

The Molecular Recognition of DNA by Novel Heterocycles

Thesis by

Michael Anthony Marques

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2005

(Defended 13 May 2005)

© 2005

Michael Anthony Marques

All Rights Reserved

...for my parents...
...for B.S.E....

Acknowledgments

I would like to begin by thanking Professor Peter B. Dervan for the opportunity to work in such an excellent research group. Peter has given me the freedom and encouragement to explore the scientific questions that I found most interesting. Furthermore, I would like to thank Peter for his continuous generosity, understanding and guidance during my stay here at Caltech. I will never forget the invaluable training I received from Peter as I make the transition into the field of medicine. I would also like to thank the members of my committee, Professors Linda Hsieh-Wilson, Steve Mayo, and Brian Stoltz for taking the time to help me with candidacy and research proposals. In particular, I would like to thank Brian Stoltz for helping me stay focused when I initially arrived at Caltech. Brian is exceptional in his ability to communicate the complexities of organic chemistry in a way that is inspirational. I am grateful for Brian's enthusiasm for science, his strong work ethic, and his willingness to share his time no matter how busy he is, including the time he has taken to talk with me throughout my time here about problems in heterocyclic chemistry.

I would also like to thank my undergraduate advisor, Professor Brian McNelis at Santa Clara University. Brian's dedication to training me as an undergrad truly prepared me for the road before me in graduate school. I am both grateful and indebted for his dedication to help develop my interest in science and research.

Next, I would like to thank my friends here at Caltech. First, let me thank the current and past group members. I owe a great deal to Adam Urbach for his patience,

training, advice, and friendship when I first joined the lab. I also thank Bobby Arora, now a professor at NYU, for lighthearted conversation and hospitality during my visits to NY. Thanks to my classmates Eric Fechter, Adam Kerstien, Raymond Doss, and Neil Garg. I am grateful to Eric for his friendship in good times and difficult ones. He is always someone that you can depend on. Thanks to Adam for many nights of hospitality when we would all unwind and play cards. Thanks to Neil for being a positive voice in my life, a dependable friend, and someone who would always make me laugh...even in graduate school.

I need to single out Raymond Doss. Ray is both a collaborator on the majority of the work described herein, as well as one of my best friends. I owe a great deal to Ray and know that life as a graduate student would not have been the same without him around. I am indebted to Ray for his time, thoughtfulness, camaraderie, support, friendship, and understanding. Here is to spending nights working till 4 am in the lab. Ray is one of the most respectable and upstanding individuals I know and I wish him all of the best in the future.

I would also like to thank the younger generation of students that I have had the pleasure of working with, namely, Carey Hsu, Jim Pucket, Ryan Stafford, Justin Cohen, and David Chenoweth. I would like to thank Dave in particular for his collaborative efforts during the end of my time here. Furthermore, I would like to thank Carey Hsu for the time he spent working with me over the course of the last couple of years, and the time he spent proofreading countless manuscripts and proposals...thanks again. I also

want to thank Nick Nickols, a friend that I hope to keep, and a future comrade in the field of medicine.

I would also like to mention Ben Edelson, to whom this thesis is in part dedicated. Ben taught me more than I could possibly write down in these few pages. I will always remember him and be grateful for those memories.

I go on to thank those friends that I have made in Los Angeles outside of Caltech. First, to my friend Maia, thank you for your constant friendship, love, and understanding. I hope we will continue to be in each other's lives forever and I am grateful for your tolerance and appreciation of my difficult brooding personality. I would also like to thank Zoe for her friendship. Zoe is amazing in her ability to be optimistic and lift my spirits even in the most trying of situations. Next, thanks to my friends Jason, Greg, and Albert of Bodega fame. I will always remember the good times, the schedule, the rules, and the occasions not appropriate for further comment here. Here is to future success.

I would like to thank my friends from back home: Dan, Brad, Cyrus, and Rob. Even though we haven't been able to keep in touch, and you are always giving me a hard time about being the perpetual student ("man....aren't you done yet?"), I love and miss you guys. Here is to good times in the future.

Finally, and most importantly, I thank my mother and father, Linda and Wes. Thank you for infinite kindness, patience, love, and self-sacrifice. You are both my inspiration for all of this hard work and I would not have been able to make it this far without you. I love you both very much and hope only for your happiness.

Abstract

With a rapid movement toward personalized genetic medicine, tailoring treatment to individual patient needs based each one's genetic code is becoming an important goal. The ability to develop small molecules capable of reprogramming the cellular machinery at the genetic level is one approach to the difficult challenge of treating diseases that result from aberrant gene expression. Inspired by the architecture of the natural products netropsin and distamycin, polyamides are capable of binding the DNA minor groove with high affinity and fidelity. Originally composed of 5-membered heterocyclic carboxamides, polyamides have evolved in both form and function. A search has been initiated to develop new DNA specific oligomers that have different electronic and geometric properties. Alteration of these properties may lead to a new class of compounds, capable of targeting DNA sequences that have previously been shown to be difficult to recognize. Second-generation compounds containing novel heterocyclic recognition elements, within the context of both 5-membered heterocyclic carboxamides and fused 6-5 benzimidazole analogues, have recently been developed. These molecules have successful DNA recognition profiles as well as favorable cell uptake properties, important considerations when searching for effective pharmacophores. These new classes of rationally designed oligomers offer one approach to the challenging problem of regulating gene expression.

Table of Contents

	Page
Acknowledgments.....	iv
Abstract.....	vii
Table of Contents.....	viii
List of Figures and Tables.....	x
Chapter 1 Programmable DNA Binding Oligomers for Control of Transcription.....	1
Chapter 2 Conformational Control of DNA Binding Oligomers.....	39
Chapter 3 Toward a Chemical Etiology for DNA Minor Groove Recognition.....	54
Chapter 4 Recognition of DNA by Hairpin Polyamides Containing N-Terminal Thiophene Residues.....	116
Chapter 5 DNA Minor Groove Recognition by Multiple Thiophene/Pyrrole Pairs.....	139
Chapter 6 Expanding the Repertoire of Heterocyclic Ring Pairs for Programmable Minor Groove DNA Recognition.....	163

Chapter 7	Programmable Oligomers for DNA Recognition.....	194
Chapter 8	Next Generation Programmable Oligomers for DNA Minor Groove Recogniton.....	224
Chapter 9	Minor Groove Recognition of T, A Base Pairs by Benzothiophene/Pyrrole Pairs.....	237
Chapter 10	Targeting G-Tetrads (5'-GGGG-3') in the DNA Minor Groove.....	255
Chapter 11	Solid Phase Synthesis of Oligomers Using Safety Catch Hydrazine Resin.....	289
Chapter 12	Synthesis and Properties of Oxazole Heterocycles.....	307

List of Figures and Tables

Chapter 1	Page
Figure 1.1 DNA Base Pairs & Basis for DNA Recognition.....	5
Figure 1.2 Origin of A,T & G,C Specificity.....	6
Figure 1.3 Selected Oligomer Binding Motifs.....	9
Figure 1.4 Family of 5-Membered Heterocyclic Carboxamides.....	11
Figure 1.5 Curvature of 4-Ring Oligomeric Subunits.....	12
Figure 1.6 Binding Mode for Thiophene Containing Polyamides.....	14
Figure 1.7 5-Membered Carboxamides vs. Fused Benzimidazole Derivatives.....	15
Figure 1.8 Hydroxybenzimidazole (Hz) and Imidazopyridine (Ip) Hairpins.....	16
Figure 1.9 NMR Structure of 1:1 Polyamide:DNA Complex.....	17
Figure 1.10 Solid Phase Synthesis Methodology for Polyamide Generation.....	19
Figure 1.11 Crystal Structure of Polyamide:NCP Complex.....	21
Figure 1.12 Important Transcription Sites Targeted by Polyamides.....	24
Figure 1.13 Activation & Repression of Transcription by Polyamides.....	25
Figure 1.14 Polyamide Conjugates.....	27
Figure 1.15 Regulation of Endogenous Genes by FITC Conjugates.....	29

Chapter 2

Figure 2.1	Equilibrium Between Hairpin & Extended Binding Modes.....	42
Figure 2.2	Chemical Structures of Conformer Study Compounds.....	43
Figure 2.3	Plasmid Sequence for pAU27.....	43
Figure 2.4	DNase I Footprinting Titrations for Conformer Study.....	45
Figure 2.5	Proposed Model for Controlling Oligomer Conformation.....	47
Table 2.1	Equilibrium Association Constants for Conformation Study.....	44

Chapter 3

Figure 3.1	High Resolution Structures of Polyamide:DNA Complexes.....	56
Figure 3.2	Family of 5-Membered Heterocyclic Carboxamides.....	58
Figure 3.3	1:1 & 2:1 Polyamide Designs for Specificity Studies.....	59
Figure 3.4	Formation of Monomer & Dimer Boc-Protected Amino Acids.....	61
Figure 3.5	Monomeric and Dimeric Boc-Protected Amino Acids for SPS.....	62
Figure 3.6	Synthesis of Furan (Fr) Monomer & Dimer.....	63
Figure 3.7	Synthesis of N-H Pyrrole (Nh) Monomer.....	64
Figure 3.8	Synthesis of Thiophene (Tn) Monomer & Dimer.....	65
Figure 3.9	Synthesis of Des-Methyl Thiophene (Dt) Monomer & Dimer.....	66
Figure 3.10	Synthesis of Methoxythiophene (Mt) Monomer.....	67
Figure 3.11	Synthesis of N-Thiazole (Nt) Monomer.....	68
Figure 3.12	Solid Phase Synthesis of 1:1 and 2:1 Binding Polyamides.....	69

Figure 3.13	DNase I Titrations for Hairpins Containing Nh, Tn, Ht & Fr.....	70
Figure 3.14	DNase I Titrations for 1:1 Compounds Containing Py, Hp, Nh & Ht.....	73
Figure 3.15	DNase I Titrations for 1:1 Compounds Containing Fr, Nt, Tn & Th.....	75
Figure 3.16	Geometric & Electronic Profiles of 5-Membered Heterocycles.....	76
Figure 3.17	Curvature of Contiguous 4-Ring Polyamides.....	77
Figure 3.18	<i>Ab Initio</i> Models of Subunits Containing Ht, Hp, Th & Tn Heterocycles...	81
Table 3.1	Equilibrium Association Constants for Hairpin Specificity Study.....	71
Table 3.2	Equilibrium Association Constants for 1:1 Specificity Study.....	74
Table 3.3	Pairing Specificity Table for Hairpins Containing Novel Heterocycles.....	86
Table 3.4	Specificity Table for 1:1 Compounds.....	87

Chapter 4

Figure 4.1	Binding Model for Hairpins Containing N-Terminal Thiophene Heterocycles.....	119
Figure 4.2	Plasmid pCW15 Sequence & Hairpin Design.....	120
Figure 4.3	Synthesis of Thiophene Cap Derivatives.....	122
Figure 4.4	Solid Phase Synthesis of Thiophene Cap Hairpins.....	123
Figure 4.5	DNase I Titrations for Thiophene Cap Hairpins.....	124
Figure 4.6	Proposed Binding Model for Chlorothiophene (Ct) Cap.....	128
Table 4.1	Equilibrium Association Constants for Hairpins Containing N-Terminal Thiophene Derivatives.....	125
Table 4.2	Molecular Modeling of N-Terminal Thiophene Caps.....	126

Chapter 5

Figure 5.1	Sequence Selectivity of Hydroxypyrrole (Hp).....	142
Figure 5.2	Sequence Selectivity of 3-Methylthiophene (Tn).....	144
Figure 5.3	Solid Phase Synthesis of Hairpins Containing Multiple Tn Rings.....	146
Figure 5.4	Synthesis of Thiophene (Tn) Monomer & Dimer.....	147
Figure 5.5	DNase I Titrations for Thiophene-Containing Hairpins.....	148
Figure 5.6	DNase I Titrations for Hairpins Containing Multiple Thiophenes.....	150
Figure 5.7	Molecular Modeling of Tn, Hp & Py Surface Areas.....	151
Table 5.1	Equilibrium Association Constants for Hairpins Containing Hydroxypyrrole (Hp).....	142
Table 5.2	Equilibrium Association Constants for Hairpins Containing Multiple Hydroxypyrrole (Hp) Heterocycles.....	143
Table 5.3	Equilibrium Association Constants for Hairpins Containing Thiophene (Tn).....	149
Table 5.4	Equilibrium Association Constants for Hairpins Containing Multiple Thiophene (Tn) Heterocycles.....	151

Chapter 6

Figure 6.1	5-Membered Carboxamides vs. Fused Benzimidazole Derivatives.....	166
Figure 6.2	DNA Minor Groove Contacts for Hydroxybenzimidazole (Hz) Containing Oligomers.....	167
Figure 6.3	Synthesized Hydroxybenzimidazole Oligomers.....	168

Figure 6.4	Synthesis of Hydroxybenzimidazole-Imidazole (Hz-Im) Dimers.....	169
Figure 6.5	Solid Phase Synthesis of Hydroxybenzimidazole (Hz) Oligomers.....	171
Figure 6.6	Plasmid Sequences for pDHN1 & pDEH10.....	172
Figure 6.7	DNase I Titrations for the Hydroxybenzimidazole/Pyrrrole (Hz/Py) Pair.....	173
Figure 6.8	DNase I Titrations for Oligomers Containing Multiple Hydroxybenzimidazole Dimers.....	175
Figure 6.9	DNase I Titrations for the Hydroxybenzimidazole/Benzimidazole (Hz/Bi) Pair.....	177
Figure 6.10	Curvature of Hydroxybenzimidazole vs. 5-Membered Heterocyclic Carboxamide.....	179
Figure 6.11	Isopotential Surface of Hydroxypyrrrole (Hp) & Hydroxybenzimidazole (Hz).....	180
Figure 6.12	Isopotential Surface of Pyrrrole (Py) & Benzimidazole (Bi).....	181
Table 6.1	Equilibrium Association Constants for Hairpins Containing Hydroxypyrrrole (Hp) & Hydroxybenzimidazole (Hz).....	174
Table 6.2	Equilibrium Association Constants for Oligomers Containing Multiple Hydroxypyrrrole (Hp) & Hydroxybenzimidazole (Hz) Ring Systems.....	176
Table 6.3	Equilibrium Association Constants for Oligomers Containing a Hydroxybenzimidazole/Benzimidazole (Hz/Bi) Pair.....	178
Table 6.4	Specificity of Hydroxybenzimidazole (Hz) Oligomers.....	178

Chapter 7

Figure 7.1	Novel Thiophene & Oxazole Caps.....	197
Figure 7.2	Internal & N-Terminal Heterocyclic Ring-Ring Pairings.....	199
Figure 7.3	Synthetic Scheme for Z-Hz Dimers (Z = Im, Tn, Ct, Is & No).....	200
Figure 7.4	Solid Phase Synthesis of Z-Hz Oligomers (Z = Im, Tn, Ct, Is & No).....	201
Figure 7.5	DNase I Titrations for Imidazole (Im) & Chlorothiophene (Ct) Cap Hairpins.....	203
Figure 7.6	DNase I Titrations for Chlorothiophene-Hydroxybenzimidazole (Ct-Hz) & Imidazole-Hydroxybenzimidazole (Im-Hz) Containing Oligomers.....	204
Figure 7.7	DNase I Titrations for Isoxazole (Is) & N-Oxazole (No) Cap Hairpins.....	205
Figure 7.8	DNase I Titrations for Isoxazole-Hydroxybenzimidazole (Is-Hz) & N-Oxazole-Hydroxybenzimidazole (No-Hz) Containing Oligomers.....	206
Figure 7.9	Molecular Modeling of Imidazole-Hydroxypyrrrole (Im-Hp) & Imidazole-Hydroxybenzimidazole (Im-Hz) Dimers.....	208
Figure 7.10	Structure of Cap Heterocycles & Calculated Partial Charges.....	209
Table 7.1	Equilibrium Association Constants for Hairpins Containing Novel Cap Heterocycles.....	198
Table 7.2	Equilibrium Association Constants for Oligomers Containing	

	Novel Cap-Hydroxybenzimidazole Dimers (Z-Hz: X = Im, Tn, Ct, Is & No).....	207
Table 7.3	Specificity of Novel Cap Containing Oligomers.....	210

Chapter 8

Figure 8.1	Comparison of Carboxamide & Carbon Linkages.....	224
Figure 8.2	Hydrogen Bonding Model & DNaseI Titration for Oligomer Targeting 5'-GTAC-3'	227
Figure 8.3	Comparison of Polyamide, Polyamide Hybrid & Programmable Oligomer Structures.....	228
Figure 8.4	Molecular Modeling of Polyamide vs. Oligomer Geometry.....	230
Figure 8.5	Solid Phase Synthesis of Programmable Oligomer.....	232
Figure 8.6	O-Methoxy Deprotection Using Boron Trichloride.....	234
Table 8.1	Equilibrium Association Constant for 5'-GTAC-3' Oligomer.....	229

Chapter 9

Figure 9.1	Proposed Binding Modes for Benzothiophene Hairpins.....	241
Figure 9.2	Plasmid Design for Benzothiophene Specificity Study.....	242
Figure 9.3	Solid Phase Synthesis of Benzothiophene Containing Hairpins.....	243
Figure 9.4	DNase I Titrations for Benzothiophene Containing Hairpins (Bt)H & (Bt)Me.....	244
Figure 9.5	DNase I Titrations for Benzothiophene Containing Hairpins	

	(Bt)Cl & (Bt)FCl.....	245
Table 9.1	Equilibrium Association Constants for Hairpins Containing N-Terminal Benzothiophene Caps.....	248
Chapter 10		
Figure 10.1	5-Membered Carboxamides vs. Fused Benzimidazole Derivatives.....	259
Figure 10.2	Ball and Stick Models of 4-G Oligomers in the Context of Their Target DNA Sequences.....	260
Figure 10.3	Chemical Structures of 4-G Oligomers.....	261
Figure 10.4	Synthesis of Imidazole-Imidazopyridine (Im-Ip) Dimers.....	262
Figure 10.5	Solid Phase Synthesis of 4-G Oligomers.....	264
Figure 10.6	Sequence of Plasmid pEF16.....	265
Figure 10.7	DNase I Titrations for 4-G Control Compounds.....	266
Figure 10.8	DNase I Titrations for 4-G Oligomers.....	267
Figure 10.9	Molecular Modeling of 4-G Oligomer Curvature.....	268
Figure 10.10	Modeled Structures of A-Form & B-Form DNA.....	272
Table 10.1	Equilibrium Association Constants for Oligomers That Target G-Tetrads.....	270
Chapter 11		
Figure 11.1	Diagram of Hairpin Tail, Turn & Core Structure.....	292
Figure 11.2	Solid Phase Synthesis of Standard Py-Im Hairpin on	

	Hydrazine Safety Catch Resin.....	293
Figure 11.3	Synthesis of Hairpins Containing Thiophene/Pyrazole Ring Pairs On Hydrazine Safety Catch Resin.....	294
Figure 11.4	Hairpin Tail Derivatives.....	295
Figure 11.5	Yields for Cleavage from Hydrazine Resin.....	297
 Chapter 12		
Figure 12.1	Collection of 5-Membered Heterocyclic Carboxamides.....	308
Figure 12.2	Synthesis of O-Oxazole (Oo).....	309
Figure 12.3	Synthesis of N-Oxazole (No).....	309
Figure 12.4	Mechanism of Oxazole Fragmentation.....	310
Figure 12.5	Determination of Oxazole Carboxamide Cleavage by H ¹ NMR.....	311
Figure 12.6	Comparison of Hydroxypyrrole (Hp) and Hydroxyisoxazole (Hi).....	312
Figure 12.7	Electronic and Geometric Profile of 5-Membered Heterocycles.....	312
Figure 12.8	General Synthesis for Hydroxyisoxazole (Hi).....	313
Figure 12.9	Alkylation of Hydroxyisoxazole Aryl Ether.....	314
Figure 12.10	Differentiation of Hydroxyisoxazole Carboxy Esters.....	314
Figure 12.11	Separation of Hydroxyisoxazole Isomers by HPLC.....	315
Figure 12.12	Isopotential Surface for Hydroxyisoxazole (Hi).....	316