position the glucose into the DNA groove.
The corresponding rank order of Phen-Ni with $^{G4}DC_{50}$ values less than 5 µM was $3$-Ni $< 10$-Ni $< 13$-Ni $< 16$-Ni $< 21$-Ni $< 17$-Ni $< 14$-Ni $< 20$-Ni $< 19$-Ni $< 8$-Ni $< 12$-Ni. This order did not exactly match the order of these complexes from the thermal denaturation experiments with G2 (Table 3-3). For instance, $21$-Ni and $20$-Ni that stabilized G2 the most were only ranked at the 5th and 8th. $3$-Ni showed no significant thermal stabilization of G2 ($T_m = 60.0 \degree C$); however, it exhibited the lowest $^{G4}DC_{50}$ value (0.92 µM). Other discrepancy was also overserved. $6$-Ni showed a very high $^{G4}DC_{50}$ value (30.5 µM) but with a good stabilization of G2 ($T_m = 66.5 \degree C$). Such discrepancy for $3$-Ni and $6$-Ni could result from the interference of fluorophore and quencher with the complexes in solution in fluorescence melting as mentioned in the previous section. 

Another plausible explanation of the observed discrepancy, especially for complexes
containing positively charged side chains such as 21-Ni and 20-Ni is the difference in binding sites. In the melting experiments monitored by fluorescence, the thermal stabilization of G-quadruplex DNA is directly correlated to the binding affinity of the ligand. 21-Ni and 20-Ni gave the highest T_m values of G2 probably because they both contain a positively charged side chain that electrostatically interacted with DNA backbones and further stabilized DNA. On the other hand, such an interaction may cause 21-Ni and 20-Ni not to bind to the exact site as TO binds. Therefore, in the FID assay, the displacement of TO by 21-Ni and 20-Ni is not as efficient as that by other ligands, resulting in slightly higher G⁴DC₅₀ values. Further evidence of this explanation comes from the observation that except for 20-Ni and 21-Ni, all the complexes with low G⁴DC₅₀ values contain an aromatic moiety as a side chain. The proper alignment of the aromatic side chains could enhance their end-stacking with G-quartets and more efficiently displace TO out of G-quadruplex DNA. Without Ni²⁺, 5-substituted phenanthrolines 20 and 21 could not act as potent G₄ ligands because their G⁴DC₅₀ values were drastically large, 83.2 μM for 20 and 72.3 μM for 21. The G⁴CD₅₀ values for 20 and 21 were 28 times and 36 times higher than those of 20-Ni and 21-Ni, respectively (Figure 3-10). The compounds 20 and 21 also showed no stabilization of G2 in the melting experiment monitored by FRET. These observations suggest that the end-stacking interaction is an important factor for the binding of phenanthroline derivatives to G-quadruplex DNA.
For eleven complexes (except for 10-Ni and 16-Ni, reason explained in the next section) with a $\text{G}^4\text{DC}_{50}$ value less than 5 µM, their corresponding $\text{dsDC}_{50}$ values (Figure 3-11) were measured using a DNA duplex (D1), and their selectivity of G-quadruplex over duplex was determined by the ratio of $\text{dsDC}_{50}/\text{G}^4\text{DC}_{50}$. The rank order for such selectivity was $\text{3-Ni} (8.2) > \text{14-Ni} (7.2) > \text{13-Ni} (5.5) > \text{17-Ni} (5.1) > \text{19-Ni} (4.1) > \text{8-Ni} (3.9) > \text{12-Ni} (2.7) > \text{20-Ni} (2.3) > \text{21-Ni} (1.4)$. 20-Ni and 21-Ni containing cationic groups at physiological pH gave the worst selectivity amongst these nine complexes. The $\text{G}^4\text{DC}_{50}$ value for 20-Ni was 2.99 µM, 2.3 fold lower than its $\text{dsDC}_{50}$ (6.94 µM). It is clear that 20-Ni more preferably binds to G-quadruplex than duplex DNA. The $\text{G}^4\text{DC}_{50}$ value of 21-Ni was 2.01 µM, 1.5 fold lower than that of 20-Ni, suggesting that 21-Ni is a slightly stronger G4 ligand. However, the selectivity of 21-Ni to G-quadruplex over duplex was worse than that of 20-Ni because the $\text{G}^4\text{DC}_{50}$ value of 21-Ni was only 1.4-fold lower than its $\text{dsDC}_{50}$.
(2.75 µM). Although the piperidine moiety in 21-Ni is a known effective side chain for G4 ligands, its ability to dictate binding selectivity is strongly pH-dependent. Kerwin and coworkers reported that PIPER (with piperidine as a side chain) had diminished binding selectivity for G-quadruplex and duplex DNA at low pH.\textsuperscript{207} The G\textsubscript{4}DC\textsubscript{50} value was 23.4 µM for 1-Ni, 7.7 fold and 11.6-fold higher than that of 20-Ni and 21-Ni, respectively. This observation clearly reveals that side chains at the 5 position are important for developing such phenanthroline-based G4 ligands.

![Graph showing G\textsubscript{4}DC\textsubscript{50} and dsDC\textsubscript{50} values of the Phen-Ni complexes from the FID assay.](image)

Figure 3-11. The G\textsubscript{4}DC\textsubscript{50} and dsDC\textsubscript{50} values of the Phen-Ni complexes from the FID assay.

Overall, the results obtained from the FID assay were somewhat in agreement with the data from the thermal denaturation experiments. However, a clear discrepancy between G\textsubscript{4}DC\textsubscript{50} and T\textsubscript{m} values could not be neglected. Thermal denaturation monitored
by fluorescence is suitable to capture a coarse picture of the effect of Phen-Ni on G-quadruplex stabilization in a high-throughput manner; however, it cannot provide precise information due to the nature of its system. Indeed, in the thermal denaturation experiments monitored by fluorescence, we were not able to observe a saturation plateau of the melting temperature ($T_m$) even when the concentration of Phen-Ni was increased to a ligand/DNA ratio of 10. Such a saturation plateau was observed when we conducted thermal denaturation of a native G-quadruplex DNA (same sequence as G2) monitored by circular dichroism, suggesting the possible interference amongst fluorescein, DABSYL and Phen-Ni. The FID assay allows us to directly compare the binding affinities of the ligands with similar binding modes semi-quantitatively. It is particularly useful to compare the binding selectivity of the ligands between G-quadruplex and duplex DNA. The results presented above further validates that cross inspection of the fluorescence melting data with other methods is important. It is necessary to point out that the observed discrepancy between the data from fluorescence thermal denaturation and FID could also result from the temperature dependence of affinities in $T_m$. Nevertheless, all the biophysical methods mentioned above clearly indicate that side chains and Ni$^{2+}$ ions are two crucial factors to make 5-substituted phenanthrolines good G4 ligands. A positively charged side chain tethered to the 5-position of phenanthroline is a structural motif for enhancement of the binding affinity G-quadruplex DNA but it has little effect on the selectivity for G-quadruplex over duplex DNA. By contrast, an arylsulfanyl group tethered to the 5-position of phenanthroline is useful to improve both the binding affinity and selectivity of the ligand because it expands the stacking surface of the ligand. The surface of G-quartets is much larger than base pairs; therefore, ligands with expanded
stacking surface bind more tightly to G-quadruplex DNA than duplex DNA.

3.3.5. 5-Substitued Phen-Ni complexes inhibited telomerase activity in vitro.

Upon confirming the binding and stabilization of telomeric G-quadruplex DNA by our novel 5-substituted phen-Ni complexes, we investigated their inhibition of human telomerase activity in vitro using a TRAP (Telomeric Repeat Amplification Protocol) assay. As discussed in the previous sections, telomerase is an enzyme that can bind and extend telomeric DNA and maintain the telomere length. Telomerase is over-expressed in 90% of tumors and helps cancer cells to compensate telomere shortening. It is well acknowledged that stabilized G-quadruplexes formed by the telomeric DNA can efficiently inhibit the telomerase binding and extension activity. Therefore, in the development of G4 ligands and telomerase inhibitors, quantitative analyses of telomerase activity are crucial, and the most frequently used method is the TRAP assay.

Before PCR became an extensively used technique, telomerase activity had been directly studied in vitro by the following method. Radioactively labeled nucleotide 5’-triphosphates were incorporated to its substrate sequence by telomerase. The resulting radioactive products were measured using autoradiography. Due to the low abundance of telomerase in living cells, the quantitation of telomerase activity with this assay was not reliable, even a large amount of radioactive materials were used. To increase the detection sensitivity, Kim and coworkers developed the first PCR-based TRAP assay. It can also be used for the detection of telomerase activity in a high throughput fashion.

A PCR-based TRAP assay can be divided into two main stages: elongation of primers by telomerase and subsequent amplification of telomerase products using PCR.
In the elongation stage, the primer oligonucleotides (TS) are extended with addition of repeat segment (5’-TTAGGG) by telomerase. The pure form of telomerase is not commercially available. The experiments are normally conducted by using cell lysate extracted from telomerase-positive cells such as HeLa cells. In the following PCR amplification stage, these extended products are used as the templates and amplified in the presence of a reverse primer ACX. Utilization of the PCR amplification significantly increases the detection sensitivity; therefore, the products can be detected using low-dose radioactive materials or even conventional fluorescent DNA dyes such as SYBR Green (Figure 3-12).

1. Elongation of the primer

![Diagram of elongation of the primer](image)

2. Amplification of the telomerase products

![Diagram of amplification of the telomerase products](image)

Figure 3-12. Scheme of a TRAP (Telomeric Repeat Amplification Protocol) assay for telomerase activity detection.

In the present work, we explored the feasibility of using two different TRAP
assays. In the first method (real-time PCR TRAP), we conducted the PCR amplification in a real-time PCR machine and monitored the amplified products using SYBR Green. This assay was carried out in a 96-well plate in a high throughput fashion, which in principle can screen hundreds of compounds in a 3-4 hours period. The biggest advantage of this method is free of radioactive materials. In the second method, we used a 5'\textsuperscript{-32}P labeled primer in the PCR amplification and the final products were measured using autoradiography. Both methods will be explicitly discussed in the following section.

**Real-time PCR TRAP.** SYBR Green is a conventional DNA dye. Upon binding to DNA, it fluoresces with an excitation wavelength 497 nm (blue light) and an emission wavelength 520 nm (green light). The intensity of the fluorescence is positively correlated to the amount of the DNA present in solution. During the real-time PCR amplification, the fluorescence of the solution is recorded after each thermal cycle, and the intensities are plotted versus the cycle number. Telomerase-positive cell lysate can extend the primer to yield products that are used as templates in the PCR amplification. The fluorescence intensity of such a sample increases as a function of increasing thermal cycle numbers in PCR. No fluorescence increment can be observed if telomerase-negative cell lysate is used. A threshold of the fluorescence intensity was set as 10 times over the deviation of the background in a DNA free sample. The thermal cycle number is needed to reach the threshold is defined as a Ct value. In the presence of a G4 ligand, telomerase is inhibited so it produces fewer extended products, namely fewer template copies for PCR amplification. Hence, more thermal cycles (Ct) are needed to reach the fluorescence threshold. The Ct value is correlated to the ability of a G4 ligand to inhibits telomerase. The greater the Ct value is, the stronger the inhibition effect is. To determine
this concept, we used a well-studied G4 ligand TMPyP4 as a model compound. The results are shown in Figure 3-13. In the absence of TMPyP4, it requires 25 thermal cycles to reach the threshold. As the concentration of TMPyP4 increases, more thermal cycles are needed to reach the same threshold. For instance, in the presence of 6.25 µM TMPyP4, it requires 34 cycles to reach the threshold. When the concentration of TMPyP4 are larger than 12.5 µM, the threshold cannot be reached even after 41 thermal cycles that is pre-defined limitation number for this method.

![Figure 3-13](image)

Figure 3-13. The results of real-time PCR TRAP concentration-dependent inhibition profile for TMPyP4.

All twenty-one Ni containing complexes and the free phenanthroline derivatives were studied by this real-time PCR TRAP assay (Figure 3-14). In the absence of Ni$^{2+}$, all compounds except for 13 exhibited no telomerase inhibition effect. This observation is
highly consistent with the data from the FRET and FID assays. 5-Substituted phenanthroline derivatives are not good G4 ligands due to the relatively small aromatic surfaces. Interestingly, compound 13 showed a telomerase inhibition effect in this TRAP assay, but it had no G-quadruplex binding ability based on other two assays. Hence, compound 13 was suspected to be a telomerase inhibitor that may directly affect the telomerase without interactions with DNA. This assumption warrants further investigation.

Ni$_2^+$ containing complexes 3-Ni, 8-Ni, 10-Ni, 12-Ni, 13-Ni, 14-Ni, 16-Ni, 17-Ni, 19-Ni, 20-Ni, and 21-Ni had Ct values that were greater than 40 (pre-defined limitation of the thermal cycle numbers), suggesting the inhibition of telomerase activity by these complexes. These results were highly consistent with the previous conclusion from the FID experiments that they are good G4 ligands. However, there are nonnegligible discrepancies between the TRAP and the FRET assay. Complexes 19-Ni, 16-Ni, 3-Ni, 14-Ni, 8-Ni and 12-Ni showed higher G-quadruplex binding affinities (determined by the FID assay) and better telomerase inhibition effects (determined by the TRAP assay) than 1-Ni. However, none of them showed a G-quadruplex stabilization in the FRET assay, compared to 1-Ni. This discrepancy further validates our previous discussion in section 3.3.3 that the FRET assay has certain disadvantages: The G-quadruplex conformation and ligand binding can be affected by fluorescence and quencher molecules; the fluorescence signals can be disturbed by ligands, temperature, and pH conditions. Interestingly, all the complexes that were underestimated in the FRET assay only contain aromatic side chains. Therefore, it is reasonable to conclude that the aromatic moieties at the 5-position of phenanthroline derivatives may have a notable interference with the fluorophore in the
Concentration-dependent studies were performed for the above effective telomerase inhibitors using the real-time PCR TRAP assay. The TRAP reactions were performed in the presence of ligands at various concentrations ranging from 0 to 100 µM, and the corresponding Ct values for each experiment were plotted versus the concentrations.
Clear concentration-dependent telomerase inhibition effects were observed for all the inhibitors (Figure 3-15). To compare their inhibition effects, a telomerase inhibition factor was defined as the integrals of the areas under the curves in Figure 3-15 for each complex. Overall, the rank order of the inhibition effects obtained from the TRAP assay (Figure 3-16) is somewhat in agreement with the data from the FID experiments (Figure 3-9). Complexes 10-Ni, 16-Ni, and 3-Ni exhibited the best inhibition effects in the TRAP assay, and they all showed strong binding affinities in the FID assay. Compound 13, in the absence of Ni$^{2+}$ ion, was a strong telomerase inhibitor without G-quadruplex interactions and addition of Ni$^{2+}$ ion even decreased its inhibition effect. The rest of the inhibitors were all observed with efficient G-quadruplex binding in the FID assay with a different order.
Figure 3-16. Telomerase inhibition factors calculated from the real-time PCR TRAP assay for effective inhibitors.

Although the real-time PCR TRAP assay is a high-throughput protocol to study telomerase inhibitors, it has a major disadvantage. In this assay, the telomerase products produced by telomerase in the first step and subsequently amplified by DNA polymerase and quantified using SYBR Green in the second step. Theoretically, the formation of the PCR products can be inhibited in either the telomerase elongation stage or the DNA replication stage. Ideally, when telomerase adds more than four 5’-TTAGGG-3’ repeats to the TS primers, a G-quadruplex can be induced and stabilized by a G4 ligand in solutions. Then, the G-quadruplexes will inhibit the binding of telomerase and stop the further extensions. This leads to fewer telomerase products as templates in the following PCR process, resulting in a larger number of thermal cycles to reach the threshold. However, the ligands of interest can potentially bind to duplex DNA as well due to the
nature of its stacking surface. Such binding could inhibit DNA polymerase activity in the PCR step. If a ligand binds to duplex DNA not G-quadruplex DNA, it will give a false positive result because more thermal cycles are still needed to reach the threshold. In this scenario, a telomerase inhibitor and a duplex DNA binder will give similar results in the real-time PCR TRAP assay. This possibility could not be neglected especially for some of our complexes such as 20-Ni and 21-Ni that function as both telomerase inhibitors and duplex DNA binders in real-time PCR experiments. To clarify, we further performed an electrophoresis-based 32P TRAP assay to get a more detailed picture of the telomerase inhibition effects of the 5-substituted phenanthroline Ni2+ complexes.

_Electrophoresis-based 32P TRAP assay._ In a radioactive TRAP assay, the TS primers are labeled at the 5’-end with a 32P phosphate group. Thus, when the PCR products are resolved in an acrylamide gel, the extended TS primers with different copies of 5’-TTAGGG-3’ can be separated by electrophoresis and form a ladder. The gel is visualized by autoradiography and the length of the ladders is correlated to the telomerase activities. To determine whether the positive effects resulting from inhibition of telomerase or inhibition of DNA polymerase, we added an internal standard oligonucleotide TSNT and its reverse primer NT into the reactions. TSNT is a non-G-rich oligonucleotide that cannot be extended by telomerase. It can be replicated by DNA polymerase using TS as the forward primer and NT as the reverse primer. Using this modified method, several scenarios could exist. If a ligand binds to G-quadruplex DNA but does not affect DNA polymerase, we should observe the PCR products of TSNT and no elongated products from telomerase extension. If the ligand has no effect on both telomerase and DNA polymerase, we should observe both TSNT products from PCR and
elongated products by telomerase. If the ligand only inhibits DNA polymerase activity, we should observe elongated products from telomerase and no TSNT products from PCR. Therefore, the presence of the TSNT products can report the ligands selectivity against G-quadruplex (telomerase) over duplex DNA (polymerase).

We first analyzed all the complexes with the $^{32}$P-TRAP at a relatively high concentration (50 µM). The internal standard TSNT is a 36 nt (nucleotide) product (IS, Figure 3-17) that has a higher mobility than the products elongated by telomerase on the gel. The length of telomerase products starts from 50 nt and they form a ladder with a 6 nt gap between two adjacent bands (Figure 3-17). Regardless the hybridization position of the primer (RP) on TS, the first three telomerase products are all 50 nt long (Figure 3-17). Thus, the band representing the 50 nt long product has a high intensity on the gel. In the presence of a G4 ligand, the formation of the first three telomerase products will not be inhibited, because the products are not long enough to form a G-quadruplex. Therefore, this band can always be observed on the gel regardless the presence of the ligands.

![Figure 3-17. Illustration of the telomerase-extended products in a TRAP assay.](image-url)
We employed two well studied small molecules BIBR and TMPyP4 (50 µM and 100 µM) as the control compounds. BIBR is a noncompetitive telomerase inhibitor without interactions with DNA.\textsuperscript{210} TMPyP4 is a well-studied DNA ligand that binds to both G-quadruplexes and duplex DNA.\textsuperscript{211-212} In the resulted gel, a telomerase positive control lane (lane 1 in Figure 3-18) gave a clear ladder for telomerase products with different lengths and an internal standard band (IS) located at the bottom. A telomerase negative control lane (lane 14 and 28) only showed an internal standard band. In the presence of BIBR, no telomerase products ladder can be found, and the intensity of the internal standard was not affected suggesting that it only inhibits telomerase not DNA replication (lane 2). When TMPyP4 was added, both telomerase products ladder and the internal standard disappeared suggesting the inhibition of both telomerase and DNA polymerase activities (lane 3). The twenty-one complexes were studied with this TRAP assay at a concentration of 50 µM, and the results are highly consistent with the data from previous experiments (lane 5 to 27). Complexes \textbf{3-Ni, 8-Ni, 12-Ni, 13-Ni, 14-Ni, 17-Ni, 19-Ni, 20-Ni,} and \textbf{21-Ni} inhibited telomerase activity with various selectivity. Complexes \textbf{3-Ni, 8-Ni, 17-Ni, 21-Ni,} and \textbf{20-Ni} had strong inhibition effects at this concentration because both the telomerase elongated products and internal standards were completely missing. For these complexes, it is necessary to perform a concentration-dependent study because when the duplex amplification was completely inhibited, it was not clear if the telomerase extension was inhibited or not. Complexes \textbf{13-Ni} and \textbf{14-Ni} gave a better selectivity at the same concentration than the others, and the internal standards were not affected. Meanwhile, the effect of \textbf{10-Ni} and \textbf{16-Ni} was not able to be analyzed using electrophoresis. Their PCR products remained in the loading wells of the acrylamide gels.
probably due to DNA condensation by these two complexes. The reason of this condensation remained unclear to us at this point.

![Figure 3-18](image.png)

Figure 3-18. $^{32}$P TRAP reactions for all Phen-Ni complexes (50 µM). Reactions in the absence or presence of Ni$^{2+}$ (50 µM), BIBR (50 µM), and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template). The reaction in the right lane has no telomerase.

The results of concentration-dependent experiments for 13 and 13-Ni are shown in Figure 3-19. Inhibition of telomerase by 13 was observed when its concentration was larger than 25 µM. No inhibition of DNA polymerase activity by 13 was observed even at the concentration of 50 µM. When 13 was coordinated with Ni$^{2+}$ ions, the resulting 13-Ni showed significantly decreased telomerase inhibition effect. The inhibition of telomerase activity was barely observed when 30 µM of 13-Ni was used. By contrast, no inhibition
of DNA polymerase by 13-Ni was observed up to 50 µM. All these results are consistent with the data from the real-time PCR TRAP assay. It is reasonable for us to conclude the 13 may inhibit telomerase activity via the direct interaction with the protein, not the G-quadruplex formation approach. The detailed mechanism of action is still unclear. It is noteworthy that the complex 13-Cu (50 µM) formed by mixing 13 and Cu$^{2+}$ in a 1:2 ratio also showed the inhibition effect of telomerase activity.

Figure 3-19. $^{32}$P TRAP reactions in the presence of 13 (left) at concentrations of 50, 25, 12.5, 6.25 and 3.125 µM (from left to right). $^{32}$P TRAP reactions in the presence of 13-Ni (right) at concentrations of 50, 40, 30, 20, 15 and 10 µM (from left to right). Reactions in the presence of 13-Cu$^{2+}$ (50 µM), Ni$^{2+}$ (50 µM), BIBR (50 µM), and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).
The complex 14-Ni also exhibited a significant telomerase inhibition effect at 25 µM (Figure 3-20). No inhibition of DNA polymerase activity was observed in the presence of 50 µM 14-Ni, suggesting selective inhibition of telomerase over DNA polymerase could be achieved under the concentration range used. In the presence of 50 µM 14, the inhibition of both telomerase and DNA polymerase was not observed. Unlike 13-Cu, the complex 14-Cu (50 µM) could not inhibit telomerase activity.

Figure 3-20. \(^{32}\text{P}\) TRAP reactions in the presence of 14-Ni at concentrations of 50, 25, 12.5, 6.25 and 3.125 µM (from left to right). Reactions in the presence of 14 (50 µM), 14-Cu\(^{2+}\) (50 µM), Ni\(^{2+}\) (50 µM) and BIBR (50 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).
In the presence of 50 µM 8-Ni, the inhibition of DNA polymerase and telomerase activity was observed. Selective inhibition of telomerase over DNA polymerase was detected in the presence of 45 µM 8-Ni. The minimal telomerase inhibitory concentration of 8-Ni was 25 µM. As a comparison, compound 8 did not exhibit any telomerase and DNA polymerase inhibition at 50 µM.

Figure 3-21. 32P TRAP reactions in the presence of 8-Ni at concentrations of 50, 45, 40, 35, 30, 25 and 20 µM (from left to right). Reactions in the presence of 8 (50 µM), BIBR (50 µM) and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).
In the presence of 30 µM 17-Ni, the inhibition of both DNA polymerase and telomerase activity was observed (Figure 3-22). Selective inhibition of telomerase over DNA polymerase was detected in the presence of 25 µM 17-Ni. The minimal telomerase inhibitory concentration of 17-Ni was 15 µM. As a comparison, compound 17 at 50 µM only exhibited limited inhibition of telomerase activity with no inhibitory effect on DNA polymerase.

Figure 3-22. \(^{32}\)P TRAP reactions in the presence of 17-Ni at concentrations of 30, 25, 20, 15 and 10 µM (from left to right). Reactions in the presence of 17 (50 µM), BIBR (50 µM) and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).
The results of 3-Ni were similar to all the complexes mentioned above except that it has worse inhibitory selectivity for telomerase over DNA polymerase (Figure 3-23). At the concentration of 30, 40 and 50 μM 3-Ni, the inhibition of both enzymes was observed. In the presence of 20 μM 3-Ni, the inhibition of telomerase activity was detected while no inhibition effect on DNA polymerase was observed. 3-Ni at 10 μM showed no inhibition effect on both enzymes, suggesting a very narrow concentration range for selectivity inhibition of telomerase over duplex DNA polymerase.

![Figure 3-23. $^{32}$P TRAP reactions in the presence of 3-Ni at concentrations of 50,40,30,20,10 and 5 μM (from left to right). Reactions in the presence of 3 (50 μM), BIBR (50 μM) and TMPyP4 (100 μM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).]
All the Phen-Ni mentioned contain an aromatic moiety (arylsulfanyl group) at 5-position of phenanthroline. In this section, we studied two complexes **20-Ni** and **21-Ni** that contain a positively charged side chain. At the highest concentration (50 µM), inhibition of both telomerase and polymerase was observed (Figure 3-24). Between 6.25 and 25 µM, the telomerase extension product bands gradually decreased as a function of increasing ligand concentration while the intensity of the PCR product remained unchanged, a strong evidence of selective inhibition of telomerase over polymerase by these two complexes. No telomerase inhibition could be detected at 3.125 µM, the lowest ligand concentration used. Interestingly, when the TRAP reactions were conducted with either **20** or **21** with Cu²⁺, in a 2:1 ratio, no inhibition effect could be observed. Phen and Cu²⁺ most likely form a geometry intermediate between the trigonal bipyramid and square pyramid²¹³ which stacks to G-quadruplex DNA less efficiently. Several complexes formed by coordination of phenanthroline and Cu²⁺ ions have been used as G4 ligands. However, they were modified by attaching side chains at the inner crescent side of phenanthroline. These side chains are closer to the chelation sites of phenanthroline can significantly affect the chelation mode with Cu²⁺ ions."
Figure 3-24. $^{32}$P TRAP reactions in the presence of 20-Ni (left) and 21-Ni (right) at concentrations of 50, 25, 12.5, 6.25 and 3.125 µM (from left to right). Reactions in the presence of 20 or 21 (50 µM), 20-Cu$^{2+}$ or 21-Cu$^{2+}$ (50 µM), Ni$^{2+}$ (50 µM), BIBR (50 µM) and TMPyP4 (100 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).

It is noteworthy that the concentration of Phen-Ni to inhibit the telomerase activity reported here were comparable with other known G4 ligands such as PIPER$^{214-215}$ and phenanthroline derivatives$^{216}$. Our TRAP results were in close agreement with the data from the FID assay-most of the complexes with a low $G_4$DC$_{50}$ value showed a good telomerase inhibition effect, implying that the observed telomerase inhibition by Phen-Ni could result from their binding to/inducing G-quadruplex structures.

3.3.6. 5-Substituted Phen-Ni complexes inhibited telomerase positive cancer
**cell growth.** By inhibiting the telomerase-mediated telomere maintenance, many G4 ligands showed cytotoxicity against telomerase-positive cancer cell lines.\textsuperscript{141} In this section, we investigated the cytotoxicity of Phen-Ni against two telomerase-positive cancer cell lines including HeLa (cervical cancer) and MCF-7 (breast cancer) using the MTT assay.\textsuperscript{217-218}

In a typical experiment, cancer cells in a growth media were incubated with the Phen-Ni complex at 100 μM for 72h at 37 °C. The reaction mixture was treated with MTT (3-[4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide) and the cell viability was determined by measuring the absorbance at 520 nm using a microplate reader. MTT can be reduced to the colorful formazan by a reductase in live cells. The number of live cells in the media is directly correlated to the absorbance.

Four complexes (3-Ni, 4-Ni, 20-Ni, and 21-Ni) significantly inhibited the proliferation of both HeLa and MCF-7 cells and two complexes (10-Ni and 14-Ni) only inhibited the growth of MCF-7 cells. Except for 14-Ni, all the complexes with detectable cytotoxicity were proved to be telomerase inhibitors in the TRAP experiments. 3-Ni, with a benzene ring as the side chain, exhibited the greatest inhibition amongst them. Though 3-Ni was the best G4 ligand, it showed moderate inhibition effect on telomerase activity. The direct correlation between cell viability and telomerase inhibition could not be fully established. It is plausible that the cytotoxicity of 3-Ni makes it readily permeable through the lipid bilayer in cell membranes.\textsuperscript{219} The accumulation of 3-Ni with relatively high concentration in cells accounts for its cytotoxicity. By contrast, 20-Ni and 21-Ni contain cationic side chains that should have low permeability. In our experiments, though 20-Ni and 21-Ni have the strongest telomerase inhibition, they both exhibit
moderate cytotoxicity toward two cancer cell lines used (Figure 3-25).

Figure 3-25. Cell proliferation of two cancer cell lines in the presence of 5-substituted Phen-Ni complexes, determined by an MTT assay.

To test if the inhibition effect came from the G-quadruplex stabilization and telomerase inhibition, the corresponding free compounds in the absence of Ni$^{2+}$ ions were studied on HeLa cells (Figure 3-26).
Interestingly, only complex 3-Ni exhibited significantly improved cell cytotoxicity compared to the free compound. This further supported that 3-Ni was able to be diffused into the cells and exhibited its inhibition effects. Since 3-Ni exhibited the inhibition of internal standards in the TRAP assay, the cytotoxicity could also be, at least partially, attributed to replication inhibition. The further studies and development of these complexes in vivo must be combined with drug delivery vehicles that can efficiently transport them into cells and cell nuclei.

3.4 Conclusion

We investigated a series of 5-substituted 1,10-phenanthroline Ni\textsuperscript{2+} complexes as telomeric G-quadruplex DNA ligands. Their binding affinity and selectivity were highly
depending on the type of their side chains. Both cationic side chains and aromatic side chains can enhance the G-quadruplex binding affinity of the ligands. Compared to those cationic diethyl amines and piperidine, arylsulfanyl side chains improved the selectivity of the ligands against G-quadruplex over dsDNA, which could be used as a better side chain motif for selective G4 ligands.

Meanwhile, we evaluated several biophysical and biochemical methods used for studying G4 ligands. Thermal denaturation monitored by fluorescence can be used as a high throughput method for screening G4 ligands from a large library; however, some deficiencies must be taken into consideration due to the interferences of fluorophores on DNA and DNA ligands. In contrast, fluorescent intercalator displacement (FID) assay is a useful method for semi-quantitatively determine the binding affinity of G4 ligands. This method is especially useful if the ligand of interest binding to the same location as TO. It also provides a meaning way to directly measure the selectivity of ligands for G-quadruplex DNA over duplex DNA. Telomeric repeat amplification protocol (TRAP) is highly useful for telomerase inhibitor studies. Real-time PCR TRAP is a high throughput assay that can evaluate telomerase activity. However, for G-quadruplex stabilizing telomerase inhibitor studies, real-time PCR TRAP is not able to differentiate the signals caused by ligand binding to duplex DNA. For this reason, $^{32}$P-TRAP is a reliable method for G4 ligand studies, in which the duplex DNA binding can be monitored by introducing an internal standard. For minimal usage of radioactive materials and reliability, we consider a combination of both methods is ideal for studying G4 ligands.
Chapter 4. Development of Thiazole Orange Derived G-quadruplex Ligands and Telomerase Inhibitors.

Thiazole Orange (TO) is a DNA dye that is frequently used to stain DNA in living cells. The aromatic rings of TO enable it to intercalate to nucleobase pairs and light-up the DNA as a fluorescence probe. The TO-DNA binding is a reversible process and the TO can be displaced off from the DNA by other strong DNA ligands. This behavior is often utilized in a Fluorescent Intercalator Displacement (FID) assay to evaluate the DNA-ligands binding affinity. In recent years, TO attracts more attention because of its G-quadruplex DNA binding property. G-quadruplex is a secondary DNA structure that is closely related to the telomeres maintenance and oncogene regulations in human genome. TO stacks on G-quadruplexes with a binding affinity that is similar to its duplex DNA binding. In order to develop novel G4 ligands with a good binding affinity and selectivity, our group synthesized twelve novel TO derived compounds (CP1-12). Twelve positively charged side chains were chosen and tethered to the quinoline ring of TO to enhance the G-quadruplex binding affinity by electrostatic interactions with DNA backbones. In the present work, several TO derivatives exhibited better G-quadruplex stabilization than TO, determined by biophysical studies including Circular Dichroism (CD) thermal denaturation, fluorometric titration, and ESI mass spectrometry. Polyamine and piperidine/pyrrolidine side chains with suitable chain length significantly improved the G-quadruplex binding affinity of TO. In a TRAP assay, CP6, with a triply protonated
polyamine side chain, exhibited a significant telomerase inhibition effect (nanomolar), with is comparable with the best reported synthetic G4 ligand BRACO-19.

4.1 Introduction

As discussed in the previous sections, G-quadruplex is a DNA secondary structure that is extensively present in DNA telomeric regions and oncogene promotor regions. Because of its crucial roles in telomere maintenance and oncogene expressions, many small molecules have been designed to stabilize G-quadruplexes. All successful G-quadruple ligands share two features: aromatic surfaces that stack onto the G-quartet and cationic sites that bind onto the phosphate backbones (DNA grooves). Thiazole Orange (TO) is one of the molecules that meet these requirements (Figure 4-1).

As discussed in Chapter 2, TO is a DNA dye that binds to both G-quadruplex and duplex DNA. TO has a similar binding affinity with single G-quadruplex structure and duplex DNA. It has been reported that TO selectively binds to consecutive G-quadruplex structures in long G-rich sequences with 700-2000 nucleotides over duplex DNA.\textsuperscript{192} Therefore, TO is an interesting scaffold for developing novel G4 ligands.
Most of the well-studied G4 ligands such as PIPER or BRACO-19 have long cationic side chains that efficiently interact with DNA backbones. In a TO molecule, there is only one positive charge, and it is located at the aromatic surface that stacking on the G-quartets. Thus, TO does not provide sufficient interactions with DNA backbones. As a result, the G-quadruplex binding of TO is relatively weak. In the present research, we investigated TO derivatives with positively charged side chains, aiming to enhance the G-quadruplex binding affinity.

Twelve TO derivatives with a range of side chains were synthesized in the earlier work of our group (Figure 4-2). The side chains were carefully chosen to study the effects of the factors including side chain length, positive charge density and chain flexibility. Accordingly, these side chains can be grouped into three categories: aliphatic amines, polyamines, and polyethylene glycol amines. The derivatives (CP1, CP2, and CP3) containing aliphatic amino side chains with two, three or six carbons were used to study the G4 stabilization effects depending on the distance between the positively charged sites and TO. The derivatives (CP4, CP5, and CP6) containing polyamine with different number of protonated sites were used to evaluate the charge density effect on the G-quadruplex binding. The derivatives (CP7 and CP8) containing polyethylene glycol side chains were designed to investigate the effect of side-chain flexibility on G4 binding. Lastly, the derivatives containing piperidine/pyrrolidine side chains (CP9, CP10, CP11, and CP12) were also studied because these side chains were used in many successful G4 ligands including PIPER and BRACO-19.
These side chains were tethered onto the nitrogen atom in the quinolinium ring with an amide linkage. A molecular docking study showed that, when TO stacks on the G-quartets, the methyl group on this location points out from the G-quartets to the groove. \(^{220}\) It is reasonable to conclude that a long side chain tethered to this nitrogen can readily interact with the DNA grooves.

In another study, a homodimeric TO molecule (TOTO) was prepared by bridging two TO molecules with a cationic chain on the designated nitrogen atoms to increase the duplex DNA binding affinity (Figure 4-3). Based on the data of an NMR study, TOTO adopted a bis-intercalation mode on a 5'-CTAG-3'/3'-CTAG-5' sequence. The benzothiazole moiety is sandwiched between two pyrimidines (CT), and the quinolinium between two purines (AG). The N-methyl group of the benzothiazole ring protrudes the
major groove, and the cationic linker lies in the minor groove crossing from edge to edge with a perfect matching size (Figure 4-3). This duplex DNA binding is greatly enhanced by the doubled aromatic stacking size of TOTO and the Columbic attraction between the side chain linker and the DNA phosphate backbones.\textsuperscript{221} Therefore, the nitrogen atom in the quinolinium ring was chosen as a suitable site to install the side chains in our present research.

Figure 4-3. The structure of TOTO and an illustration of bis-intercalation of TOTO. \textsuperscript{221}

In the present work, the G-quadruplex stabilization effects of these TO derivatives were studied using a thermal denaturation monitored by Circular Dichroism. \textbf{CP6}, with a polyamine side chain, and \textbf{CP9}, with a pyrrolidine side chain, exhibited the
best G4 stabilization amongst all the compounds tested. The binding affinities of these compounds toward a human telomeric G-quadruplex DNA were determined using fluorometric titration experiments. Results from ESI mass spectrometry suggested that, similar to TO, the TO derivatives produced 1-1, 1-2 and 1-3 complexes with a telomeric G-quadruplex. In addition, we evaluated their inhibition of telomerase activity in vitro using the telomerase repeat amplification protocol (TRAP) assay. CP6 exhibited a remarkable telomerase inhibition effect which is comparable to the most advanced synthetic G4 ligand BRACO-19. Thus, a polyamine side chain with a suitable length could be a potential structural motif for developing more potent G4 ligands and telomerase inhibitors.

4.2 Experimental

4.2.1 Materials and general methods. All chemicals for synthesis were purchased from Sigma-Aldrich or Fisher Scientific and used without further purification. All the enzymes were purchased from New England Biolabs. Oligonucleotides were synthesized on an Applied Biosciences 392 DNA/RNA synthesizer using standard phosphoramidite protocols and purified using polyacrylamide gel electrophoresis. All the chemicals and phosphoramidites for DNA synthesis were purchased from Glen Research (Sterling, VA). The concentrations of oligonucleotide solutions were determined by a Varian Cary 100 Bio UV-Vis spectrophotometer (Walnut Creek, CA), using the molar extinction coefficients (in units of mol of nucleotide/L⁻¹·cm⁻¹) obtained from Oligo Analyzer 3.1 (www.idtdna.com). DNA labeling was performed by incubating [γ-³²P]-ATP (300 µCi) and T4 polynucleotide kinase (20 units) in the presence of an oligonucleotide (1 µg) at 37 °C for 45 min. Unreacted [γ-³²P]-ATP was removed using a MicroSpin G-25
column (GE Health care). [$γ^{32}P$]-ATP was purchased from PerkinElmer (Waltham, MA). Quantification of 5′ $^{32}P$-labeled oligonucleotides was carried out using a Storm 860 phosphor imager and ImageQuant 5.1 software (Molecular Dynamics). Cell culture media and supplements (fetal bovin serum, L-glutamine, and antibiotic-antimycotic) were purchased from Invitrogen.

4.2.2. Circular dichroism G-quadruplex denaturation analysis. A stock solution containing 100 µM human telomeric oligonucleotides PCG1 (5′-AGG GTT AGG GTT AGG GTT AGG GT-3′), 20 mM lithium cacodylate buffer (pH=7.0) and 150 mM NaCl was prepared and incubated at 95°C for 5 min, slowly cooled to room temperature and stored at 4 °C for overnight to form the corresponding G-quadruplex structure.

CD thermal denaturation spectra of solutions containing 5 µM PCG1, 20 µM TO derivatives, 20 mM lithium cacodylate buffer (pH=7.0) and 150 mM NaCl were collected at 290 nm as a function of temperature (25-95 °C) at a heating rate of 0.5 °C/min with a Jasco J-810 spectropolarimeter. The data in CD mdeg units were converted into molar ellipticity using the Spectra Manager (Jasco). Melting profiles were smoothed and analyzed using Originlab and the temperature with the lowest first derivative was determined as the melting temperature ($T_m$).

4.2.3. Fluorometric scan of TO derivatives interacting with G-quadruplex. A stock solution containing 100 µM human telomeric G-quadruplex sequence PCG1 (5′-AGG GTT AGG GTT AGG GTT AGG GT-3′), 10 mM lithium cacodylate buffer (pH=7.0) and 50 mM KCl was prepared and incubated at 95 °C for 5 min, slowly cooled
to room temperature and stored at 4 ˚C for overnight.

Reaction solutions containing 0.25 μM PCG1, 10 mM lithium cacodylate buffer (pH=7.0) and 50 mM KCl were prepared from the stock solution before use. This solution was incubated with 0.5 μM TO derivatives for 30 min in the dark. Spectra (excitation: 501 nm, slit width: 5.0 nm, scan speed 120 nm/min) were recorded in the range of 510-650 nm.

4.2.4. Fluorometric titration of TO derivatives interacting with G-quadruplex and duplex DNA. Stock solutions containing 200 μM human telomeric G-quadruplex sequence PCG1 (5’-AGG GTT AGG GTT AGG GTT AGG GT-3’)) or 100 μM 17 mer dsDNA S1 (5’-GCCAGTTCTAGTAAACC-3’/3’-GGTTACTACGAACTGGC-5’), 10 mM lithium cacodylate buffer (pH=7.0) and 100 mM NaCl were prepared and incubated at 90 ˚C for 5 min, slowly cooled to room temperature and stored at 4 ˚C for overnight. A solution containing 0.5 µM TO derivatives, 10 mM sodium cacodylate buffer (pH=7.0) and 100 mM NaCl was prepared and aliquots of the stock solution of PCG-1 or 17 mer duplex DNA were gradually added. The spectra of fluorescence (excitation: 501 nm, slit width: 5.0 nm, scan speed 120 nm/min) were recorded from 510-650 nm after each addition. The maximum fluorescence intensities were plotted against the equivalent concentrations of the DNA and the binding parameters were extrapolated using a software Prism 7.0.

4.2.5. Mass Spectrometry. A human telomeric DNA sequence PCG1 (5’-AGG GTT AGG GTT AGG GTT AGG GT-3’) and two complementary DNA sequences S1(5’-AGC GAT GAG AGG CCA CGA GGA ATC GCT GGT ACC G-3’/ 5’-CGG TAC CAG
CGA TTC CTC GTG GCC TCT CAT CGC T -3’
were heated to 95 °C for 5 mins
followed by cooling slowly to room temperature and stored in 4°C for overnight in 150
mM NH₄OAc solution. TO derivatives (250 µM) were mixed with the DNA (50 µM) and
the resulting complex solutions were injected into a Thermo Fisher Orbitrap Fusion
Tribrid Mass Spectrometer (San Jose, CA) with 50% methanol-water at a rate of 5
µl/min. The ion spray voltage was set to -2,250 V and ion transfer tube temperature to
300 °C. Nitrogen sheath and auxiliary gas flows were 45 and 0 arbitrary units,
respectively. The collision energy was adjusted to 5 or 10 arbitrary unit (au). MS spectra
were recorded in the m/z range from 1400 to 1800 with an average of 150 scans. Data
were collected and analyzed using the Xcalibur software.

4.2.6. TRAP assay. Cell lysates containing telomerase were extracted from
cultured HeLa cells. The cells from a 90% confluent 225 cm² culture flask were washed
with cold phosphate-buffered saline(PBS) and scraped off from the culture flask in 1 ml
cold PBS. The cells were pelleted in 1.5 ml conical tubes by spinning for 5 min at
1000rpm and suspended in 1ml cold 1× CHAPS Lysis Buffer. The suspension was
incubated on ice for 30mins and centrifuged at 12,000g for 10 minutes at 4°C. The
supernatant was extracted and stored at -80°C.

A reaction mixture (50 µL) was prepared containing 5’-[³²P]-labeled TS primer
(0.1 µg 5’-AATCCGTCGAGCAGAGTT-3’), RP (0.1 µg, 5’-
GCGCGGCTTACCCTACCCTACCCAACC-3’), NT (0.1 µg, 5’-
ATCGCCTTCTCGGCTTTTT-3’), TSNT (0.01 amol, 5’-
AATCCGTCGAGCAGTTAAAAGGCGAGAAGCGAT-3’), buffer: [Tris-HCl (20
mM, pH 8.3), MgCl₂ (1.5 mM), KCl (63 mM), 0.05% Tween 20, EGTA (1 mM)], cell
lysate (1 μL), Taq polymerase (2 units), dNTP mix (50 μM), and TO derivatives (3.125-100 μM). This solution was incubated at 30°C for 30 min, 94°C for 2 min followed by a 3-step PCR method (94 °C for 20 s, 59 °C for 30 s, and 72 °C for 30 s) with 25 cycles. The reaction solution was mixed with 10 μL 10× loading dye and resolved with 10% denaturing polyacrylamide gel [5% crosslink, 45% urea w/w].

4.2.7. MTT cell viability. MCF-7 cells (human breast adenocarcinoma cells) and HeLa cells (cervical cancer cells) (a generous gift from Dr. William. K. Chan, UOP) were maintained in advanced DMEM medium solutions supplemented with 5% fetal bovine serum, L-glutamine and antimycotic at 37 °C in a humid atmosphere containing 5% CO₂.

The TO derivatives were pre-dissolved in sterilized water. The cells were seeded in a 96-well plate (3000 cells/well). After 24h incubation, the supernatant was removed followed by adding 200 μl FBS free media containing TO derivatives. After 72h incubation, 20 μl MTT (3-(4,5-dimethylthiazole-2-yl)-2,5-di- phenyl tetrazolium bromide) (0.5 mg/ml) was added for further 4h incubation. The supernatant was removed, and the formazan product was dissolved in 150 μl of DMSO before measurement. The absorbance was measured at 590nm using a TriStar LB 941 Multimode Microplate Reader. Untreated cells and media without cells were used as controls. To calculate the viability (%), the following equation was used:

$$\text{Viability (\%)} = \frac{(A_{\text{treated}} - A_{\text{media}})}{(A_{\text{untreated}} - A_{\text{media}})} \times 100\%$$

$A_{\text{treated}}$: The absorbance of the solution containing treated cells

$A_{\text{media}}$: The absorbance of the media
4.3 Results and Discussion

4.3.1. Thiazole orange derivatives stabilized human telomeric G-quadruplexes. In the previous studies of our group, all twelve derivatives were examined by a thermal denaturation experiment monitored by fluorescence. CP6 containing a polyamine side chain and CP9 containing a pyrrolidine side chain exhibited strong G4 stabilization effect. At the ratio of 10:1 (ligand:DNA), CP6 and CP9 elevated the melting temperature (T_m) of the human telomeric G-quadruplex with 8.60 °C and 8.85 °C, respectively. However, there were some concerns with this assay. A saturation of the T_m increment induced by the compounds was never reached even in the presence of highly concentrated compounds. As discussed in chapter 3, thermal denaturing monitored by Fluorescence has certain drawbacks due to the interferences between the fluorescence molecules and the tested compounds. Specifically, TO and TO derivatives have an excitation/emission spectra at 501 nm/510 nm which is partially overlapping with the spectra of the fluorophore probe in the thermal denaturation experiment (492 nm/517 nm). Thus, to obtain more reliable data, we conducted a thermal denaturation monitored by Circular Dichroism (CD) in the present research.

Circular Dichroism (CD) spectrometry is a common technique to analyze DNA G-quadruplex conformations. Different G-quadruplex structures can be distinguished from the distinct features of their CD spectra. The conformation of the human telomeric G-quadruplex DNA (G4, 5’-AGGGTAGGGTTAGGGTTAGGG-3’) is cation dependent. It forms two distinguished conformations in the presence of Na⁺ and K⁺ ions respectively (Figure 4-4). In the presence of K⁺ ions, a hybrid-type telomeric G-
quadruplex is observed with a positive peak around 290 nm and a negative peak around 240 nm.\textsuperscript{58,169} When Na\textsuperscript{+} ion is used, a positive peak around 290 nm and a negative peak around 265 nm are observed indicating the presence of a basket-type G-quadruplex.\textsuperscript{58} Under both conditions, the positive peak at 290 nm represents the G-quadruplex structures, and its intensity will decrease when a G-quadruplex dissociates into a random coil (Figure 4-4). Thus, thermal denaturation of G-quadruplex DNA can be conducted by monitoring the CD ellipticity at 290 nm. The melting temperature (T\textsubscript{m}) is defined as the temperature, at which 50\% of the G-quadruplexes dissociate into random coils. It can be extrapolated by the first derivatives from the plots of the molar CD ellipticity against the temperatures.

CD thermal denaturation experiments in the presence of twelve derivatives were conducted under two conditions: 100 mM K\textsuperscript{+} ions and 150 mM Na\textsuperscript{+} ions. These are the most frequently used conditions in G-quadruplex related studies because they produce reliable data. The physiological intracellular concentrations for these two ions in a normal cell are approximately 14 mM for Na\textsuperscript{+} and 150 mM for K\textsuperscript{+}.\textsuperscript{223} In cancer cells, a 30\% decrement of K\textsuperscript{+} ion concentration and a 30\% elevation of Na\textsuperscript{+} ion concentrations compared to those in normal cells are observed, which are closely related to carcinogenesis and related signaling pathways.\textsuperscript{224} The concentration elevation of these ions in cancer cells could directly regulate the G-quadruplex formation and are considered as a sign of aggressive cancer cells.\textsuperscript{225-226} Thus, even not exactly same as the intracellular concentrations in cancer cells, the conditions in our experiments are still suitable for G4 ligands studies.
Figure 4-4. CD spectra of G4 in the presence of 150 mM Na\(^+\), 100 mM K\(^+\), and in the absence of cation.

Under the designated conditions, CD thermal denaturation experiments were conducted in the presence of twelve derivatives, together with TO and TO acid in the ratio of 1:4 (DNA:ligand). TO acid was tested here because it was the precursor used to tether the side chains (Figure 4-1). The T\(_m\) values were extrapolated from the thermal denaturing profiles (Figure A-10) and plotted in the bar graphs (Figure 4-5).
Figure 4-5. G-quadruplex $T_m$ values determined by the CD thermal denaturation experiments in the presence of TO derivatives.

Table 4-1. $T_m$ values (°C) of G4 in the presence of TO derivatives.

<table>
<thead>
<tr>
<th></th>
<th>Blank</th>
<th>CP1</th>
<th>CP2</th>
<th>CP3</th>
<th>CP4</th>
<th>CP5</th>
<th>CP6</th>
<th>CP7</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 mM K$^+$</td>
<td>52.60</td>
<td>59.24</td>
<td>60.23</td>
<td>62.28</td>
<td>60.48</td>
<td>62.44</td>
<td>71.02</td>
<td>58.68</td>
</tr>
<tr>
<td>150 mM Na$^+$</td>
<td>57.54</td>
<td>57.53</td>
<td>56.53</td>
<td>60.03</td>
<td>61.53</td>
<td>58.54</td>
<td>65.04</td>
<td>56.03</td>
</tr>
<tr>
<td>CP8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TO acid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 mM K$^+$</td>
<td>59.41</td>
<td>69.44</td>
<td>64.03</td>
<td>59.67</td>
<td>64.88</td>
<td>63.99</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>150 mM Na$^+$</td>
<td>57.03</td>
<td>62.53</td>
<td>56.04</td>
<td>60.52</td>
<td>60.53</td>
<td>60.03</td>
<td>57.05</td>
<td></td>
</tr>
</tbody>
</table>

In the absence of any ligands (blank), the $T_m$ value in 150 mM Na$^+$ solution (57.54 °C) was significantly higher than that in 100 mM K$^+$ solution (52.60 °C). It is well-known that Na$^+$ and K$^+$ ions have similar stabilization effects on G-quadruplexes and a higher cation concentration can elevate the $T_m$ of G-quadruplexes.$^{227}$
For the same ligand, the $T_m$ increments in 100 mM K$^+$ solution were always greater than that in 150 mM Na$^+$. All the compounds stabilized the G-quadruplexes (produced a higher $T_m$ than the blank) in 100 mM K$^+$ and the highest $T_m$ was acquired by CP6 (71.02°C). In 150 mM Na$^+$, only seven compounds stabilized the G-quadruplexes (CP3, CP4, CP5, CP6, CP9, CP11, CP12, and TO), and the highest $T_m$ was 65.04 °C produced by CP6. As discussed in previous chapters, cations sit between the G-quartets and coordinate with the guanines. Meanwhile, cations also form a hydrated spine in the DNA grooves. Thus, the highly concentrated salts in 150 mM Na$^+$ solution will compete with the ligand binding. This observation was consistent with many reports that highly salted conditions significantly reduce the DNA-ligand binding.\textsuperscript{192}

Regardless of the different extent of $T_m$ values, two conditions produced similar trends of G-quadruplex stabilization for the compounds. Lower stabilization effects were observed for aliphatic amine side chain containing derivatives CP1 (59.24 °C), CP2 (60.23 °C), and CP3 (62.28 °C) than TO (63.99 °C) in 100 mM K$^+$ and nearly no stabilization effect in 150 mM Na$^+$. Aliphatic chains sometimes can provide hydrophobic interactions with DNA in some reported ligands. However, in these TO derivatives, no improvement was observed. This may occur due to the geometry of the side chains which are not efficiently fitting with the G-quadruplex structure and consequently provides no additional interactions. CP3, with the longest aliphatic chain, produced the highest $T_m$ in both conditions amongst the aliphatic amine derivatives, and suggested that the longer side chain with higher flexibility may provide better interactions of the positive charge with the grooves.

Amongst the compounds with polyamine side chains CP4 (60.48 °C /61.53 °C),
**CP5** (62.44 °C /58.54 °C) and **CP6** (71.02 °C /65.04 °C), **CP6**, with the longest side chain, had a much stronger stabilization effects than **TO** (63.99 °C /60.03 °C). Similarly, a longer side chain is more favored for the G-quadruplex stabilization. **CP6** was the strongest ligand amongst all twelve ligands under both conditions and significantly improved the binding affinity of **TO**. **CP6** consists of three positively charged amine groups on a long flexible side chain. Positively charged sites in **CP6** interact with the phosphate groups of the DNA backbones and enhanced the G-quadruplex binding. **CP4** with one less positive charge exhibited a weaker stabilization effect than **CP6**. Meanwhile, **CP5** with similar two positive charges but a shorter chain length did not exhibit the same stabilization effect as **CP6**, suggesting that a flexible long side chain is indispensable for the binding of **TO** derivatives to G-quadruplexes. This was also consistent with the stabilization behaviors of aliphatic amine derivatives **CP1**, **CP2** and **CP3**.

Polyethylene glycol side chains are more flexible than aliphatic side chains. However, **CP7** (58.68 °C/56.03 °C) and **CP8** (59.41 °C/57.03 °C) gave the worst stabilization effects under both conditions. The electronegative oxygen atoms have charge repulsion effects against the DNA phosphate backbone and consequently decrease the stabilization effects. For the same reason, **TO** acid also exhibited poor stabilization effect.

**CP9**, **CP10**, **CP11**, and **CP12** contain piperidine or pyrrolidine side chains. Both piperidine and pyrrolidine motifs can be protonated under physical conditions and form positive charges. **CP9** with a pyrrolidine ring at the end of the side chain was the second strongest stabilizer amongst all twelve ligands. A pyrrolidine group in **CP9** (69.44
°C/62.53 °C) gave better G-quadruplex stabilization than a piperidine group in **CP10** (64.03 °C/56.04 °C). A similar observation was also reported by another research group. Ligands with polyethylene glycol side chains, **CP11** (59.67 °C/60.52 °C) and **CP12** (64.88 °C/60.53 °C), did not significantly enhance the G-quadruplex stabilization effects of **TO** (63.99 °C/60.03 °C) because of the charge repulsion forces from the oxygen atom(s), even coupled with a pyrrolidine or piperidine ring.

Overall, **CP6** and **CP9** exhibited the best stabilization among twelve compounds. The pyrrolidine and polyamine side chain significantly improved the G-quadruplex stabilization effect of **TO**. Meanwhile, a long flexible linear geometry side chain of the ligands is essential for efficient interactions of the terminal amino groups with the G-quadruplexes.

### 4.3.2. Thiazole orange derivatives preserve G-quadruplex light-up probe properties

DNA dyes such as **TO** has significant fluorescence enhancements upon binding to DNA. **TO**-DNA complexes emit intensive fluorescence at around 530-540 nm with a significant quantum yield when excited at 500 nm. This light-up feature enables **TO** and some of its derivatives to be used as probes for both duplex and G-quadruplex DNA in living cells and electrophoresis gels. Meanwhile, **TO** is frequently used in a Fluorescence Intercalation Displacement (FID) assay to study other DNA ligands, in which a pre-bond **TO** molecule can be displaced off from the DNA by a strong ligand. In order to study the side chain effects on the fluorescence signal of **TO**-DNA complexes, we incubated all the derivatives with a human telomeric G-quadruplex (G4, 5’-AGGGTTAGGGTTAGGGTTAGGGT-3’) and recorded the yielded fluoresce (Figure 4-6).
Figure 4-6. Fluorescence scan of TO derivatives in the presence of human telomeric G-quadruplex. $C_{\text{G-quadruplex}}=0.25$ µM; $C_{\text{TO}}=0.5$ µM.

The background signals of the derivatives in the absence of DNA are nearly zero in solution. The maximum fluorescence intensities of the compounds were all present at around 535 nm. No significant shift of the maximum fluorescence emission was observed for any of the derivatives because the conjugated chromophore structures of the derivatives were not greatly affected by the side chains.
The maximum fluorescence intensities of the ligand-DNA complexes were plotted in a bar graph and compared with TO (Figure 4-7). Firstly, the overall quantum yield of twelve derivatives and TO acid in the presence of G-quadruplexes were notably lower than TO (342.7 au). Most of the side chains have more than one positive charges and may decrease the hydrophobicity of the local environment which can partially quench the fluorescence signal. Another previous reported TO conjugate with a positively charged side chain at the same position was also found diminishing the fluoresce signal of TO. 189
Amongst all the derivatives, a general trend was observed: the derivatives that exhibited good stabilization effects in the CD thermal denaturation experiments also showed a relatively higher quantum yield including **CP9** (288.7 au), **CP4** (265.9 au), **CP10** (288.7 au), **CP12** (262.8 au), **CP6** (250.9 au), **CP11** (224.7 au) and **CP3** (181.8 au). Electronegative groups containing compounds **CP7** (31.7 au), **CP8** (96.7 au), and **TO acid** (76.9 au) showed extremely low quantum yield. However, the rank order of G-quadruplex stabilization and fluorescence intensity are not exactly same. This discrepancy may be partially explained by the different charge locations and charge densities of the side chains. For instance, **CP6** exhibited the best stabilization effect in CD thermal denaturation experiment but was only at the fifth for the fluorescence intensity measurements (250.9 au). **CP6** has a side chain with three positive charges and its fluorescence response can be significantly attenuated by the hydrophilicity of this side chain.

Even with lower quantum yield compared to **TO** (342.7 au), the derivatives still preserve the light-up probe properties, and several of them such as **CP6** (250.9 au) and **CP9** (288.7 au) can be used to detect the presence of G-quadruplex with stronger binding affinity than **TO**.
4.3.3. Thiazole orange derivatives interact with human telomeric G-quadruplex with enhanced affinities. Because TO derivatives still preserve fluorescence light-up properties, the binding constants of these compounds can be conveniently extrapolated from the fluorescence titration curves. In a typical experiment, a stock solution of human telomeric G-quadruplex (G4, 5’-AGGGTTAGGTTAGGTTAGGT-3’) was prepared and gradually titrated into a solution containing the designated compound. The fluorescence spectra were recorded after each addition and the maximum intensities were plotted against the concentrations of the G-quadruplex in the mixture. The consequent curves were processed by least-squared fitting and the equilibrium dissociation constant (K_d) values were calculated accordingly. Monchaud and co-workers utilized this titration experiment to compare the binding affinities of TO against G-quadruplexes with different conformations and a one site binding mode was discovered for the TO-G-quadruplex complex.\textsuperscript{188} Similar to the reported results, our titration indicated that TO binds to G-quadruplex with a one site binding mode with the K_d value around 0.87 \textmu M (Figure 4-8).

Six compounds (CP4, CP6, CP9, CP10, CP11, and CP12) were studied with this titration method because they yielded relatively higher fluorescence intensities than the others upon binding to DNA. CP4, CP9, CP10, CP11, and CP12 exhibited a similar one site binding mode as TO. The K_d values of CP4 (0.32 ± 0.01 \textmu M), CP6 (0.27 ± 0.02 \textmu M), and CP9 (0.77 ± 0.01 \textmu M) were lower than that of TO (0.87 ± 0.02 \textmu M), while CP10 (1.13 ± 0.04 \textmu M), CP11 (1.65 ± 0.05 \textmu M), and CP12 (2.18 ± 0.07 \textmu M) produced higher K_d values than TO (0.87 ± 0.02 \textmu M) (Figure 4-8). This observation was highly consistent with the results of CD thermal denaturation in the presence of Na\textsuperscript{+} ions, which
was a similar condition used in this titration. In the CD thermal denaturing experiments, 
**CP10**, **CP11**, and **CP12** had similar or lower stabilization effects compared to **TO**, while 
**CP4**, **CP6** and **CP9** had a higher stabilization effects than **TO**.

Interestingly, the binding mode evidenced by the least-square fitting of **CP6** was 
different from those of **TO** and other derivatives (Figure 4-8). By fitting with one binding 
site mode, **CP6** produced a K<sub>d</sub> value of 0.27 µM which was the lowest amongst all the 
derivatives (Figure 4-9). However, the chi-square value of the fitting was 0.97 which was 
notably worse than those for other ligands (>0.99). Therefore, the same titration curve for 
**CP6** was fitted using the two binding sites mode and a chi-square value of 0.999 was 
obtained. Two K<sub>d</sub> values of **CP6** obtained from this fitting mode were 0.07 ± 0.04 µM 
and 1.73 ± 0.11 µM. The K<sub>d</sub> value of the main binding of **CP6** with the G-quadruplex 
was remarkably lower than the value of **TO** (0.87 ± 0.02 µM), suggesting a very strong 
binding. This result is consistent with the previous observation that **CP6** stabilized the G-
quadruplex most among all the **TO** derivatives. This fitting also suggested that there is a 
second binding site in the G-quadruplex for **CP6**. Since **CP6** is tethered with a side chain 
containing three positive charges, groove binding could occur due to the charge 
interactions between the ligand and the DNA backbones.
Figure 4-8. Fluorometric titration curves of TO and TO derivatives against human telomeric G-quadruplex. ---two binding site mode; —one binding site mode.
Figure 4-9. $K_d$ values of TO and derivatives binding to G4 (solid) and ds17 (blank) determined by the fluorometric titration experiments.

For the development of G4 ligands, selectivity against G-quadruplex over duplex DNA is always a contradictory objective when improving the binding affinity. The most efficient strategy to elevate the binding affinity of a ligand is to incorporate positively charges that interact with DNA phosphate backbones. However, electrostatic interactions do not usually provide sufficient selectivity for the ligands. In order to study the selectivity of the twelve derivatives, the same titration experiments were conducted against a 17 mer non-G-rich duplex DNA (ds17, 5'-GCCAGTTCGTAGTAACC-3'/3'-GGTTACTACGAACTGGC-5') (Figure 4-10). It was reported that TO has a similar binding affinity against G-quadruplex and duplex DNA. In our experiments, the $K_d$ value of TO binding to ds17 was determined as $0.40 \pm 0.00 \, \mu\text{M}$ which was within the same range as that of G-quadruplex binding ($0.87 \pm 0.02 \, \mu\text{M}$) under the same experimental conditions. This $K_d$ value is close to the reported values from other research
Figure 4-10. Fluorometric titration curves of TO and TO derivatives against ds17 ——two binding site mode; ——one binding site mode.
The $K_d$ values of $\text{CP4}$, $\text{CP6}$, $\text{CP9}$, $\text{CP10}$, $\text{CP11}$ and $\text{CP12}$ binding to ds17 were also determined. Their selectivity for G-quadruplex over duplex DNA was compared with the ratio of $K_d$-G4/$K_d$-ds17 (Table 4-3). The smaller the ratio is, the better the binding selectivity is. As expected, the derivatives did not show a noteworthy improved selectivity compared to TO because of the additional positively charged side chains. As discussed in chapter 1, highly selective G4 ligands such as telomestatin usually contain enlarged aromatic surfaces. The addition of multiple positively charged side chains improves their binding affinity through the charge interactions with the DNA phosphate backbones. This Coulombic attraction does not differentiate G-quadruplex from duplex DNA. $\text{CP6}$ ($K_d$-G4/$K_d$-ds17=3.20) exhibited the worst selectivity amongst the compounds because of the multiple charges in the side chain. $\text{CP9}$ ($K_d$-G4/$K_d$-ds17=2.04) was the only compound that produced a slightly better selectivity than TO ($K_d$-G4/$K_d$-ds17=2.17). This was consistent with other researcher’s work that pyrrolidine side chains can produce better selectivity than linear positively charged side chains in a G4 ligand.230

<table>
<thead>
<tr>
<th>Compound</th>
<th>TO</th>
<th>CP6</th>
<th>CP4</th>
<th>CP9</th>
<th>CP10</th>
<th>CP11</th>
<th>CP12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_d$-G4/$K_d$-ds17</td>
<td>2.17</td>
<td>3.20</td>
<td>2.59</td>
<td>2.04</td>
<td>4.14</td>
<td>2.72</td>
<td>3.05</td>
</tr>
</tbody>
</table>

4.3.4. Thiazole orange derivatives associate with human telomeric G-quadruplexes with various ratios. Electro spray-ionization mass spectrometry is a
highly sensitive analytical tool for biomolecules including oligonucleotides, peptides and their non-covalent ligand complexes.\textsuperscript{231} Compared to other analytical technologies including NMR and X-ray crystallography, ESI-MS has the advantages of easy sample preparation and low sample consumption. Because of the known ion suppression effect from non-volatile salts, ammonium acetate is one of the few salts that is compatible with ESI-MS oligonucleotides analysis. Similar to K\textsuperscript{+} and Na\textsuperscript{+} ions, NH\textsubscript{4}\textsuperscript{+} ions can also coordinate with guanines and stabilize G-quadruplexes. NH\textsubscript{4}\textsuperscript{+} ions locate themselves between two adjacent G-quartets, and the observed number of the NH\textsubscript{4}\textsuperscript{+} ions (n) in the MS spectra can be used to estimate the number of G-quartets layers (n+1).\textsuperscript{232} Thus, the presence of NH\textsubscript{4}\textsuperscript{+} ions in the G-quadruplexes is very useful to report the formation of G-quadruplexes. Meanwhile, an intercalating G4 ligand may expel the bound HN\textsubscript{4}\textsuperscript{+} ions from the G-quadruplexes which can be observed in the MS spectra. Therefore, ESI-MS is capable to generate a plenty of information of the ligand-DNA association including the binding affinity, stoichiometry and binding mode.\textsuperscript{233} In the present research, we examined all the compounds mixed with the G-quadruplex (G4, 5\textsuperscript{	extprime}-AGGGTTAGGGTTAGGGTTAG GGT-3\textsuperscript{	extprime}) with negative-ion ESI-MS and Higher Energy Collisional Dissociation (HCD) experiments. All the spectra are shown in relative abundances with the most abundant ion in each spectrum corresponding to 100\% ion intensity (base peak).

ESI-MS spectra obtained from a solution only contained G-quadruplex (MW=7270.77) showed the most abundant ions at \textit{m/z} 1459.85, corresponding to [G4+2 NH\textsubscript{4}\textsuperscript{+} -7H\textsuperscript{+}]\textsuperscript{5-} (calculated \textit{m/z} 1459.95) (Figure 4-11, top graph). There were also two smaller peaks representing [G4+1 NH\textsubscript{4}\textsuperscript{+} -6H\textsuperscript{+}]\textsuperscript{5-} at \textit{m/z} 1456.46 and [G4+5H\textsuperscript{+}]\textsuperscript{5-} at \textit{m/z} 1453.04. The latter peaks represent the unfolded, or half unfolded G-quadruplexes that
were produced in the process of ionization. Then, the same G-quadruplex solution was mixed with TO or a TO derivative in the ratio of 1-5 (DNA-ligand), and the corresponding spectra were obtained under the same condition.

Figure 4-11. Negative ion ESI mass spectra of solutions containing CP1-3 and G4. *: Free G4, α: 1-1 complexes, β: 1-2 complexes γ: 1-3 complexes.

In the spectra obtained from the solution of TO and the G-quadruplex, 1-1, 1-2 and 1-3 complexes were observed, and the 1-1 complex was the most abundant ion (Figure 4-11-TO). This observation was consistent with previously reported studies that TO forms complexes with a human telomeric G-quadruplex in the same three ratios.\textsuperscript{188}
Both CP1 and CP2 formed 1-1 and 1-2 complexes and the 1-3 complex was not detected in the mass spectra. Similar to TO, CP3 formed three complexes with a ratio of 1-1, 1-2 and 1-3, respectively. The relative intensity of free G-quadruplex in the spectra of CP3 was significantly lower than those in the spectra of TO, CP1, and CP2, suggesting stronger binding of CP3 to the G-quadruplex. This observation is consistent with the CD thermal denaturation results. CP3 showed better stabilization of the same G-quadruplex than TO, CP1, and CP2. It is plausible the longer side chain (six carbons) of CP3 allows the terminal amino group to interact with the G-quadruplex groove and thus stabilize the ligand-DNA complex. The side-chain length for CP1 and CP2 are two carbons and four carbons respectively, which may not be long enough to optimize the interaction between the amino group and G-quadruplex.

The binding of CP1, CP2, and CP3 to G-quadruplex could also result from the hydrophobic interaction between aliphatic chains and G-quadruplex groove. However, such interactions should be reduced or not exist without solvent in the gas phase condition of ESI mass spectrometry. Therefore, the enhanced binding of CP3 came from the electrostatic attraction between the terminal amino group and negatively charged DNA backbones. In addition, the longer aliphatic chain of CP3 may allow it to have an optimized fit to stack with the end G-quartet.
Figure 4-12. Negative ion ESI mass spectra of solutions containing CP4-6 and G4. *: Free G4, α: 1-1 complexes, β: 1-2 complexes γ: 1-3 complexes.

**CP4, CP5, and CP6** all contain a side chain of polyamines. **CP4** that contains a short side chain with two amino groups formed three complexes with a ratio of 1-1, 1-2 and 1-3 (DNA-ligand), respectively (Figure 4-12). The mass spectrum of **CP4** was similar to that of **CP3**, with a decreased intensity of the peak representing free G-quadruplex. Surprisingly, **CP5** and **CP6** each containing three amino groups formed much less abundant complexes. The intensities of 1-1, 1-2, and 1-3 complexes for these two compounds were much less than those in the spectrum of **CP4**. The peaks representing free G-quadruplex were dominant in the mass spectra of **CP5** and **CP6**.
These observations were not consistent with our previous studies that side chain with multiple positive changes should enhance the binding of a ligand to DNA. This discrepancy could be attributed to the following factors: 1) it was reported that for the compounds with a polyamine side chains, when the charge number and charge distance increase, their solubility will drop significantly.\textsuperscript{235-237} 2) side chains of CP\textsubscript{5} and CP\textsubscript{6} are similar to spermine which was known to cause DNA condensation at high concentrations.\textsuperscript{238} Indeed, DNA condensation was observed when the concentration of CP\textsubscript{6} was over 80 µM in our experiments, which could reduce the detectable amount of the complexes. 3) In the ESI mass spectrometric experiments, the interference between the high concentration (150 mM) of NH\textsubscript{4}\textsuperscript{+} and CP\textsubscript{5} or CP\textsubscript{6} may exist. A recent study suggested that ammonium ions competitively bind to DNA phosphate backbones and can significantly decrease the binding of multiply charged compounds to the target DNA.\textsuperscript{239} In our case, the formation of low abundance of complexes between CP\textsubscript{5} or CP\textsubscript{6} with the G-quadruplex in ESI mass spectra could result from the same reason.
The formation of complexes between **CP7** and the G-quadruplex was negligible (Figure 4-13). Only a very small amount of 1-1 complex was observed (Figure 4-13). **CP8** formed good amounts of 1-1 and 1-2 complexes with the G-quadruplex. The short ethylene glycol linker of **CP7** significantly decreased the binding of **CP7** to the G-quadruplex probably due to the repulsion between the electron-dense oxygen atom and negatively charged DNA backbones. The longer ethylene glycol linker for **CP8** enables the interaction between the terminal amino group and DNA grooves; therefore, no noticeable decrease of **CP8** in the formation of the complexes was observed. TO acid
formed less abundance of 1-1 and 1-2 complexes because of the repulsion between the 
carboxylic acid group and the DNA backbones.

![Figure 4-14. Negative ion ESI mass spectra of solutions containing CP9-12 and PCG1. 

**CP9** with a pyrrolidine side chain exhibited a much higher affinity than **TO** with 
three complexes (1-1, 1-2, and 1-3) formed (Figure 4-14). The most abundant ions were 
1-2 complexes and almost no free G-quadruplex was observed in the presence of **CP9**. 
**CP10, CP11** and **CP12** also produced similar MS spectra comparable to **CP9**. The strong
binding CP9 to the G-quadruplex determined by ESI mass spectrometry was consistent with the results from CD thermal denaturation. The strong binding of these four complexes with the G-quadruplex was probably due to the known binding effect of pyrrolidine and piperidine to G-quadruplex DNA. In addition, the ESI-MS spectra revealed that the most abundant complexes of these four ligands were with a ratio of 1-2 suggesting there are two main binding sites in G-quadruplex suitable for these four ligands.

The relative binding affinity of ligands with DNA can be quantitatively determined using the Intensity Ratio (IR), which is defined as an abundance ratio of all the complex ions to the sum of all free G-quadruplexes and complex ions in an ESI-MS spectrum (Equation 1).\(^{233}\)

\[
IR = \frac{\sum I_{r(G+nL)}}{\sum I_{r(G)}+\sum I_{r(G+nL)}}
\]

\(^{(1)}\)

Here, \(\sum I_{r(G+nL)}\) is the total intensities of the complexes where \(n=1, 2\) and 3; \(\sum I_{r(G)}\) is the intensity of free G-quadruplexes that was not associated with the compounds. Therefore, the IR value reflect the binding affinity of the ligands to the G-quadruplex DNA (Figure 4-15).
Figure 4-15. IR values for TO derivatives binding with G4.

Overall, all the derivatives except for CP7, produced abundant peaks in the ESI-MS spectra representing the DNA-ligand complexes in the ratios of 1-1, 1-2, and 1-3. The weak binding of CP7 to the G-quadruplex was also observed in the previous CD thermal denaturation experiments. CP9, with a pyrrolidine side chain was proved to be a good G4 ligand binding to G-quadruplex with 1-1, 1-2 and 1-3 mode. The trend of the complex formation by CP1, CP2, CP3, CP7, and CP8 suggested that a long flexible side chain is essential to enable the terminal positively charged group to interact with G-quadruplex phosphate backbones.
Table 4-4. IR values for TO derivatives binding with the G4 and the corresponding complex compositions.

<table>
<thead>
<tr>
<th>Compound</th>
<th>CP1</th>
<th>CP2</th>
<th>CP3</th>
<th>CP4</th>
<th>CP5</th>
<th>CP6</th>
<th>CP7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free G4 (%)</td>
<td>100</td>
<td>100</td>
<td>37.17</td>
<td>68.39</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>1-1 (%)</td>
<td>97.81</td>
<td>77.43</td>
<td>96.79</td>
<td>100</td>
<td>42.66</td>
<td>46.61</td>
<td>3.38</td>
</tr>
<tr>
<td>1-2 (%)</td>
<td>36.74</td>
<td>24.60</td>
<td>100</td>
<td>61.20</td>
<td>5.69</td>
<td>5.96</td>
<td></td>
</tr>
<tr>
<td>1-3 (%)</td>
<td>0</td>
<td>0</td>
<td>22.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR (%)</td>
<td>57.36</td>
<td>50.50</td>
<td>85.50</td>
<td>70.21</td>
<td>32.59</td>
<td>34.46</td>
<td>3.27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compound</th>
<th>CP8</th>
<th>CP9</th>
<th>CP10</th>
<th>CP11</th>
<th>CP12</th>
<th>TO</th>
<th>TOacid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free G4</td>
<td>91.32</td>
<td>8.80</td>
<td>11.16</td>
<td>31.28</td>
<td>14.58</td>
<td>85.32</td>
<td>100</td>
</tr>
<tr>
<td>1-1 (%)</td>
<td>100</td>
<td>41.81</td>
<td>46.19</td>
<td>78.70</td>
<td>52.71</td>
<td>100</td>
<td>67.36</td>
</tr>
<tr>
<td>1-2 (%)</td>
<td>38.89</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>48.21</td>
<td>31.77</td>
</tr>
<tr>
<td>1-3 (%)</td>
<td>60.20</td>
<td>45.10</td>
<td>40.75</td>
<td>13.16</td>
<td>11.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR (%)</td>
<td>60.33</td>
<td>95.83</td>
<td>94.49</td>
<td>87.52</td>
<td>91.28</td>
<td>65.41</td>
<td>52.63</td>
</tr>
</tbody>
</table>

IR (%) = \([1-1] + [1-2] + [1-3]/ G4 + [1-1] + [1-2] + [1-3]\)

Similar experiments were performed with CP3, CP6 and CP9 binding to a 17mer duplex DNA (ds17, 5’-CCTCGGCCGCGCCCACC-3'/3’GGAGCCCGGCCGCTGG-5’, MW=10999.25) to study the binding selectivity. These three compounds were chosen because they are the representative derivatives binding to the G-quadruplex efficiently in each category.

The ESI-MS spectra showed that TO binds to ds17 with 1-1, 1-2 and 1-3 ratios (Figure 4-16). CP3, and CP6 bind to ds17 with 1-1 and 1-2 ratios. In addition to these
three ratios, CP9 binds to ds17 with a 1-4 binding ratio. In all the spectra, the peaks at 
$m/z$ 1833.00 representing intact ds17 [ds17-6H]$^6$ were observed (calculated $m/z$
1832.21).

Figure 4-16. Negative ion ESI mass spectra of solutions containing ds17, CP3, CP6, and

The IR values for duplex binding were calculated using the following equation
(equation 2).

\[
IR_{ds} = \frac{\sum lr(ds + nL)}{\sum lr(ds) + \sum lr(ds + nL)}
\]  \(2\)

Here, $\sum lr(ds + nL)$ is the total intensities of the complexes with ds17 where
n=1, 2, 3, and 4; $\sum lr(ds)$ is the intensity of free ds17 that was not associated with the compounds. The IR value reflects the binding affinity of the ligands to ds17.

The obtained IR$_{ds}$ values (Figure 4-17) were compared with their IR$_{G4}$ values. TO exhibited minimal selectivity as the IR$_{ds}$ value is slightly higher the IR$_{G4}$ value. CP9 showed a better selectivity for G4 over duplex DNA, indicated by its higher IR$_{ds}$ than IR$_{G4}$. The selectivity of CP6 was worse than TO, suggested by the relatively higher IR$_{ds}$/IR$_{G4}$ ratio (Table 4-5). This observation is highly consistent with the results of fluorometric titration, except for CP3 which was not studied by fluorometric titration due to the low quantum yield. These data further confirmed that the pyrrolidin side chain of CP9 enhances a moderate selectivity for G-quadruplex over duplex DNA. The overall rank order for the strength of G4 binding is CP9 > CP3 > TO > CP6 and the rank order for the strength of duplex DNA binding is CP9 > TO > CP3 > CP6. It is noteworthy that the binding of CP6 could be underestimated because of its multiple positive charges as stated in the previous sections.
4.3.5. Thiazole orange derivatives form complexes with human telomeric G-quadruplexes with various binding modes. The binding modes of TO derivatives with the G-quadruplex DNA were further investigated by studying the gas-phase dissociation
behavior using the High Energy Collision Dissociation (HCD). Since CP9 gave the most abundant complex ions in the ESI-MS spectra amongst all compounds, we chose CP9 as a representative compound to demonstrate the binding modes. The study of the binding mode of G4 ligands is based on previously reported collisional activated dissociation (CAD) behaviors of the ligand-G-quadruplex complexes. The conclusions from this report are as follows. A ligand with a low binding affinity can easily dissociate from the G-quadruplexes at a relatively low collision energy and produces an intact G-quadruplex (three G-quartets) with two NH$_4^+$ ions. In gas phase, G-quartets stacking ligands are often observed for this dissociation behavior. By contrast, a strong ligand will be retained on the G-quadruplex even when a high collision energy is applied. Usually, the loss of guanine nucleobase(s) will happen before the strong ligand dissociates from the oligonucleotide. If the ligand is a groove binder, the G-quadruplex complex can fully or partially unfold into a complex structure with the groove binder attached, which can be detected in the HCD. The distinct feature of the complex is that it has zero or one NH$_4^+$ ions, instead of two NH$_4^+$ ions as observed in the ligand-G-quadruplex complexes. Meanwhile, some ligands form dimers to bind with the G-quadruplexes which should be taken into the considerations.

In the HCD experiments, the energy-dependent product ion spectra for 1-1, 1-2 and 1-3 G4-ligand complexes were obtained at the collision energies of 5 and 10 arbitrary units (au), respectively. For clarity, the rule of annotations below will be used for all the spectra. The Greek letters $\alpha$, $\beta$, and $\gamma$ represent 1-1, 1-2 and 1-3 complexes, respectively; The symbol nN denotes the number of NH$_4^+$ ions in the designated complexes. For example, $\gamma2N$ represents a 1-3 complex with two NH$_4^+$ ions. The symbols 0N and 1N
represent zero or one NH$_4^+$ in the complexes. As indicated above, these two cases mostly reflect a unfolded or partially unfolded complex structure (not an intact G-quadruplex secondary structure). The possible fragment pathways of G4-CP9 complexes were proposed based on the patterns we observed in MS/MS spectra, which will be explained in detailed (Figure 4-18).

Figure 4-18. The proposed fragmentation pathway of G4-CP9 complexes in MS/MS.
Figure 4-19. Energy dependent product ion spectra of G4-CP9 1-3 complexes (γ2N). From top to bottom: a) ESI-MS spectrum of a solution containing CP9 and G4; MS/MS product ion spectra of G4-CP9 [1-3] complexes at the collision energy of b) 5 au and c) 10 au. α: 1-1 complexes, β: 1-2 complexes, γ: 1-3 complexes; nN, n= the number NH₄⁺ ions.

In this experiment, the 1-3 complexes γ2N [3CP9+G+2NH₄⁺-7H]⁵⁻ at m/z 1726.57 were isolated and collisionally activated at the energy of 5 au (Figure 4-19a). As expected, the relative intensity of the precursor ion decreased to 24.13% compared with the base peak (Figure 4-19b). The dissociation of the ligand yielded a 1-2 complex β2N [2CP9+G+2NH₄⁺-7H]⁵⁻ at m/z 1637.71 (Figure 4-19b). The presence of two NH₄⁺ ions suggested that this was a complex between two ligands and one intact G-quadruplex. In addition, the binding of the lost ligand was relatively weak. The binding of the remaining two ligands was stronger than the dissociated one; otherwise, we would observe a 1-1
complex or fully dissociated complex. The binding mode of the dissociated ligand could be end-stacking. It was reported that a G-quartets stacking ligand has relatively low affinity in the gas phase and even the strongest stacking ligand can be easily removed from the G-quadruplex at a low collision energy in MS/MS.\textsuperscript{240} Under the HCD condition at the 5 au collision energy, $\beta_2N$ further dissociated into $\beta_1N [2CP9+G+NH_4^+-6H]^5-$ at $m/z$ 1634.33 and $\beta_0N [2CP9+G-5H]^5-$ at $m/z$ 1630.94. The binding mode of the remaining two ligands could be groove binding because $\beta_1N$ and $\beta_0N$ both contained unfolded or partially folded G-quadruplexes. No observation of the 1-1 complex in the HCD experiment also supported this conclusion (step c and d in Figure 4-18).

The binding mode of the third ligand could also be groove binding. There were also 1-3 complexes $\gamma_1N [3CP9+G+NH_4^+-6H]^5-$ at $m/z$ 1723.15 and $\gamma_0N [3CP9+G-5H]^5-$ at $m/z$ 1719.75 present in the product ions of 1-3 complexes at the 5 au energy level. These product ions were produced by a precursor 1-3 complexes with three strong groove binders ($\gamma_2N^*$ in Figure 4-18). Three groove binders bind to the G-quadruplex with a high affinity and they were all retained on the unfolded G-quadruplex (Step e and f in Figure 4-18). The evidence for three groove binders also came from the HCD experiment of $\gamma_2N [3CP9+G+2NH_4^+-7H]^5-$ with the collision energy of 10 au (Figure 4-19c). Under this condition, G-quadruplex fully dissociated, forming a complex ion $\gamma_0N [3CP9+G-5H]^5-$ at $m/z$ 1719.75 with no NH$_4^+$ ions. The complex ion $\beta_0N [2CP9+G-5H]^5-$ at $m/z$ 1630.93 became the dominant peak amongst the complex ions containing two ligands and DNA. Almost no 1-1 complex or free G-quadruplex was observed with the collision energy of 10 au, further suggesting that two or three ligands of the 1-3 complexes had a high binding affinity.
Figure 4-20. Energy dependent product ion spectra of G4-CP9 1-2 complexes (β2N). From top to bottom: a) ESI-MS spectrum of a solution containing CP9 and G4; MS/MS product ion spectra of G4-CP9 [1-2] complexes at the collision energy of b) 5 au and c) 10 au. α: 1-1 complexes, β: 1-2 complexes, γ: 1-3 complexes; nN, n=the number NH4+ ions.

1-2 complexes with two NH4+ ions β2N [2CP9+G+2NH4+-7H]5− at m/z 1637.53 were also isolated and collisional activated at the energy of 5 au (Figure 4-20a). No precursor ions were observed at the 5 au collision energy (Figure 4-20b). The dissociation pattern was similar to that of 1-2 product complexes in the HCD experiments of 1-3 complexes. At both energy levels, peaks representing 1-1 complexes were observed with very low abundance. With the collision energy of 10 au, only β0N [2CP9+G-5H]5− at m/z 1630.93 was observed as the dominant peak, suggesting two groove binders present in the parent ion. The low abundance of α2N[CP9+G+2NH4+-7H]5− at m/z 1545.08 represented a complex formed between an intact G-quadruplex and one CP9 that strongly
binds to G-quadruplex via groove binding (step d and k in Figure 4-18).

Interestingly, at the 10 au collision energy, a fragment ion at $m/z$ 1600.89 was observed, representing the loss of a guanine from the complex ion $\beta0N \ [2\text{CP9} \ + \ \text{G-5H}]^{5-}$ at $m/z$ 1630.90 (step o and k in Figure 4-18). The same guanine loss fragment pathway was previously reported for several other strong G4 ligands especially those with positively charged side chains.\textsuperscript{242} This observation indicated that the groove binding of CP9 was remarkably strong and will be retained by the oligonucleotide even when the fragmentation of DNA occurred.

Figure 4-21. Energy dependent product ion spectra of G4-CP9 1-1 complexes ($\alpha2N$). From top to bottom: a) ESI-MS spectrum of a solution containing CP9 and PCG1; MS/MS product ion spectra of G4-CP9 [1-1] complexes at the collision energy of b) 5 au and c) 10 au. $\alpha$: 1-1 complexes, $\beta$: 1-2 complexes, $\gamma$: 1-3 complexes; nN, n=the number NH$_4^+$ ions.
The 1-1 complexes $\alpha_2N[1\text{CP9}+G+2\text{NH}_4^+-7\text{H}]^{5-}$ at $m/z$ 1548.75 were isolated and activated at the 5 au collision energy (Figure 4-21a). The single ligand CP9 remained associated with the DNA, but all the secondary G-quadruplex structures were unfolded to give $\alpha_0N[1\text{CP9}+G-5\text{H}]^{5-}$ at $m/z$ 1541.89 (step m in Figure 4-18) (Figure 4-21b). At the 10 au collision energy, $\alpha_0N$ still remained as the most abundant peak. There was still no free G-quadruplex was produced but instead, many fragment ions were observed (Figure 4-21c). The major peak was a guanine loss fragment ion at $m/z$ 1512 that was produced from the ion $\alpha_0N[1\text{CP9}+G-5\text{H}]^{5-}$ $m/z$ 1541.89 (step n in Figure 4-18).

Based on all the results from the HCD experiments of CP9, we can reasonably conclude that three CP9 molecules could bind to the G-quadruplex. Two of them have stronger binding affinities than the third one. The binding of the third CP9 to the G-quadruplex could be either end stacking or groove binding or both. Therefore, the complexes formed by CP9 with different binding modes in the same solution may exist. Such complexes cannot be differentiated via $m/z$ values. Two CP9 molecules with stronger binding affinities could also bind to the G-quadruplex via different modes. They can bind to the G-quadruplex in two separate grooves with similar interactions. It is also possible that they bind to the same groove of the G-quadruplex as a dimeric form. The binding of DNA by cyanine dimers have been previously reported. Nevertheless, the binding affinities of these two CP9 molecules should be similar because the dissociation of one CP9 from the 1-2 complex was minimal in the HCD experiments.

There is a noteworthy discrepancy between ESI-MS and fluorometric titration experiments (section 4.3.3). MS/MS suggested multiple binding sites for CP9, while only 1-1 binding mode was observed in the fluorometric titration experiments. In a
fluorometric titration, the emission of the fluorescence majorly comes from the stacking of the ligand to the G-quadruplexes. The groove binders does not provide as efficient quantum yields as stacking ligands because the structure of bound groove binders is relatively flexible. For most of the derivatives and TO, the quantum yields from the groove binding were negligible. In addition, a perfect agreement between results from fluorescence and ESI-MS should not be expected due to the differences in experimental conditions.

We further studied CP9 with an intermolecular G-quadruplex [d(T2G5T)]4(G5) to investigate the binding modes. Different from the human telomeric G-quadruplex PCG1 (G4), G5 was formed by four short G-rich sequences (S1, 5’-TTGGGGGT-3’, MW=2496.70) containing five G-quartet layers. Therefore, four NH₄⁺ ions were incorporated in the structure. The use of the intermolecular G-quadruplex in a MS/MS study allows us to determine the nature of the complex ions. For instance, if the ligand non-selectively binds to DNA random coils, a complex ion (S1+ CP9) will be observed. If the ligand can bind to the G-quadruplex, only the complex ion (G5 +CP9) can be observed. Unlike the experiments using intramolecular G4, the m/z values of these two complex ions are drastically different. In addition, if a ligand adopts the intercalation mode, the ejection of NH₄⁺ ions from the G5 should be expected.
Figure 4-22. Energy dependent product ion spectra of G5-CP9 complexes. From top to bottom: a) ESI-MS spectrum of a solution containing CP9 and G5; MS/MS product ion spectra of G5-CP9 [1-3] complexes at the collision energy of b) 2 au and c) 6 au; d) MS/MS product ion spectra of G5-CP9 [1-2] complexes at the collision energy of 10 au. α: 1-1 complexes, β: 1-2 complexes, γ: 1-3 complexes;

The ESI-MS spectra revealed that CP9 predominantly binds to G5 with the ratio of 1-2 β [2CP9+G5+4NH₄⁺-10H]⁶⁻ at m/z 1822.88 and 1-3 γ [3CP9+G5+4NH₄⁺-10H]⁶⁻ at m/z 1896.91 (Figure 4-22 a). A small amount of 1-1 α [CP9+G5+4NH₄⁺-10H]⁶⁻ at m/z 1748.68 was also present. No free G5 was observed, suggesting the strong binding affinity of CP9. Different from the binding of CP9 with G4, no ion peaks representing partially folded or unfolded G-quadruplexes were present in the spectra. All the G-quadruplex exist as intact G-quadruplex with four NH₄⁺ ions.
The complex ion 1-3 $\gamma [3\text{CP9}+\text{G5}+4\text{NH}_4^+-10\text{H}]^{6-}$ at $m/z$ 1896.91 was isolated and subjected to the collision-activated dissociation at the energy of 2 au (Figure 4-22b). Removal of one CP9 ligand from the λ ion yielded a 1-2 complex $\beta [2\text{CP9}+\text{G5}+4\text{NH}_4^+-10\text{H}]^{6-}$. The 1-1 complex, free G-quadruplex and S1 strand were not observed in the product ions. A similar spectrum was obtained even when the collisional energy increased to 6 au (Figure 4-22c). This observation suggested that, consistent with previous results of CP9 binding with G4, one of the ligands in the 1-3 complex has a relatively low binding affinity as compared to other two bound CP9.

The 1-2 complex $\beta [2\text{CP9}+\text{G5}+4\text{NH}_4^+-10\text{H}]^{6-}$ at $m/z$ 1822.88 was also isolated and subjected to the collision-activated dissociation at a relatively high energy (10 au) (Figure 4-22d). No 1-1 complex or free G5 was observed, indicating a remarkably strong binding of these two ligands. The base peak was at $m/z$ 1722.03. Based on the charge state and the $m/z$ value, it represents the complex ion resulting from the loss of a thymidine (nucleotide) from the $\beta$ ion. The exact mechanism of the loss of thymine remains unclear to us at this point. Interestingly, reanalyzing the ESI-MS spectra obtained from the HCD experiment of the CP9-G4 mixture showed a similar ion with a very low intensity. Apparently, the loss of thymidine from G5 is much easier than that from G4 probably because the thymidines in G5 are more accessible (located at the both ends of the sequence). The intramolecular G4 only has one thymine at the end, and its accessibility is restricted by the loops on top of the G-quartets.

Among the product ions of 1-2 complexes, the ions representing S1 single strand [S1-3H]$^{3-}$ at $m/z$ 830.81 and [S1-2H]$^{2-}$ at $m/z$ 1247.22 were observed. Interestingly, these two strands were found as free oligonucleotides without the association of any ligands.
The sequence of S1 is too short to form certain secondary structures for the binding of CP9.

All the CP9 containing ions in the spectra have all four NH$_4^+$ ions. It was reported that only ligands intercalating to the G-quadruplex can expel NH$_4^+$ ions from the G-quadruplex.$^{140,244}$ The preserved NH$_4^+$ ions indicated that CP9 must adopt groove binding and/or end stacking mode.

4.3.6. Thiazole orange derivatives inhibited the telomerase activity in vitro.

Since significant human telomeric G-quadruplex binding was discovered by different biophysical methods for the TO derivatives, inhibition of telomerase-mediated telomere extension was studied by a modified $^{32}$P-TRAP assay as illustrated in section 3.3.5.

Firstly, the concentration-dependent telomerase inhibition of TO was examined (Figure 4-23) and inhibition of telomerase was observed when its concentration was larger than 6.25 µM. To compare the telomerase inhibition of TO derivatives to TO, 6.25 µM was chosen as the concentration to be studied for all the derivatives (Figure 4-24).
Figure 4-23. $^{32}$P-TRAP reactions in the presence of TO at concentrations of 50, 25, 12.5, 6.25, 3.125, 1.56, 0.78 and 0.39 µM (from left to right). Reactions in the presence BIBR (50 µM) and TMPyP4 (50 µM) were carried out as controls. IS: Internal Standard (TSNT as template).

For aliphatic amine side chain containing compounds CP1, CP2, and CP3, telomerase inhibition effect increased with the increasing length of the aliphatic side chains. CP3 exhibited a higher inhibition effect compared to TO at the same concentration. This observation is in agreement with the results from our previous biophysical studies. Hydrophobic side chains cannot provide sufficient interactions for
the ligands binding. However, a better inhibition effect can be observed when the length of the side chain is long enough to accommodate the terminal positive charge to the DNA phosphate backbones.

Polyamine side chain containing compounds CP4, CP5 and CP6 exhibited the best inhibition effects amongst all twelve derivatives. At 6.25 µM CP6 inhibited both telomerase elongation and polymerase replication activity. Thus, a concentration-dependent study was necessary for CP6 for the reason explained in section 3.3.5.

Similar to our previous conclusions, CP7 and CP8 with polyethylene glycol amine side chains and TO acid did not inhibit telomerase activity at 6.25 µM. The electron-rich oxygen atoms in polyethylene glycol, and carboxylic acid group in TO acid have the repulsive force with the DNA negatively charged backbones.

Piperidine/pyrrolidine side chains containing compounds CP9, CP10, CP11 and CP12 inhibited telomerase activity at various levels. CP9, with a pyrrolidine side chain was the best inhibitor amongst these four ligands which was consistent with the previous results. Our observations suggested again that pyrrolidine rings had a better G-quadruplex stabilization compared to piperidine rings.228 CP11 and CP12 showed less inhibition effect than CP9 and CP10 due to their electronegative side chains.
Figure 4-24. $^{32}$P TRAP reactions in the presence of TO derivatives at a concentration of 6.25 µM. Reactions in the presence BIBR (50 µM), TMPyP4 (50 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).

We conducted concentration studies for the effective compounds including CP4, CP5, CP6, CP9, CP10, CP11 and CP12. The minimum inhibitory concentration for CP4 and CP5 were 0.63 µM and 1.25 µM respectively (Figure 4-25), which were significantly lower than the minimum inhibitory concentration of TO (6.25 µM).
Figure 4-25. $^{32}$P TRAP reactions in the presence of CP4 (left) and CP5 (right) at concentrations of 10, 5, 2.5, 1.25, and 0.625 µM (from left to right). Reactions in the presence BIBR (50 µM) and TMPyP4 (50 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).

**CP6** showed the strongest inhibitory effect on telomerase activity. The minimum inhibitory concentration was determined as 0.16 µM which was significantly lower than TO. It is noteworthy that, **CP6** adopted a similar structure of polyamine molecule spermine. The minimum telomerase inhibitory concentration of spermine that determined by the same assay was higher than 10 µM (Figure 4-26). Therefore, the significant G-quadruplex binding and telomerase inhibition effect of **CP6** resulted from the synergistic effect of the TO core structure and the polyamine side chain (Figure 4-26).
Because of the observed remarkable telomerase inhibition effect of CP6, we compared its activity with one of the best synthetic telomerase inhibitors BRACO-19\textsuperscript{155}, \textsuperscript{215,245} under the same condition (Figure 4-27).

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Concentration (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMPyP4</td>
<td>- - + + + + + + + +</td>
</tr>
<tr>
<td>BIBR</td>
<td>- - + + + + + + + +</td>
</tr>
<tr>
<td>CP-6</td>
<td>- - + + + + + + + +</td>
</tr>
<tr>
<td>Spermine</td>
<td>- - + + + + + + + +</td>
</tr>
</tbody>
</table>

Figure 4-26. $^{32}$P TRAP reactions in the presence of CP6 at concentrations of 10, 5, 2.5, 1.25, 0.625, 0.313, 0.156, and 0.078 µM (from left to right). Reactions in the presence BIBR (50 µM), TMPyP4 (50 µM) and spermine (100 and 10 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).
Figure 4-27. $^{32}$P TRAP reactions in the presence of BRACO-19 (left) and CP6 (right) at concentrations of 10, 2.5, 0.625, 0.156 and 0.039 µM (from left to right). Reactions in the presence BIBR (50 µM) and TMPyP4 (50 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).

The results suggested that CP6 had a slightly better inhibition effect compared to BRACO-19. At the minimum inhibitory concentration of 0.625 µM, the reaction in the presence of CP6 produced less abundant telomerase elongated products than the reaction in the presence of BRACO-19. However, CP6 had slightly worse selectivity compared to BRACO-19. At the concentration of 10 µM, CP6 had a stronger inhibition effect for the polymerase replication.
Concentration-dependent studies were also performed for CP9, CP10, CP11 and CP12 (Figure 4-28). CP9 (minimum inhibitory concentration 1.25 µM) exhibited the strongest inhibition effect among these four ligands but not as good as CP6, and this observation is highly consistent with the results of CD thermal denaturation and fluorometric titration experiments. The minimum inhibitory concentrations were around 2.5 µM for CP10, CP11, and CP12. Since piperidine and pyrrolidine side chains were incorporated into these compounds, PIPER, a known G4 ligand with a similar piperidine side chain, was examined under the same conditions. At 10 µM, all four TO derivatives exhibited much greater telomerase activity inhibition than PIPER.

Figure 4-28. $^{32}$P TRAP reactions in the presence of CP9, CP10, CP11 and CP12 (from left to right) at concentrations of 10, 5, 2.5, 1.25, and 0.625 µM (from left to right). Reactions in the presence PIPER (10 µM), BIBR (50 µM) and TMPyP4 (50 µM) were carried out as controls. IS: PCR product as internal standard (TSNT as template).
Overall, the TRAP assay clearly showed that the TO derivative CP6 containing a polyamine side chain shows excellent inhibition of telomerase activity at a nanomolar concentration which is comparable to the best developed synthetic telomerase inhibitor BRACO-19. Conjugating with polyamine side chains at the nitrogen of the quinoline ring remarkably improved the telomerase inhibition activity of TO and this activity is a synergistic effect from the aromatic stacking of TO core structure and the electrostatic interaction from the charged side chain. TO is a suitable scaffold for developing more potent G4 ligands in the future.

4.3.7. Positively charged side chains decrease the cytotoxicity of thiazole orange. TO was originally developed as a DNA dye for reticulocyte count in a flow cytometry. One of the advantages of TO being a live-cell dye is its good cell membrane diffusion ability, which is also one of the reasons that it was chosen as the core structure in the present research. After discovering the remarkable telomerase inhibition effect of CP6 by the TRAP assay, we investigated the cytotoxicity of twelve derivatives against two telomerase-positive cancer cell lines: HeLa (cervical cancer) and MCF-7 (breast cancer).

At the concentration of 100 µM, TO was found with cell viability of 20.22% for HeLa and 40.09% for MCF-7. However, all our TO derivatives exhibited a lower cytotoxicity value compared to TO (Figure 4-29).

This decreased cytotoxicity was closely related to the high polarity of the molecules and the consequent low uptake efficiency. Although there are some native endogenous polar molecules like polyamine present in living cells and can be transported
into the cell nucleus, specific transport proteins have to be recruited by cells helping them to pass through the cell membrane. The positively charged side chains increase the polarity of the compounds and consequently decrease their cell membrane diffusion efficiency. Thus, these TO derivatives cannot easily reach the cell nucleus where the telomeres and telomerase are located. In order to develop these good TO based telomerase inhibitors into anti-cancer drugs, we propose to explore the drug delivery system to increase cellular uptake efficiency in collaboration with other groups in the future.

Figure 4-29. Cell proliferation of two cancer cell lines treated by TO derivatives at 100 μM. The cell proliferation in the absence of a ligand was normalized to 100%.
4.4. Conclusion

Twelve newly synthesized TO derivatives with positively charged side chains at the nitrogen of quinoline ring were evaluated by various biophysical techniques for their G-quadruplex DNA stabilization effects. Marked stabilization effects against human telomeric G-quadruplexes were found for two of the derivatives with either a polyamine side chain (CP6) or a pyrrole side chain (CP9). In contrast to TO, these TO derivatives form 1-1, 1-2 and 1-3 complexes with a human telomeric G-quadruplex with both stacking and groove binding modes. The binding affinity was significantly enhanced by these positively charged side chains through electrostatic interactions with DNA backbones.

In the TRAP assay, CP6 exhibited remarkable telomerase inhibition effect comparable to BRACO-19, which is the most potent synthetic telomerase inhibitor. Elongation of telomeric DNA by telomerase was significantly inhibited at nanomolar concentrations by CP6.

Introducing positively charged side chains to TO decreased the cytotoxicity due to the higher polarity of the molecules. To utilize CP6, as a strong G4 ligand and telomerase inhibitor, we will explore suitable drug delivery systems. TO derivatives have slightly lower quantum yields compared with TO when they bind to DNA. However, the quantum yields of these compounds are still much higher than other fluorescent G4 ligands; therefore, it is still feasible to use these compounds as highly sensitive G-quadruplex probes with significantly stronger binding affinities than TO.
REFERENCE


(18), 6151-60.
103. Cerone, M. A.; Londono-Vallejo, J. A.; Bacchetti, S., Telomere maintenance by


162. Pennarun, G.; Granotier, C.; Gauthier, L. R.; Gomez, D.; Hoffschir, F.; Mandine,


212. Ruan, T. L.; Davis, S. J.; Powell, B. M.; Harbeck, C. P.; Habdas, J.; Habdas, P.; Yatsunyuk, L. A., Lowering the overall charge on TMPyP4 improves its selectivity for G-


242. Li, H.; Yuan, G., Collision-Induced Dissociation of Dimeric G-Quadruplexes of
HIV-1 Integrase Inhibitors and Their Complexes by Tandem-in-Time Mass Spectrometry.  


APPENDIX
Figure A-1. MS spectrum of Phen12-Ni_50% MeOH
Figure A-2. MS spectrum of Phen15-Ni_50% MeOH
Figure A-3. MS spectrum of Phen18-Ni_50% MeOH
Figure A-4. MS spectrum of Phen20-Ni_50% MeOH.
Figure A-5. MS spectrum of Phen21-Ni_50% MeOH
Figure A-6. FID G4 titration course of 5-substituted Phen Ni(II) complexes.
Figure A-7. $^{13}DC_{50}$ plots for Phen-Ni
Figure A-8. $\text{G}^4\text{DC}_{50}$ plots for 13, 20 and 21 in the absence of Ni.
Figure A-9. Concentration study of effective 5-substituted Phen-Ni complexes on tumor cells.
Figure A-10. CD thermal denaturation smoothed profiles of TO derivatives.
Figure A-11. Fluorometric titration of TO derivatives against a duplex 17mer.